Investigation of the metabolism of rare nucleotides in plants

Von der Naturwissenschaftlichen Fakultät der

Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

genehmigte Dissertation

von

Henryk Straube, **M. Sc.**

2023

Referent: Prof. Dr. Claus-Peter Witte

Korreferent: Prof. Dr. rer. nat. Jakob Franke

Korreferent: Prof. Dr. Alain Tissier

Tag der Promotion: 09.09.2022

Contributing publications

The following publications contributed to this thesis:

1. **Straube, H.**; Niehaus, M.; Zwittian, S.; Witte, C.-P.; Herde, M. (2020): Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. *The Plant Cell,* Volume 33, Issue 2, February 2021, Pages 270–289. DOI: 10.1093/plcell/koaa028

2. **Straube, H.**; Herde, M. (2022): Purification and analysis of Nucleotides and Nucleosides from Plants.
Methods in Molecular Biology, In: Affinity Chromatography: Methods and Protocols. New York, NY: Springer US, May 2022, Pages 145-155. DOI: 10.1007/978-1-0716-2176-9_11

3. **Straube, H.**; Witte, C.-P.; Herde, M. (2021): Analysis of nucleosides and nucleotides in plants: An update on sample preparation and LC–MS techniques. *Cells,* Volume 10, Issue 3, March 2021, Article number 689. DOI: 10.3390/cells10030689

4. Niehaus., M.; **Straube, H.**; Künzler, P.; Rugen, N.; Hegermann, J.; Giavalisco, P.; Eubel, H.; Witte, C.-P.; Herde, M. (2020): Rapid affinity purification of tagged plant mitochondria (Mito-AP) for metabolome and proteome analyses. *Plant Physiology,* Volume 182, Issue 3, March 2020, Pages 1194–1210 DOI: 10.1104/pp.19.00736

5. Niehaus, M.; **Straube, H.**; Specht, A.; Baccolini, C.; Witte, C.-P.; Herde, M. (2022): The nucleotide metabolome of germinating Arabidopsis thaliana reveals a central role of thymidine for chloroplast development. *The Plant Cell*, Volume 34, Issue 10, Otober 2022, Pages 3790-3813. DOI: 10.1093/plcell/koac207

6. Heinemann, K.J.; Yang, S.-Y.; **Straube, H.**; Medina-Escobar, N.; Varbanova-Herde, M.; Herde, M.; Rhee, S.; Witte, C.-P. (2021): Initiation of cytosolic plant purine nucleotide catabolism involves a monospecific xanthosine monophosphate phosphatase. *Nature Communications,* Volume 12, Article number 6846 DOI: 10.1038/s41467-021-27152-4

7. **Straube, H.**; Straube, J.; Rinne, J.; Niehaus, M.; Rose, H.; Witte, C.-P.; Herde, M. (2022): An inosine triphosphate pyrophosphatase safeguards nucleic acids from aberrant purine nucleotides. *New Phytologist*, Volume 237, Issue 5, December 2022, Pages 1759-1775. DOI: 10.1111/nph.18656

Abstract

Nucleotides are metabolites involved in primary metabolism, and specialized metabolism and have a regulatory role in various biochemical reactions in all forms of life. While in other organisms, the nucleotide metabolome was characterized extensively, comparatively little is known about the cellular concentrations of nucleotides in plants.

The aim of this dissertation was to investigate the nucleotide metabolome and enzymes influencing the composition and quantities of nucleotides in plants. For this purpose, a method for the analysis of nucleotides and nucleosides in plants and algae was developed (**Chapter 2.1**), which comprises efficient quenching of enzymatic activity, liquid-liquid extraction and solid phase extraction employing a weak-anion-exchange resin. This method allowed the analysis of the nucleotide metabolome of plants in great depth including the quantification of low abundant deoxyribonucleotides and deoxyribonucleosides. The details of the method were summarized in an article, serving as a laboratory protocol (**Chapter 2.2**).

Furthermore, we contributed a review article (**Chapter 2.3**) that summarizes the literature about nucleotide analysis and recent technological advances with a focus on plants and factors influencing and hindering the analysis of nucleotides in plants, i.e., a complex metabolic matrix, highly stable phosphatases and physicochemical properties of nucleotides.

To analyze the sub-cellular concentrations of metabolites, a protocol for the rapid isolation of highly pure mitochondria utilizing affinity chromatography was developed (**Chapter 2.4**).

The method for the purification of nucleotides furthermore contributed to the comprehensive analysis of the nucleotide metabolome in germinating seeds and in establishing seedlings of *A. thaliana*, with a focus on genes involved in the synthesis of thymidilates (**Chapter 2.5**) and the characterization of a novel enzyme of purine nucleotide degradation, the XANTHOSINE MONOPHOSPHATE PHOSPHATASE (**Chapter 2.6**).

Protein homology analysis comparing *A. thaliana*, *S. cerevisiae,* and *H. sapiens* led to the identification and characterization of an enzyme involved in the metabolite damage repair system of plants, the INOSINE TRIPHOSPHATE PYROPHOSPHATASE (**Chapter 2.7**). It was shown that this enzyme dephosphorylates deaminated purine nucleotide triphosphates and thus prevents their incorporation into nucleic acids. Loss-of-function mutants senesce early and have a constitutively increased content of salicylic acid. Also, the source of deaminated purine nucleotides in plants was investigated and it was shown that abiotic factors contribute to nucleotide damage.

Keywords: Plant nucleotide metabolism, nucleotide analysis, metabolite damage

Zusammenfassung

Nukleotide sind Metaboliten, die am Primärstoffwechsel und an spezialisierten Stoffwechselvorgängen beteiligt sind und eine regulierende Rolle bei verschiedenen biochemischen Reaktionen in allen Lebensformen spielen. Während bei anderen Organismen das Nukleotidmetabolom umfassend charakterisiert wurde, ist in Pflanzen vergleichsweise wenig über die zellulären Konzentrationen von Nukleotiden bekannt.

Ziel dieser Dissertation war es, das Nukleotidmetabolom und die Enzyme zu untersuchen, die die Zusammensetzung und Menge der Nukleotide in Pflanzen beeinflussen. Zu diesem Zweck wurde eine Methode zur Analyse von Nukleotiden und Nukleosiden in Pflanzen und Algen entwickelt (**Kapitel 2.1**), die ein effizientes Stoppen enzymatischer Aktivität, eine Flüssig-Flüssig-Extraktion und eine Festphasenextraktion unter Verwendung eines schwachen Ionenaustauschers umfasst. Mit dieser Methode konnte das Nukleotidmetabolom von Pflanzen eingehend analysiert werden, einschließlich der Quantifizierung von Desoxyribonukleotiden und Desoxyribonukleosiden mit geringer Abundanz. Die Einzelheiten der Methode wurden in einem Artikel zusammengefasst, der als Laborprotokoll dient (**Kapitel 2.2**).

Darüber hinaus wurde ein Übersichtsartikel (**Kapitel 2.3**) verfasst, der die Literatur über die Analyse von Nukleotiden und die jüngsten technologischen Fortschritte zusammenfasst. Der Schwerpunkt lag hierbei auf Pflanzen und Faktoren, die die Analyse von Nukleotiden in Pflanzen beeinflussen oder behindern, d. h. eine komplexe Matrix, hochstabile Phosphatasen und physikalisch-chemische Eigenschaften von Nukleotiden.

Um die subzellulären Konzentrationen von Metaboliten zu analysieren, wurde ein Protokoll für die schnelle Isolierung hochreiner Mitochondrien unter Verwendung einer Affinitätschromatographie entwickelt (**Kapitel 2.4**).

Die Methode zur Analyse von Nukleotiden trug außerdem zu einer umfassenden Analyse des Nukleotidmetaboloms in keimenden Samen und in sich etablierenden Keimlingen von *A. thaliana* bei, wobei der Schwerpunkt auf Genen lag, die an der Synthese von Thymidilaten beteiligt sind (**Kapitel 2.5**), sowie zu der Charakterisierung eines neuen Enzyms des Purinnukleotidabbaus, der XANTHOSINE MONOPHOSPHATE PHOSPHATASE (**Kapitel 2.6**).

Eine Proteinhomologieanalyse, die *A. thaliana*, *S. cerevisiae* und *H. sapiens* miteinander verglich führte zur Identifizierung und Charakterisierung eines Enzyms, das an der Reparatur von geschädigten Metaboliten in Pflanzen beteiligt ist, der INOSINE TRIPHOSPHATE PYROPHOSPHATASE (Kapitel 2.7). Es konnte gezeigt werden, dass dieses Enzym desaminierte Purinnukleotidtriphosphate dephosphoryliert und so deren Einbau in Nukleinsäuren verhindert. Funktionsverlustmutanten altern früh und weisen einen konstitutiv erhöhten Gehalt an Salicylsäure auf. Außerdem wurde die Quelle der desaminierten Purinnukleotide in Pflanzen untersucht, und es wurde gezeigt, dass abiotische Faktoren zur Nukleotidschädigung beitragen.

Schlagwörter: Pflanzlicher Nukleotidmetabolismus, Nukleotid-Analyse, geschädigte Metabolite

Abbreviations

Table 1: Explanation of abbreviations and terms used in this thesis

Abbreviation	Explanation
ADP	adenosine diphosphate
AEC	adenylate energy charge
AMP	adenosine 5′-monophosphate
AMPD	adenosine 5′-monophosphate deaminase
AMPK	adenosine 5′-monophosphate kinase
ATC	aspartate carbamoyltransferase
ATP	adenosine 5′-triphosphate
CDA	cytidine deaminase
CDP	cytidine 5′-diphosphate
CMP	cytidine 5′-monophosphate
CMPP	cytidine 5′-monophosphate phosphatase
cpDNA	chloroplast deoxyribonucleic acid
CSN7	COP9 signalosome complex subunit 7
CTP	cytidine 5′-triphosphate
CTPS	cytidine 5′-triphosphate synthase
dA	deoxyadenosine
dATP	deoxyadenosine 5′-triphosphate
DAU	5,6-diaminouracil
dC	deoxycytidine
DCD	deoxycytidine 5′-monophosphate deaminase
dCDP	deoxycytidine 5′-diphosphate
DCM	dichloromethane
dCMP	deoxycytidine 5′-monophosphate
dCTP	deoxycytidine 5′-triphosphate
dG	deoxyguanosine
dGTP	deoxyguanosine 5′-triphosphate
DHOase	dihydroorotate synthase
DHODH	dihydroorotate dehydrogenase
dITP	deoxyinosine 5′-triphosphate
DNA	deoxyribonucleic acid
dNDP	deoxyribonucleoside 5′-diphosphate
DNK	deoxyribonucleoside kinase
dNTP	deoxyribonucleoside 5′-triphosphate
DPYD	dihydropyrimidine dehydrogenase
dT	deoxythymidine
dTDP	deoxythymidine 5′-diphosphate
dTMP	deoxythymidine 5′-monophosphate
dTTP	deoxythymidine 5′-triphosphate
dU	deoxyuridine
dUDP	deoxyuridine 5′-diphosphate
dUMP	deoxyuridine 5′-monophosphate
DUT1	deoxyuridine 5′-triphosphate pyrophosphatase
dUTP	deoxyuridine 5′-triphosphate
GDP	guanosine 5′-diphosphate
GMP	guanosine 5′-monophosphate
GMPK	guanosine 5′-monophosphate kinase
GMPP	guanosine 5′-monophosphate phosphatase

GMPS	guanosine 5′-monophosphate synthetase
GSDA	guanosine deaminase
GTP	guanosine 5′-triphosphate
HAM1	hydroxylaminopurine mutagenesis 1
HTP	4-hydroxy-2,5,6-triaminopyrimidine
IDP	inosine 5′-diphosphate
IMP	inosine 5′-monophosphate
IMPDH	inosine 5′-monophosphate dehydrogenase
ITPA	inosine 5′-triphosphate pyrophosphatase
LC-MS	liquid chromatography-mass spectrometry
MeOH	methanol
mtDNA	mitochondrial deoxyribonucleic acid
ncDNA	nuclear deoxyribonucleic acid
NDP	nucleoside 5′-diphosphate
NDPK	nucleoside 5′-diphosphate kinase
NSH 1 and 2	nucleoside hydrolase 1 and 2
NUDIX1	8-oxo-deoxyguanosine 5′-triphosphate pyrophosphatase
NH_4Ac	ammonium acetate
PRPP	5′-phosphoribosyl-1′-pyrophosphate
R1	large subunit of the ribonucleotide reductase
R2	small subunit of the ribonucleotide reductase
RNA	ribonucleic acid
RNR	ribonucleotide reductase
SA	salicylic acid
TCA	trichloroacetic acid
TK1a and b	deoxythymidine kinase 1a and b
TMPK	deoxythymidine 5′-monophosphate kinase
TOA	trioctylamine
TOR	target of rapamycin
UDP	uridine 5′-diphosphate
UMP	uridine 5′-monophosphate
UMPK	uridine 5′-monophosphate kinase
UMPP	uridine 5′-monophosphate phosphatase
UMPS	uridine 5′-monophosphate synthase
UTP	uridine 5′-triphosphate
XDH	xanthine dehydrogenase
XMP	xanthosine 5′-monophosphate
XMPP	xanthosine 5′-monophosphate phosphatase
XTP	xanthosine 5′-triphosphate

Table of contents

1. Introduction

Nucleotides are fundamental components of metabolism in all forms of life. Nucleotide metabolism is a vital process that affects primary and specialized metabolism, gene expression and general regulation of biochemical processes in the cell. Most prominently, nucleotides serve as building blocks for the DNA in the nucleus and DNA-synthesizing organelles, mitochondria, and chloroplasts, and convey the genetic information in the form of transcripts, the RNA. Nucleotides are an essential part of protein synthesis in the form of ribosomal and transfer RNAs. They serve as direct precursors for the synthesis of several vitamins and coenzymes, such as thiamin, riboflavin, folates, nicotinamide adenine dinucleotide, flavin adenine dinucleotide, and S-adenosylmethionine and partake in energy metabolism by conserving chemical energy as adenosine triphosphate or guanosine triphosphate. Nucleotide sugars are precursors for the biosynthesis of polysaccharides like starch or cellulose, but also for the synthesis of sucrose, glycolipids and for protein glycosylation (Zrenner et al., 2006; Ashihara et al., 2020; Witte and Herde, 2020).

Nucleotide metabolism in plants can generally be divided into three main branches: nucleotide *de novo* biosynthesis, nucleotide degradation and the salvage of nucleosides and nucleobases (Zrenner et al., 2006; Witte and Herde, 2020). The molecular aspects of nucleotide *de novo* biosynthesis and nucleotide degradation will be discussed later in more detail. Additionally, nucleotide metabolism in plants is connected to specialized metabolism, resulting in the synthesis of purine alkaloids like caffeine or theobromine and pyridine alkaloids like nicotine or trigonelline (Ashihara et al., 2020), as well as phytohormone metabolism, i.e., cytokinin biosynthesis (Zrenner et al., 2006; Ashihara et al., 2020; Witte and Herde, 2020).

Nucleotides are defined as molecules that consist of a nucleobase, a heterocyclic ring structure containing nitrogen atoms (Figure 1A), that is connected via a glycosidic bond to a phosphorylated pentose, which is in most cases a deoxyribose or ribose. The degree of phosphorylation and position of the phosphoryl groups can vary and frequently occur at the 5'-OH group of the sugar (Figure 1B, Witte and Herde, 2020). Nucleosides are structured like nucleotides but do not contain phosphate groups. The nucleobases can be grouped in purines, like adenine and guanine, and pyrimidines, like cytosine, thymine and uracil. Nucleobases can be modified, either by enzymatic reactions, such as methylation, or as a consequence of spontaneous chemical reactions leading to damaged nucleosides and nucleotides (Galperin et al., 2006;

1. Introduction

Lerma-Ortiz et al., 2016; Witte and Herde, 2020; Griffith et al., 2021). These metabolites can be incorporated into nucleic acids, potentially causing detrimental effects.
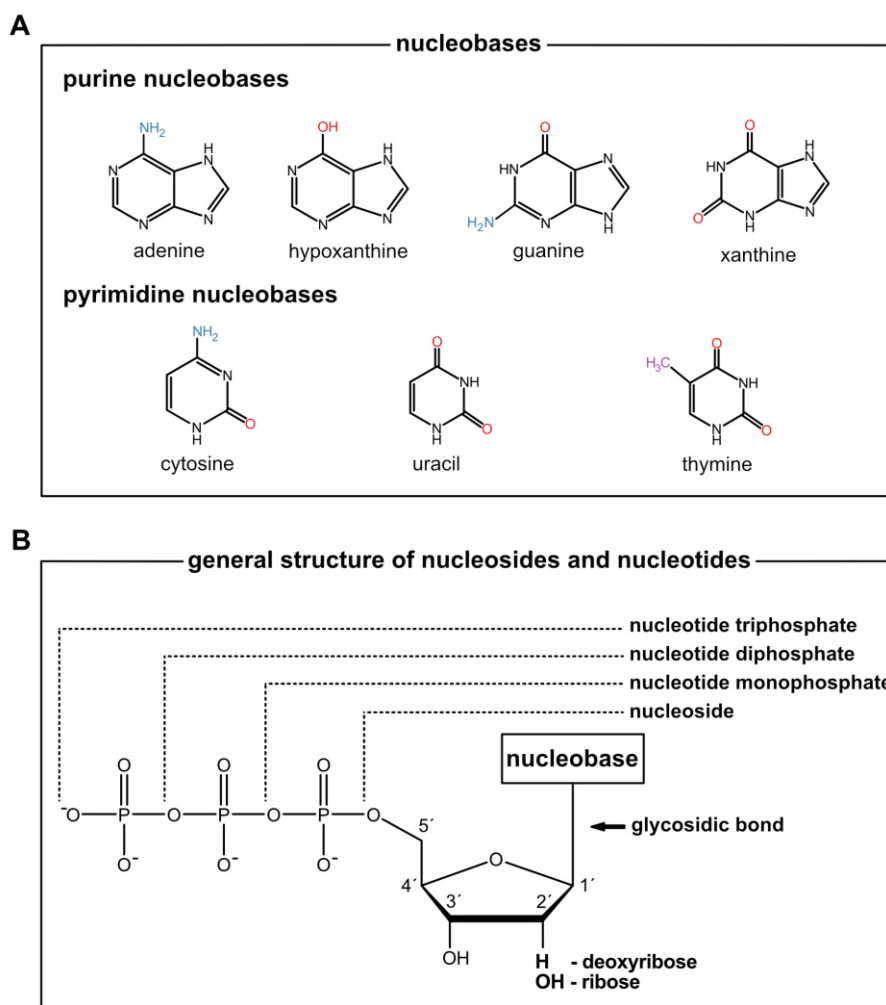

Figure 1: Structure of nucleobases, nucleosides and nucleotides. A, purine and pyrimidine nucleobases that occur as part of canonical nucleotides in plants. Groups attached to the heterocycle are shown in shown in blue (amino groups), red (oxo- or hydroxyl groups) or purple (methyl groups). Nucleobases are shown in their neutral form. **B**, general structure of nucleosides and nucleotides. A nucleoside is a nucleobase that is linked to a pentose moiety via a glycosidic bond. Naturally occurring nucleosides in plants contain a ribose or a deoxyribose as pentose. Nucleotides are nucleosides that contain usually up to three phosphate moieties at the hydroxyl group of the 5´carbon of the respective sugar. Phosphorylation at other hydroxyl groups is possible but not shown here.

While the nucleotide metabolism in mammals is well studied, given its importance for cancer and autoimmune diseases, the nucleotide metabolome in plants is largely uncharacterized. One reason for this is likely the lack of an adequate method to detect and quantify nucleotides and nucleosides in plants. Plants possess a more complex metabolome compared to mammals, creating so-called matrix effects that greatly reduce the sensitivity of the detection of nucleotides and nucleosides in mass spectrometry (Dixon, 2003; Fang et al., 2019; Wang et al., 2019). In addition, plants

1. Introduction

have a plethora of specific and nonspecific phosphatases that are dephosphorylating nucleotides during the extraction process if not properly quenched (Straube et al., 2021a; Straube et al., 2021b; Straube and Herde, 2022). To investigate the nucleotide metabolome of plants, we developed a sensitive method for the analysis of nucleotides and nucleosides from several plant species and algae (Figure 2A-C).
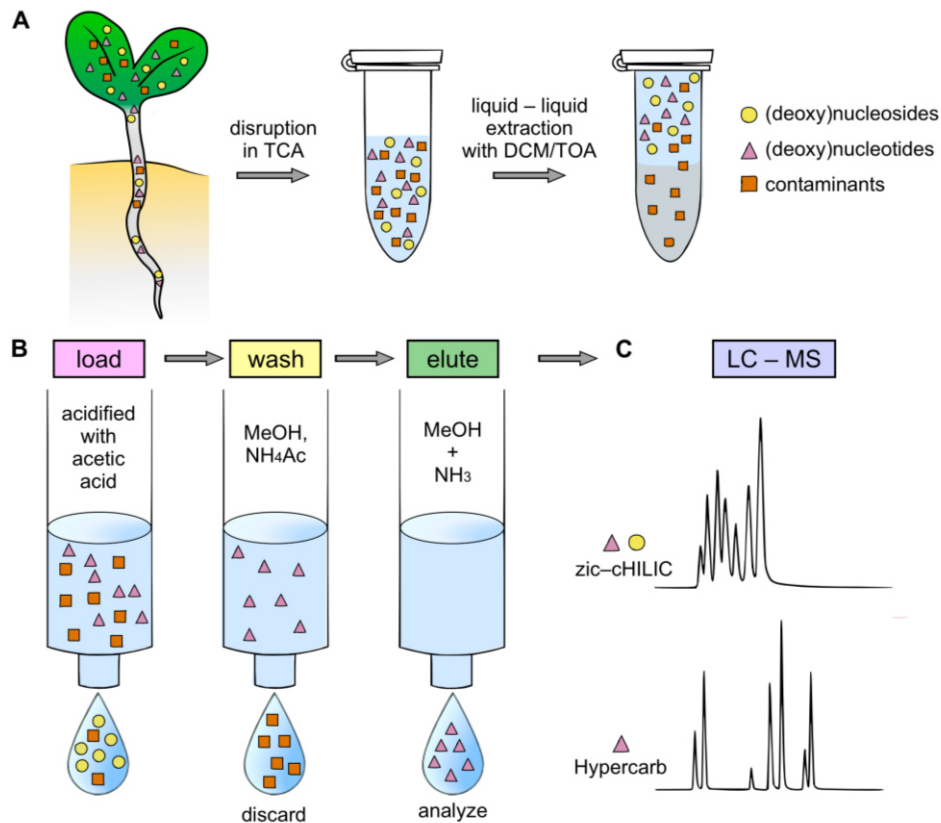

Figure 2: Schematic overview of the method for the extraction and analysis of nucleotides and nucleosides. A, plant material is disrupted and quenched with trichloroacetic acid (TCA), which is removed together with apolar contaminants by liquid-liquid extraction with dichloromethane (DCM) and trioctylamine (TOA). **B**, the extract is loaded onto a weak-anion exchange SPE cartridge, the flow-through contains (deoxy)ribonucleosides for analysis. Subsequently, contaminants are depleted by washing with methanol (MeOH) and ammonium acetate (NH$_4$AC), resulting in the elution of (deoxy)ribonucleotides with MeOH and ammonia (NH$_3$). **C**, isolated fractions are analyzed by liquid chromatography-mass spectrometry (LC-MS) using a zic-cHILIC column (nucleotides and nucleosides) or a Hypercarb column (nucleotides). Reproduced and modified from Straube et al., 2021a (www.academic.oup.com/plcell, Copyright American Society of Plant Biologists).

It provided for the first time a comprehensive overview of the plant nucleotide metabolome, including low abundant ribonucleotides like xanthosine monophosphate (XMP; Heinemann et al., 2021), deoxyribonucleotides and deoxyribonucleosides (Straube et al., 2021a). Comparative genomics led to the characterization of an enzyme involved in the metabolite damage repair system in plants, the INOSINE TRIPHOSPHATE PYROPHOSPHATASE (ITPA). In the mutants lacking this enzyme

1. Introduction

we were able to detect even less abundant damaged nucleotides like inosine di- and triphosphate (IDP and ITP; Straube et al., 2021a; Straube et al., 2022). The newly developed method was used to analyze the nucleotide metabolome in the context of seed germination with a special focus on genes involved in thymidilate biosynthesis (Niehaus et al., 2022). As an approach to unravel the molecular mechanisms involved in the compartmentalization of nucleotide metabolism, a method was developed for rapid isolation of mitochondria from plant tissues using affinity chromatography, paving the way for the comparison of relative concentrations of mitochondrial metabolites in different genotypes or differentially treated plants (Niehaus et al., 2020).

The following chapters describe purine and pyrimidine biosynthesis and degradation in plants, how disturbances in nucleotide metabolism affect plants and what is currently known about metabolite damage in the context of plants and nucleotides. In all chapters, the research contributions of the dissertation will be shortly summarized and put into perspective, discussing the metabolism of nucleotides in plants with a focus on rare nucleotides.

1.1 Nucleotide *de novo* biosynthesis

Biosynthetic pathways forming purine and pyrimidine nucleotides are utilizing precursor molecules from central metabolism like 5-phosphoribosyl-1-pyrophosphate (PRPP), glutamine, aspartate, glycine, bicarbonate, and formyl tetrahydrofolate to generate ribonucleotide monophosphates. These biosynthetic pathways involve several enzymes that are localized in the chloroplasts, mitochondria, and cytosol, and also require transport processes between these compartments.

1.1.1 Purine nucleotide *de novo* biosynthesis

Purine nucleotide *de* novo biosynthesis results in the formation of adenosine monophosphate (AMP) by 12 reactions that are catalyzed by 11 enzymes. From AMP guanosine monophosphate (GMP) is made (Figure 3A; Smith and Atkins, 2002; Ashihara et al., 2020; Witte and Herde, 2020).

Figure 3: Purine *de novo* biosynthesis and subsequent phosphorylation of purine nucleotides. A, purine *de* novo biosynthesis. **B**, phosphorylation of purine nucleotide monophosphates to their respective nucleotide triphosphates. Involved enzymes include: AMP KINASEs (AMPKs), NUCLEOSIDE DIPHOSPHATE KINASEs (NDPKs), AMP DEAMINASE (AMPD), IMP DEHYDROGENASE (IMPDH), GMP SYNTHETASE (GMPS) and GMP KINASEs (GMPKs). The subcellular localization of the respective enzymes is indicated by the colored background.

These enzymes are in most plants localized in the chloroplasts (Witte and Herde, 2020). An exception are cowpea nodules, where enzyme activity assays showed that both plastids and mitochondria harbor the enzymes of purine *de novo* biosynthesis

1.1 Nucleotide *de novo* biosynthesis

(Atkins et al., 1997; Smith and Atkins, 2002). AMP is then exported to the cytosol. GMP is synthesized in the cytosol from AMP by three reactions. AMP is converted by an AMP DEAMINASE (AMPD) to inosine monophosphate (IMP), which is then metabolized to xanthosine monophosphate (XMP) by INOSINE MONOPHOSPHATE DEHYDROGENASE (IMPDH). In a final step, guanosine monophosphate (GMP) is synthesized from XMP by a GUANOSINE MONOPHOSPHATE SYNTHETASE (GMPS). GMP can be phosphorylated by guanosine monophosphate kinases (GMPKs) to guanosine diphosphate (GDP) and AMP can be further phosphorylated by AMP KINASES (AMPKs) to adenosine diphosphate (ADP; Figure 3B). These purine nucleotide diphosphates can subsequently be phosphorylated to their respective nucleotide triphosphates, adenosine triphosphate (ATP) and guanosine triphosphate (GTP), by NUCLEOSIDE DIPHOSPHATE KINASEs (NDPKs). The enzymes for these phosphorylation reactions are localized in the mitochondria, the chloroplasts and the cytosol.

Besides being responsible for the initiation of guanylate synthesis, AMPD also contributes to adenylate catabolism and regulation of the adenylate energy charge (AEC; an index that is used to measure the energy status of the cell). Studies showed that AMPD is feedback regulated by GTP *in vitro* (Yabuki and Ashihara, 1991; Sabina et al., 2007). Analyzing mutants with elevated concentrations of GTP, we could provide evidence that AMPD is also inhibited by GTP *in vivo* because concentrations of ATP were increased and those of IMP, the product of AMPD, decreased (Straube et al., 2021a).

1.1.2 Pyrimidine nucleotide *de novo* synthesis

Pyrimidine nucleotide *de* novo synthesis catalyzes the formation of UMP from aspartate and carbamoyl phosphate by four reactions that are catalyzed by four enzymes (Figure 4A; Witte and Herde, 2020). The respective enzymes are localized in the chloroplast, the cytosol, and the intermembrane space of the mitochondrion. The first step of pyrimidine *de novo* synthesis, catalyzed by ASPARTATE TRANSCARBAMOYLASE (ATC) is regulated through feedback inhibition by uridylates, mainly UMP, but not cytidylates (Doremus and Jagendorf, 1985). A recent study confirmed that ATC is uniquely regulated by UMP, binding to a loop structure that is exclusively conserved in plants (Bellin et al., 2021). UMP is phosphorylated subsequently to UTP by URIDINE MONOPHOSPHAT KINASES (UMPKs) and NDPKs

(Figure 4B). Cytidine triphosphate (CTP) is synthesized by CYTIDINE TRIPHOSPHATE SYNTHETASES (CTPS) from UTP, ATP and glutamine (Witte and Herde, 2020).
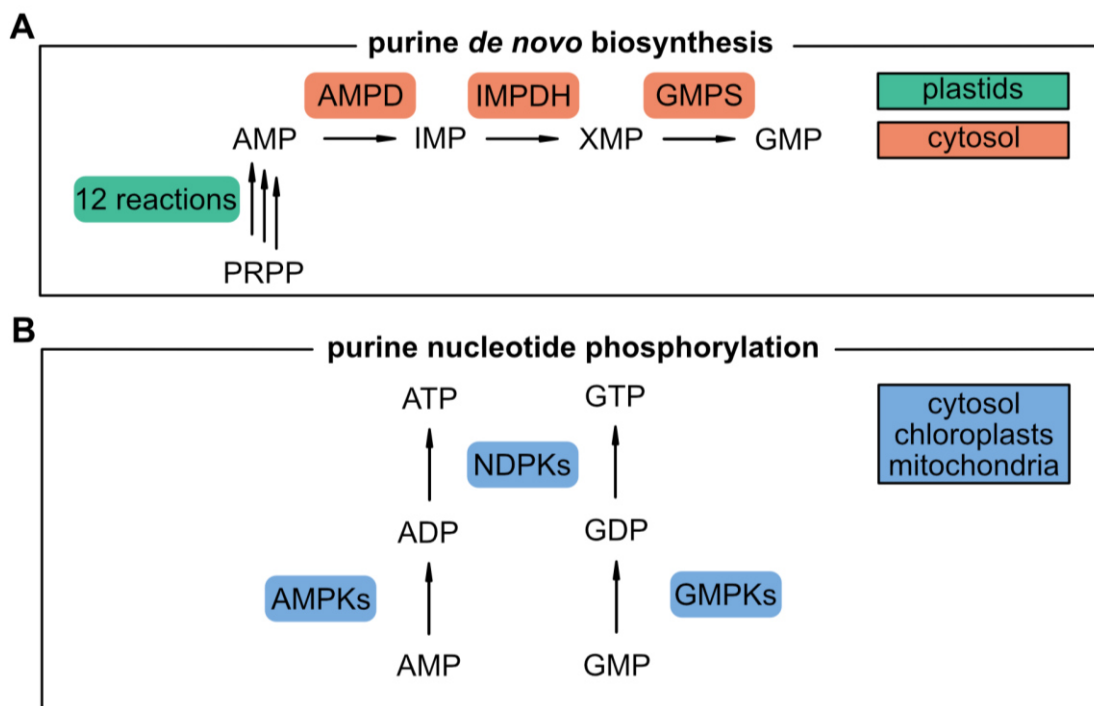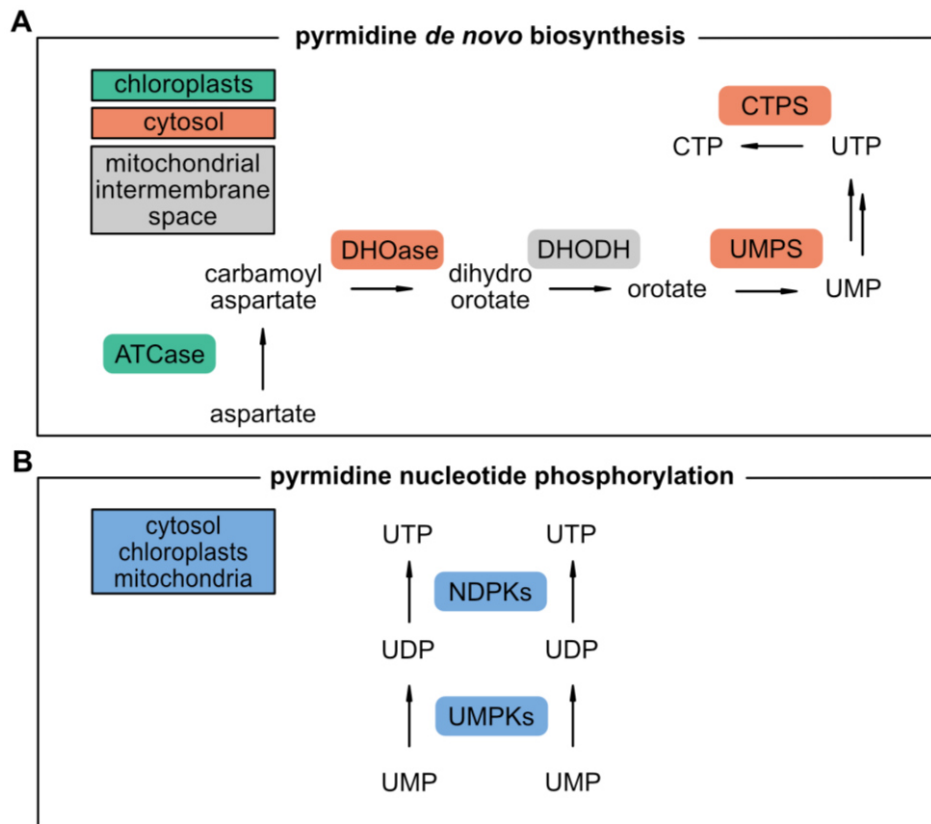

Figure 4. Pyrimidine de novo biosynthesis and subsequent phosphorylation of pyrimidine nucleotides. A, pyrimidine de novo biosynthesis. **B**, phosphorylation of pyrimidine nucleotide monophosphates to their respective nucleotide triphosphates. Enzymes, ASPARTATE TRANSCARBAMOYLASE (ATCase), DIHYDROOROTATE SYNTHASE (DHOase), DIHYDROOROTATE DEHYDROGENASE (DHODH), UMP SYNTHASE (UMPS), UMP KINASES (UMPKS), NUCLEOSIDE DIPHOSPHATE KINASES (NDPKs) and CTP SYNTHASES (CTPS). The subcellular localization of the respective enzymes is indicated by the colored background.

Although in *Arabidopsis thaliana* five enzymes are predicted to have CTPS activity, recent publications showed that CTPS2, which is localized in the cytosol, is most likely the main enzyme for CTP synthesis, as mutations in this gene result in embryo lethality (Alamdari et al., 2021; Hickl et al., 2021).

1.1.3 Deoxyribonucleotide *de novo* synthesis

The *de novo* biosynthesis of deoxyadenosine triphosphate (dATP), deoxycytidine triphosphate (dCTP), deoxyguanosine triphosphate (dGTP) and deoxythymidine triphosphate (dTTP) is dependent on RIBONUCLEOTIDE REDUCTASE (RNR). RNR is a multimeric enzyme complex that catalyzes the reduction of the 2′-hydroxyl moiety on the ribose of ribonucleoside diphosphates (NDPs), producing deoxyribonucleoside

1.1 Nucleotide *de novo* biosynthesis

diphosphates (dNDPs), a rate-limiting step in the synthesis of DNA precursors (Wang and Liu, 2006; Garton et al., 2007, Witte and Herde, 2020). In *A. thaliana,* RNR is a heterocomplex that consists of two large subunits (R1) with regulatory functions and two small subunits (R2) harboring the di-iron tyrosyl radical cofactor for the reduction of NDP to dNDPs. The large subunits of the RNR are encoded by one gene, *RNR1*, while the small subunits are encoded by three genes, *TSO2, RNR2A* and *RNR2B*. A mutation resulting in a loss-of-function in genes encoding R1 is lethal. However, plants carrying *R1* alleles with weak missense mutations are viable but show severe growth defects and chlorosis (Garton et al., 2007; Tang et al., 2019). By contrast, loss-of-function of R2 encoding genes as single mutants are viable. Mutants of *TSO2* show severe developmental abnormalities and white spots on the leaves that are completely lacking chloroplasts, whereas single mutants of *RNR2A* and *RNR2B* show no developmental abnormalities. Double mutants of *TSO2* with either *RNR2A* or *RNR2B* are not viable. The exact stoichiometry of the RNR subunits is currently not known for the plant enzyme. Interestingly, the transcripts of the genes coding for the R2 subunit are differentially expressed in distinct tissues and developmental stages (Wang and Liu, 2006). Furthermore, it was shown that subunit 7 (CSN7) of the CONSTITUTIVE PHOTOMORPHOGENESIS 9 signalosome complex interacts and determines the localization of the R2 subunit of the RNR in *A. thaliana*. Usually, in quiescent cells the R2 subunit is localized in the cytoplasm, but in meristematic and actively dividing cells the R2 subunit was detected mostly in nuclei. Nuclear localization of R2 was also found in cells that respond to DNA damage. In mutants lacking a functional CSN7 the R2 unit was constitutively localized in the nucleus (Halimi et al., 2011).

In general, less is known about how RNR is regulated in plants than in other organisms like yeast or mammals. In those organisms, the R1 subunit has two sites for allosteric regulation, while ATP bound at one of the sites increases the enzymatic activity, dATP decreases the activity. The reduction of NDPs in other organisms by RNR is also allosterically regulated by another site, depending on the concentrations of dATP, dGTP, and dTTP, while dCTP has no effect on the activity (Torrents, 2014). The R1 activity also seems to be dependent on thioredoxin (Sauge-Merle et al., 1999). The products of the NDP reduction via RNR, dNDPs, are phosphorylated to deoxyribonucleoside triphosphates (dNTPs) by NDPKs, analogous to the phosphorylation of NDPs. NDPKs are multisubstrate enzymes and the *in vivo* stoichiometry of NDP phosphorylation is currently unknown (Witte and Herde, 2020).

1.1 Nucleotide *de novo* biosynthesis

Interestingly, we discovered that the ratios of dCTP and dGTP to dATP and dTTP correlate positively with the GC content of DNA in several different plant and algae species (Straube et al., 2021a). Whether this is the result of a differential regulation of RNR in different species or an unknown regulatory mechanism remains to be elucidated.

While the precursor molecules for RNA synthesis are ATP, CTP, GTP and UTP, canonical DNA precursors comprise dATP, dCTP, dGTP and dTTP, but not deoxyuridine triphosphate (dUTP; Wang and Liu, 2006; Witte and Herde, 2020). RNR catalyzes the reduction of uridine diphosphate to deoxyuridine diphosphate (dUDP), which serves as a precursor for dTTP synthesis. Thymidylate synthesis involves several reactions and enzymes in the cytosol, the mitochondria, and the chloroplasts (Figure 5).
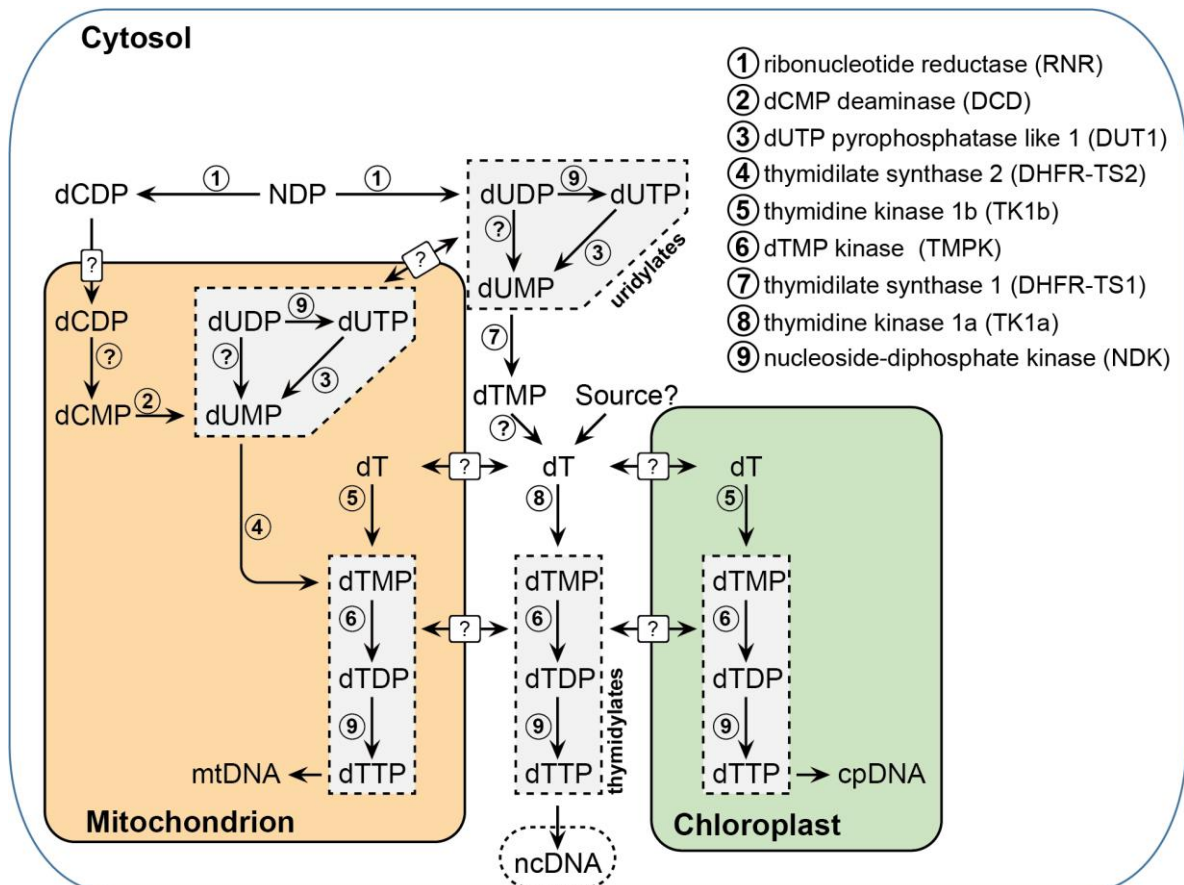
Figure 5: Proposed model of thymidilate metabolism in *Arabidopsis thaliana*. Enzymes putatively involved in the formation of thymidylate pools and their subcellular localization are shown. Question marks indicate currently unknown enzymes or transporters allowing an exchange of nucleosides and nucleotides between the cytosol, the chloroplast and the mitochondrion. Reproduced and modified from Niehaus et al., 2022.

Most likely, dUDP is phosphorylated to dUTP and successively dephosphorylated by the DEOXYURIDINE TRIPHOSPHATE PYROPHOSPHATASE (DUT1) to deoxyuridine monophosphate (dUMP), as DUT1 is strongly regulated by TARGET OF

RAPAMYCINE (TOR), a protein kinase whose activity promotes anabolism (Busche et al., 2021). Alternatively, dUDP is directly dephosphorylated to dUMP by an unknown phosphatase (Witte and Herde, 2020). It is unlikely that dUDP is dephosphorylated by DUT1, as DUT1 belongs to a family of dUTP pyrophosphatases that have no biochemical activity with dUDP (Vértessy and Tóth, 2009). In a second reaction, a methyl group is transferred to dUMP from methylentetrahydrofolate by the enzyme DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE (DHFR-TS). The product of this reaction, deoxythymidine monophosphate (dTMP), is then phosphorylated by a dTMP KINASE (TMPK, previously referred to as ZEU1, see Ronceret et al., 2008) to deoxythymidine diphosphate (dTDP), which is subsequently phosphorylated by NDPKs to dTTP (Witte and Herde 2020). An alternative entrance to dTTP biosynthesis via dUMP is the dephosphorylation of dCTP or deoxycytidine diphosphate (dCDP) to deoxycytidine monophosphate (dCMP) and its following deamination by DEOCYTIDINE MONOPHOSPHATE DEAMINASE (DCD) to dUMP (Xu et al., 2014; Niu et al., 2017). Interestingly, the enzymes involved in the synthesis of dTTP are located in several different compartments: RNR resides in the cytosol, DUT1 is localized in the cytosol, the chloroplasts and the mitochondria, DCD is present in the mitochondria, DHFR-TS can be found in the cytosol and mitochondria, TMPK is localized in the cytosol, chloroplasts and mitochondria and NDPKs are localized in the cytosol, mitochondria and chloroplasts (Figure 5, Xu et al., 2014; Witte and Herde, 2020; Niehaus et al., 2022). Studying a mutant collection of all available mutants in enzymes involved in the synthesis of dTTP, we could show that especially organellar enzymes are important in early seedling establishment, whereas in older seedlings cytosolic enzymes can compensate for a loss of organellar enzymes involved in dTTP biosynthesis (Niehaus et al., 2022). A potential regulation of enzymes involved in dTTP biosynthesis remains unknown.

1.1.4 Regulation of nucleotide *de novo* biosynthesis

Previous reports state that genes involved in pyrimidine nucleotide *de novo* synthesis are not strongly transcriptionally regulated in contrast to transcripts of genes involved in purine nucleotide *de novo* synthesis, besides changes in transcript abundance during different developmental stages (Zrenner et al., 2006). This is contrary to what was shown recently. Evidence suggests that TOR coordinates purine and pyrimidine nucleotide *de novo* biosynthesis and, conversely, is itself regulated by the availability of nucleotides. The same study also showed that TOR has an important function in

1.1 Nucleotide *de novo* biosynthesis

harmonizing nucleotide availability and ribosome biogenesis (Busche et al., 2021). Furthermore, transcriptomic profiling revealed that especially during germination, transcripts related to nucleotide metabolism are differentially regulated (Narsai et al., 2011; Law et al., 2012). Analyzing the abundances of nucleotides and nucleosides during a time course from dry seeds to established seedlings, we could show that indeed the concentrations of these metabolites vary strongly during germination. It was also shown that nucleotides are synthesized very early in germination, before the onset of DNA replication. During seedling establishment, the rapid increase of deoxyribonucleotide triphosphates (dNTPs) concentrations correlated with the onset of chloroplast DNA replication (Niehaus et al., 2022).

1.2 Salvage of deoxyribonucleosides

Salvage of nucleobases and nucleosides is a general concept in nucleotide metabolism that also applies to ribonucleosides. This chapter focuses on what is known about the salvage of deoxyribonucleosides. Salvage reactions are thought to be involved in energy conservation by converting nucleosides and nucleobases to their respective nucleotide monophosphates. The nucleosides or nucleobases may originate from the excision repair of nucleic acids, are the products of nucleic acid degradation, the aberrant dephosphorylation of nucleotides, or may be taken up from the environment (Witte and Herde, 2020).

In plants, two enzymes are known that are specifically involved in the salvage of deoxyribonucleosides, DEOXYNUCLEOSIDE KINASE (DNK) and THYMIDINE KINASE (TK; Clausen et al., 2012; Xu et al., 2015; Witte and Herde, 2020). *In vitro* studies showed that DNK has a kinase activity for deoxyadenosine (dA), deoxycytidine (dC), and deoxyuanosine (dG), but not for deoxythymidine (dT). In *A. thaliana*, TK has two isoforms that are encoded by two genes, *TK1A* and *TK1B*. Both enzymes have an enzymatic activity with dT and deoxyuridine (dU), while having no activity with any other naturally occurring deoxyribonucleoside. Loss-of-function mutants in *TK1A* show no phenotypic abnormalities under normal growth conditions, whereas *TK1B* mutants have chlorotic cotyledons (Xu et al., 2015). A mutation in *TK1A* and *TK1B* is lethal and can only be partially rescued when a carbon source is applied exogenously (Clausen et al., 2012; Xu et al., 2015). Mutants lacking a functional copy of *DNK* are not described, but a loss-of-function of this gene is likely lethal.

The localization for TK1A and TK1B was clearly shown, TK1A is localized in the cytosol, while TK1B is localized in mitochondria and plastids (Xu et al., 2015). The localization of DNK is unclear, but investigations of deoxyribonucleoside kinase activities in different subcellular fractions of *A. thaliana* and *S. tuberosum* showed that DNK activity is exclusively observed in mitochondria (Clausen et al., 2014). We analyzed the nucleotide and nucleoside concentrations in mutants of *TK1A* and *TK1B* during germination and could show that TK1B has an influence on the concentrations of all dNTPs except dTTP. The effect of the abolished TK1B function seems to be counterintuitive at first, but we hypothesized that the replication of chloroplast DNA (cpDNA) is severely slowed down in these mutants as there is not enough dTTP provided, whereas the other dNTPs are continuously synthesized but not consumed by the polymerase. Indeed, when analyzing the content of cpDNA, *TK1B* mutants

1.2 Salvage of deoxyribonucleosides

harbor significantly less cpDNA than wild type plants, further indicated by chlorotic cotyledons. These results provide evidence that TK1B is crucial for the synthesis of dTTP from dT, fueling cpDNA replication during seedling establishment. Analysis of the deoxyribonucleoside concentrations in germinating seeds furthermore revealed a continuously increasing concentration of dT until seedling establishment, while deoxyadenosine, deoxycytidine, and deoxyguanosine concentrations were either so low in concentration that they could not be detected or in the case of deoxycytidine the abundance did not change during germination and seedling establishment. Taken together with the fact that an abolished function of TK1A and TK1B is lethal and that the concentration of dT is comparatively high during germination, it is tempting to speculate that dT is the main precursor for dTTP synthesis during germination (Niehaus et al., 2022). However, the source for dT in plants remains puzzling. The high concentration of the dT in germinating seeds and its central importance for dTTP synthesis suggest a role for salvage in plant physiology that goes beyond the mere conservation of energy.

1.3 Nucleotide catabolism

While nucleosides and nucleobases can be salvaged to their respective nucleotides, plants possess, in contrast to many other organisms, the ability to fully degrade purine and pyrimidine bases (Figure 6A and B; Werner and Witte, 2011; Witte and Herde, 2020).
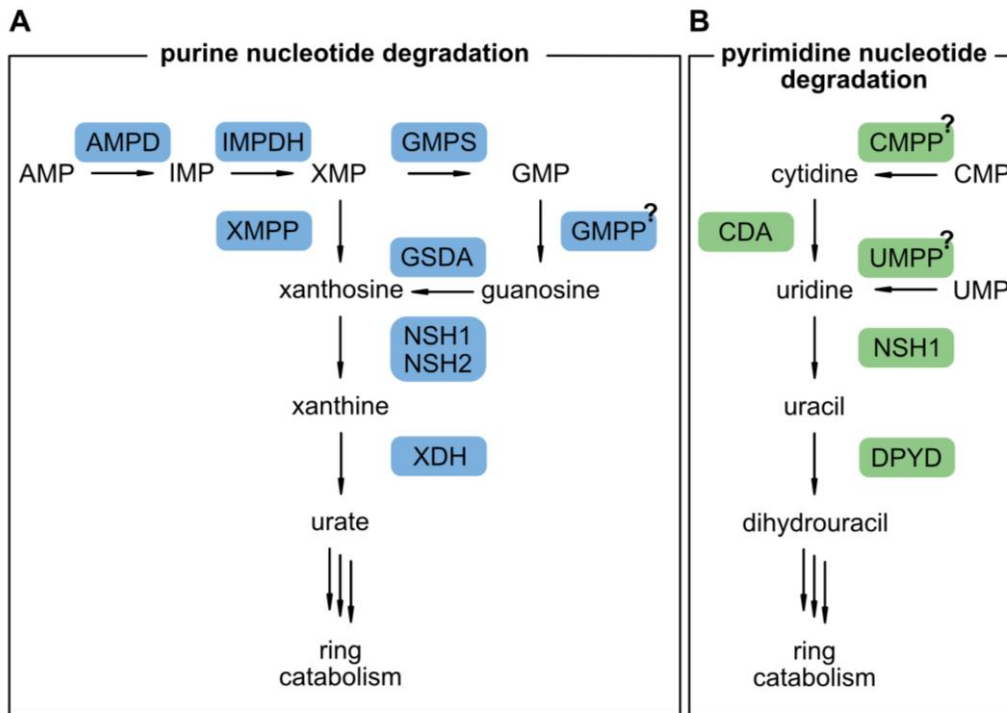

Figure 6: Overview of the plant nucleotide catabolism. A, enzymes involved in the purine nucleotide degradation include: AMP DEAMINASE (AMPD), IMP DEHYDROGENASE (IMPDH), GMP SYNTHETASE (GMPS), XANTHOSINE MONOPHOSPHATE PHOSPHATASE (XMPP), GUANOSINE MONOPHOSPHATE PHOSPHATASE (GMPP), NUCLEOSIDE HYDROLASES 1 and 2 (NSH1 and NSH2) and XANTHINE DEHYDROGENASE (XDH). **B,** enzymes involved in the pyrimidine nucleotide degradation include: CMP PHOSPHATASE (CMPP), UMP MONOPHOSPHATASE (UMPP), CYTIDINE DEAMINASE (CDA), NUCLEOSIDE HYDROLASE 1 (NSH1) and DIHYDROPYRIMIDINE DEHYDROGENASE (DPYD). A question mark indicates that the respective gene for this function is currently not known in plants.

These reactions recycle the nitrogen that is contained in the heterocyclic nucleobases (Werner and Witte, 2011; Soltabayeva et al., 2018). Intermediates of purine degradation like allantoin and uric acid are also involved in alleviating oxidative stress (Brychkova et al., 2008; Watanabe et al., 2014; Ma et al., 2016; Witte and Herde, 2020), whereas intermediates of pyrimidine degradation are thought to be utilized by enzymes of specialized metabolism, leading to the synthesis of specialized compounds like lathyrine, willardiine or 5-aminouracil (Kafer et al., 2004). Biochemical studies suggest that not only ribonucleosides are degraded by enzymes of the nucleotide degradation pathway, but also deoxyribonucleosides and modified

nucleosides (Vincenzetti et al., 1999; Jung et al., 2011; Dahncke and Witte, 2013; Chen et al., 2016).

Purine nucleotide catabolism is initiated with the dephosphorylation of GMP by an unknown GMP PHOSPHATASE. The product of this reaction, guanosine, is then further metabolized by a GUANOSINE DEAMINASE (GSDA) to xanthosine (Dahncke and Witte, 2013). Xanthosine is hydrolyzed by a heterocomplex of NUCLEOSIDE HYDROLASE 1 and 2 (NSH1/ 2), resulting in xanthine and ribose (Jung et al., 2011; Baccolini and Witte, 2019). Xanthine is oxidized by the XANTHINE DEHYDROGENASE (XDH) to urate (Hesberg et al., 2004; Urarte et al., 2015) which is further degraded by several enzymatic steps that release the nitrogen in form of ammonia from the nucleobase and result in glyoxylate (Werner and Witte, 2011). It was previously thought possible that in plants all purine nucleotides have to be converted to GMP to be degraded (Witte and Herde, 2020). Recently, a highly specific cytosolic enzyme that dephosphorylates xanthosine monophosphate (XMP) to xanthosine was discovered, the XANTHOSINE MONOPHOSPHATE PHOSPHATASE (XMPP), which is exclusive to plant nucleotide catabolism (Heinemann et al. 2021). Employing our new method for the analysis of nucleotides, we could show that loss-of-function mutants of *XMPP* accumulated XMP, even more so in the background of a *NSH1* mutation, proving that XMPP has a significant role in the dephosphorylation of XMP *in vivo*. The reason for a second entry point to purine nucleotide catabolism and its physiological relevance remain unknown, but it is speculated that XMPP plays a major role in the degradation of adenylates, whereas the purine degradation via GSDA is more important for the degradation of guanylates, keeping the guanylate and adenylate concentrations in balance. Inosine and hypoxanthine are no major intermediates of purine catabolism, they are probably derived from modified RNA turnover or damaged nucleobases resulting from DNA excision repair (Baccolini and Witte, 2019; Witte and Herde, 2020). Inosine is hydrolyzed to hypoxanthine by the NSH1/ 2 heteromer and oxidized to xanthine by the XDH (Urarte et al., 2015; Baccolini and Witte, 2019).

Pyrimidine nucleotide catabolism is initiated with the dephosphorylation of UMP and cytidine monophosphate (CMP) by unknown phosphatases, resulting in the corresponding ribonucleosides, uridine and cytidine, respectively. Cytidine is then deaminated by a CYTIDINE DEAMINASE (CDA) to uridine (Faivre-Nitschke et al., 1999; Vincenzetti et al., 1999; Chen et al., 2016). Uridine can be hydrolyzed by NSH1

to uracil and ribose, but in contrast to xanthosine and inosine, a homomer of NSH1 is sufficient for the hydrolysis of uridine (Baccolini and Witte, 2019). Subsequently, uracil is degraded by four enzymatic steps to malonate-semialdehyde (Tintemann et al., 1985; Walsh et al., 2001; Cornelius et al., 2011; Witte and Herde, 2020).

While the molecular details of the degradation of ribonucleosides have been well investigated, it remained elusive how deoxyribonucleosides and modified nucleosides are turned over (Witte and Herde, 2020). Biochemical data suggested that deoxyribonucleosides are degraded by the same enzymes that catabolize ribonucleotides. *In vitro* experiments showed that CDA also deaminates deoxycytidine (Chen et al., 2016) and 5-methylcytidine (Vincenzetti et al., 1999). It was also shown that NSH1 can hydrolyze methyluridine (Jung et al., 2011), making it likely that these are also *in vivo* substrates derived from RNA turnover. It could also be shown that GSDA deaminates deoxyguanosine *in vitro* (Dahncke and Witte, 2013). We analyzed mutants of *GSDA*, *NSH1* and *CDA* to examine whether deoxyribonucleosides are substrates for these enzymes *in vivo* and found that mutants of NSH1 accumulate deoxythymidine and CDA accumulates deoxycytidine, whereas for GSDA conclusive evidence could not be obtained because deoxyguanosine was not detected. Interestingly, the blockage of nucleotide degradation led not only to increased concentrations of the direct substrates of the enzymes but also to increased concentrations of nucleotides. This result suggests that ribo- and deoxyribonucleosides are directly phosphorylated to their respective nucleotides *in vivo*. Current literature concerned with plant nucleotide metabolism suggests a temporal and/or spatial separation of nucleotide synthesis, degradation, and salvage (Ashihara et al., 2020; Witte and Herde, 2020). Our data provide evidence that metabolic compartments in plants possess all these enzymes at once, likely being strictly regulated under natural conditions. Summarizing our work and recent discoveries in plant nucleotide metabolism, we suggest that the interplay of degradation and synthesis is part of a process that purges the direct precursor pools for the nucleic acid synthesis of aberrant nucleotides. In this process, only canonical metabolites are phosphorylated again to serve as polymerase substrates whereas aberrant intermediates are degraded (Witte and Herde, 2020; Straube et al., 2021a; Straube et al., 2022).

1.4 Effects of disturbed nucleotide metabolism on plant viability

As nucleotides are an integral part of metabolism, disturbances in biosynthesis, salvage, or degradation pathways of nucleotides can have detrimental effects on organisms. In non-plant organisms, it is well established that, for example, the ratios of dNTPs to one another must be fine-tuned, otherwise, polymerase stalling, DNA damage, mutations, and, as a consequence, cell death will result (Kumar et al., 2011; Poli et al., 2012). In plants, on the other hand, comparatively little is known about the consequences of disturbed nucleotide metabolism and the underlying mechanistic principles.

Mutants lacking key enzymes of purine and pyrimidine nucleotide *de novo* biosynthesis are non-viable. It was shown, that loss-of-function mutations in genes involved in deoxyribonucleotide biosynthesis have deleterious consequences for the plant, similar to observations in mammals (Wang and Liu, 2006; Garton et al., 2007; Xu et al., 2015; Alamdari et al., 2021). Even mutants that possess slightly altered dNTP concentrations show drastic phenotypic alterations, when compared to the wild type, like bleached leaves, increased cell death, decreased fertility, and overall stunted growth (Wang and Liu, 2006; Garton et al., 2007), whereas mutants that alter the dNTP concentrations strongly are usually non-viable (Wang and Liu, 2006; Garton et al., 2007; Xu et al., 2015; Hickl et al., 2021).

A mutant of VENOSA4, a homolog of a dNTP triphosphatase from mammals recently discovered in plants, shows also significant phenotypic alterations, bearing chloroplasts with increased size but overall more yellowish leaves and an impairment in photosynthetic traits, likely a result of increased dNTP concentrations (Yoshida et al., 2018; Xu et al., 2020). Mutants with defective nucleotide degradation like *cda* and *gsda* show altered growth and germination rates (Chen et al., 2016; Schroeder et al., 2018). Interestingly, a common observation is that mutants in nucleotide metabolism often exhibit changes in chloroplast size and abundance, and consequently exhibit altered photosynthetic traits (Wang and Liu, 2006; Yoo et al., 2009; Xu et al., 2015; Yoshida et al., 2018; Niehaus et al., 2022). This is likely the consequence of impaired chloroplast DNA replication, as several of these mutants show strongly altered cpDNA to nuclear DNA ratios (Garton et al., 2007; Pedroza-García et al., 2019; Niehaus et al., 2022). During seedling establishment, it even seems as if the dNTP ratios are specifically tailored for cpDNA replication, this is consistent with the observation that cpDNA synthesis occurs prior to nuclear and mitochondrial DNA replication (Niehaus

1.4 Effects of disturbed nucleotide metabolism on plant viability

et al., 2022). The mechanistic reasons why perturbations in nucleotide concentrations have a stronger effect on the chloroplast genome than the nuclear or mitochondrial genome remains to be elucidated.

1.5 Metabolite damage and molecular countermeasures

In addition to the concentrations and the ratios of metabolites to each other, the integrity of the metabolite pools is of great importance for the proper course of metabolism. While the damage to macromolecules like nucleic acids or proteins is a widely recognized phenomenon, the concept of damaged metabolites is a comparatively new one (D'Ari and Casadesús, 1998; Linster et al., 2013). It was suggested that biological systems are more chaotic than is generally assumed and that errors like metabolite damage, resulting from the inherent limitation of accuracy and fidelity of biological systems, are one of the main drivers of evolution (Tawfik, 2010). Metabolites can be damaged by spontaneous non-enzymatic reactions or by enzymatic promiscuity (Galperin et al., 2006; Linster et al., 2013; Lerma-Ortiz et al., 2016; Crécy-Lagard et al., 2018; Griffith et al., 2021). It is suggested, that as many as 5000 metabolites, which could be classified as 'damaged' might exist in organisms (Lerma-Ortiz et al., 2016). Damaged metabolites can be defined as metabolites, on whose formation no positive evolutionary selection pressure is exerted, as they are useless or even toxic, and whose removal represents a selection advantage (van Schaftingen et al., 2013; Griffith et al., 2021). This also distinguishes damaged metabolites from metabolites that can arise from spontaneous, non-enzymatic reactions but are important intermediates of metabolic pathways (Keller et al., 2015). Organisms possess a biochemical machinery that fixes these metabolic accidents, the so-called damage control system, consisting of several metabolite repair enzymes. As organisms contain a core set of metabolites that share similar susceptibility for certain enzymatic or chemical damage, many of the enzymatic functions of the damage control system are highly conserved across the kingdoms of life (Hanson et al., 2016; Straube et al., 2022). Only few of these enzymes have been characterized in plants, like L-2-HYDROXYGLUTERATE DEHYDROGENASE or METHIONINE SULFOXIDE REDUCTASE (Hanson et al., 2016). Albeit the rapidly increasing information about plant genes, many remain uncharacterized. Also, the metabolites behind many peaks in plant metabolomic profiles remain unknown. It is thus speculated, that many of the hitherto uncharacterized genes and features, especially in plants, are involved in metabolite damage repair (Hanson et al., 2016).

A group of metabolites that is prone to be damaged are nucleotides (Galperin et al., 2006). Usually, the form of aberrant nucleotides that are most toxic are the nucleotide triphosphates, as they serve as direct precursors for nucleic acid synthesis. As such,

1.5 Metabolite damage and molecular countermeasures

the repair system of nucleotides often involves pyrophosphatases that dephosphorylate the nucleotide triphosphates to their respective monophosphates. While in other organisms several examples of these enzymes are characterized, comparatively little is known about these pyrophosphatases in plants (Hanson et al., 2016). In plants, only a few enzymes involved in degrading damaged nucleotides have been characterized, like the DEOXYURIDINE TRIPHOSPHATE PYROPHOSPHATASE (DUT1; Dubois et al., 2011; Niehaus et al., 2022) the 8-OXO DEOXYGUANOSINE TRIPHOSPHATE PRYOPHOSPHATASE (NUDX1; Yoshimura et al., 2007) and the INOSINE TRIPHOSPHATE PHOSPHATASE (ITPA; Straube et al. 2022). Although acting directly on the nucleotide triphosphates, it could be shown that a loss of NUDIX1 led to increased amounts of 8-oxo deoxyguanosine in DNA (Yoshimura et al., 2007). The incorporation of damaged nucleotides into DNA is thought to cause DNA damage that has to be repaired. As errors can also happen during DNA repair, DNA damage increases mutation rates. Interestingly, environmental factors were shown to influence the abundance of damaged metabolites. When loss-of-function mutants of *NUDIX1* were treated with 3 µM paraquat, an herbicide that leads to the generation of reactive oxygen species, the amount of 8-oxo deoxyguanosine in DNA rose significantly when compared to the wild type (Yoshimura et al., 2007). In non-plant organisms, it has been shown that such damage also affects other nucleotides.

We characterized a novel enzyme involved in the damage repair system of plants, the INOSINE TRIPHOSPHATE PYROPHOSPHATASE (ITPA; Straube et al., 2022; Figure 7). This enzyme belongs to the family of ITPases or *HYDROXYLAMINOPURINE MUTAGENESIS* 1 (HAM1) enzymes, which are known to dephosphorylate deaminated purine nucleotide triphosphates and thus prevent their incorporation into nucleic acids. These aberrant purines can arise from the deamination of canonical intermediates like ATP and GTP, or the aberrant phosphorylation of IMP and XMP. An analysis of the biochemical properties of ITPA demonstrated that the plant enzyme has similar kinetic properties to known ITPases from other organisms, dephorsphorylating inosine and xanthosine triphosphate (ITP and XTP), as well as deoxyinosine triphosphate (dITP). Interestingly, ITPA of *A. thaliana* also dephosphorylated inosine diphosphate (IDP), in contrast to the ITPase of humans and yeast (Davies et al., 2012). Furthermore, the enzyme is localized in the cytosol, nucleus, and chloroplasts. Mutants abolished in ITPA function display an earlier onset of senescence. They also have increased

concentrations of IDP and ITP, whereas XTP and dITP were not detectable. It was shown that ITP and dITP (despite being undetectable) are incorporated into nucleic acids, as the concentration of deoxyinosine and inosine in DNA and RNA, respectively, was higher in the mutants when compared to the wild type. This result is indicative for an activity of ITPA with dITP, although we were not able to directly detect this metabolite. Additional experiments indicated that oxidative stress induced by cadmium treatment led to significantly increased concentrations of IDP and ITP, suggesting that oxidative damage of ADP and ATP is likely the major source of IDP and ITP. By contrast, studies of the pathways of aberrant IMP phosphorylation indicate that they are unlikely to be a major source of IDP and ITP in plants.
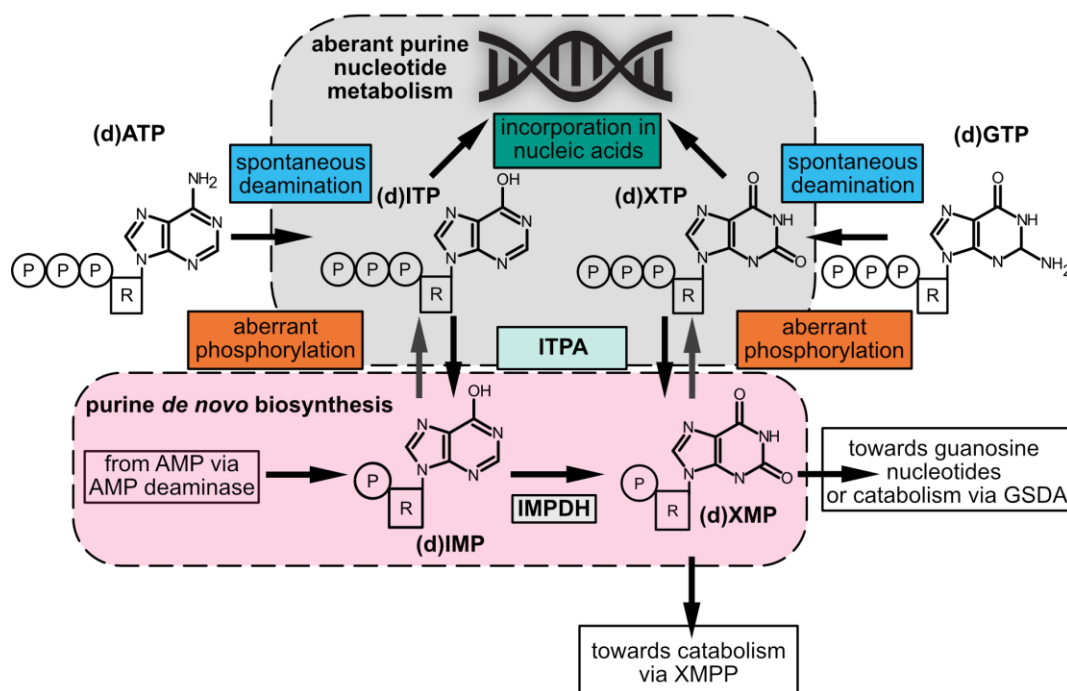

Figure 7. Overview of deaminated purine nucleotide metabolism and its interplay with purine *de novo* synthesis. Enzymatic and non-enzymatic steps that lead to the generation of deaminated purine nucleotides are shown, as well as their removal by the INOSINE TRIPHOSPHATE PYROPHOSPHATASE (ITPA) and reintegration into canonical purine nucleotide metabolism or potential incorporation into nucleic acids. (d)ATP, (deoxy)adenosine triphosphate; (d)ITP, (deoxy)inosine triphosphate; (d)XTP, (deoxy)xanthosine triphosphate; (d)GTP, (deoxy)guanosine triphosphate; AMP, adenosine monophosphate; (d)IMP, (deoxy)inosine monophosphate; (d)XMP, (deoxy)xanthosine monophosphate; IMPDH, INOSINE MONOPHOSPHATE DEHYDROGENASE; XMPP, XANTHOSINE MONOPHOSPHATE PHOSPHATASE.

RNA-Seq analysis of *itpa* and wild type plants led to the discovery that genes involved in the response to salicylic acid (SA), systemic acquired resistance and senescence are upregulated in this mutant background. Analyzing the concentrations of phytohormones revealed that mutants abolished in ITPA function have slightly increased levels of SA. The mechanistic connection remains to be elucidated, but it is possible that *itpa* plants suffer from DNA damage. Similarly, the mutants showed

1.5 Metabolite damage and molecular countermeasures

increased concentrations of inosine diphosphate (IDP) and inosine triphosphate (ITP) when grown in the presence of 10 µM cadmium, a heavy metal that is known to cause oxidative stress, when compared to mutants grown without the presence of heavy metals. No significant phosphorylation of inosine monophosphate (IMP) could be observed, neither *in vitro* nor *in vivo* (Straube et al. 2022), indicating that chemical damage, not enzyme promiscuity is the source for IDP and ITP. These results furthermore demonstrate that abiotic stress can lead to and enhance metabolite damage in plants. In summary we demonstrated that ITPA is required to dephosphorylate deaminated purine nucleotides to protect nucleic acids from their incorporation.

2. Publications and manuscripts

2.1 Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism

Henryk Straube[1], Markus Niehaus [1], Sarah Zwittian[1], Claus-Peter Witte[1] and Marco Herde[1]

[1] Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany

Type of authorship:	First author
Type of article:	Research article
Contribution to the publication:	Planned and performed experiments, analyzed all data, prepared the figures, participated in writing the manuscript

THE PLANT CELL

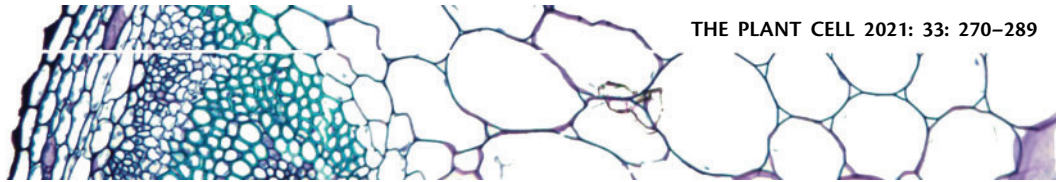**Research Article**

Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism

Henryk Straube ![ORCID],[1] Markus Niehaus ![ORCID],[1] Sarah Zwittian ![ORCID],[1] Claus-Peter Witte ![ORCID] [1] and Marco Herde ![ORCID] [1,*]

1   Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany

*Author for correspondence: mherde@pflern.uni-hannover.de

H.S. and M.H. designed the study. H.S., M.N., S.Z., and M.H. acquired the experimental data. H.S. analyzed the data and H.S., C.-P.W., and M.H. interpreted the data. H.S., C.-P.W., and M.H. wrote the manuscript. All authors read and revised the manuscript and agreed on the final version.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell) is: Marco Herde (mherde@pflern.uni-hannover.de).

Abstract

Detecting and quantifying low-abundance (deoxy)ribonucleotides and (deoxy)ribonucleosides in plants remains difficult; this is a major roadblock for the investigation of plant nucleotide (NT) metabolism. Here, we present a method that overcomes this limitation, allowing the detection of all deoxy- and ribonucleotides as well as the corresponding nucleosides from the same plant sample. The method is characterized by high sensitivity and robustness enabling the reproducible detection and absolute quantification of these metabolites even if they are of low abundance. Employing the new method, we analyzed *Arabidopsis thaliana* null mutants of *CYTIDINE DEAMINASE*, *GUANOSINE DEAMINASE*, and *NUCLEOSIDE HYDROLASE 1*, demonstrating that the deoxyribonucleotide (dNT) metabolism is intricately interwoven with the catabolism of ribonucleosides (rNs). In addition, we discovered a function of rN catabolic enzymes in the degradation of deoxyribonucleosides in vivo. We also determined the concentrations of dNTs in several mono- and dicotyledonous plants, a bryophyte, and three algae, revealing a correlation of GC to AT dNT ratios with genomic GC contents. This suggests a link between the genome and the metabolome previously discussed but not experimentally addressed. Together, these findings demonstrate the potential of this new method to provide insight into plant NT metabolism.

Introduction

Metabolomics of nucleotides (NTs) and nucleosides (Ns) in plants is a notably understudied area, in part is due to technical challenges concerning sample preparation and chromatographic separation. In particular, research on deoxyribonucleotide (dNT) and deoxyribonucleoside (dN)

metabolism would greatly benefit from methods allowing the comprehensive quantification of these metabolite classes in plant samples.

Several enzymes of plant dNT metabolism have been functionally characterized in *Arabidopsis thaliana*, but a direct impact of loss-of-function mutants on the dNT pools

was hitherto only proven in one case with a PCR-based approach that allows only relative quantification of dNT triphosphates (dNTPs) in vivo (Wang and Liu, 2006; Garton et al., 2007; Yoo et al., 2009). Severe phenotypes like aberrant leaf morphology, growth inhibition, white spots in leaves, and reduced seed yield result from mutation of plant dNT metabolism genes (Wang and Liu, 2006; Garton et al., 2007; Dubois et al., 2011; Pedroza-García et al., 2015, 2019; Le Ret et al., 2018). Although not shown, these mutations likely cause altered cellular dNT concentrations, which in turn are probably often the reason for the phenotypic alterations. Changes of NT quantities affecting DNA and RNA replication were observed in similar mutants of non-plant organisms (Nick McElhinny et al., 2010a, 2010b; Gon et al., 2011; Kumar et al., 2011). Furthermore, the catalytic function of plant dN kinases could so far only be validated in vitro (Stasolla et al., 2003; Clausen et al., 2012; Pedroza-García et al., 2015; Le Ret et al., 2018; Pedroza-García et al., 2019), because possible changes of dN concentrations in the corresponding mutants cannot be assessed with current methods (Pedroza-García et al., 2015, 2019; Le Ret et al., 2018). The inability to detect dNs is also the reason why the metabolic fate of dNs is currently unknown in plants. dNs might be degraded by the same enzymes that also catabolize ribonucleosides (rNs), because at least some of these enzymes were shown to catabolize both types of substrates in vitro (Dahncke and Witte, 2013; Chen et al., 2016).

In contrast to ribonucleotides (rNTs), dNTs, and dNs, the detection and quantification of rNs and their degradation products is less challenging; this spurred a comprehensive characterization of the rN degradation pathways in vivo (Baccolini and Witte, 2019; Witte and Herde, 2020). The catabolism of rNs is integrated into a complex network including the de novo biosynthesis of rNTs and the salvage (recycling) of rNs and nucleobases ultimately giving rise to ratios and amounts of rNTs suitable for all downstream processes (Zrenner et al., 2006; Ashihara et al., 2020; Witte and Herde, 2020).

Unlike most animals, plants are able to fully catabolize purine and pyrimidine rNs including the nucleobases, using the released nitrogen for amino acid biosynthesis (Werner et al., 2010; Werner and Witte, 2011; Ashihara et al., 2020; Witte and Herde, 2020). Three key enzymes participate in the initial steps of rN catabolism: cytidine deaminase (CDA; Vincenzetti et al. 1999; Kafer and Thornburg, 2000; Chen et al., 2016; Witte and Herde, 2020), guanosine deaminase (GSDA; Dahncke and Witte, 2013; Witte and Herde, 2020), and nucleoside hydrolase 1 (NSH1; Jung et al., 2009, 2011; Baccolini and Witte, 2019; Witte and Herde, 2020). CDA was shown to deaminate cytidine and deoxycytidine in vitro and mutants of *CDA* accumulate cytidine in vivo (Chen et al., 2016). GSDA deaminates guanosine and deoxyguanosine in vitro and the mutants accumulate guanosine in vivo (Dahncke and Witte, 2013; Baccolini and Witte, 2019), while NSH1 participates in the hydrolysis of uridine, xanthosine,

and inosine in vitro and in vivo (Jung et al., 2009; Riegler et al., 2011; Baccolini and Witte, 2019).

Mutants of *GSDA*, *CDA*, and *NSH1* show phenotypical abnormalities (Jung et al., 2011; Dahncke and Witte, 2013; Chen et al., 2016; Baccolini and Witte, 2019) during germination, development, and dark stress. The abnormal phenotypes are thought to result from the accumulation and toxicity of rNs or the lack of degradation-derived metabolites (Stasolla et al., 2003; Schroeder et al., 2018; Baccolini and Witte, 2019). So far, only metabolomic data for rNs and their degradation products have been obtained in *gsda*, *cda*, and *nsh1* mutants, whereas rNT, dNT, and dN pools have not been investigated due to technical limitations. However, it would be interesting to quantify dNs in these mutants to clarify whether the corresponding enzymes are indeed involved in dN degradation in vivo. The substantial accumulation of rNs (and perhaps also dNs) in these mutants might also result in altered concentrations of rNTs and dNTs because plants possess kinases for the salvage of rNs and dNs (Witte and Herde, 2020).

The primary metabolome comprises four main classes of metabolites: the amino acids, the carbohydrates, the lipids, and the NTs. While the first three classes are routinely characterized comprehensively in plant metabolome studies using well-established methods (Salem et al., 2020), this does not apply to NTs and NT-derived metabolites, which are often not analyzed at all or are highly underrepresented. A main reason lies in the lack of suitable methods for a comprehensive analysis of NTs in plant samples with modern mass spectrometry (MS) techniques. So far, NT analysis in plants (see literature survey summarized in Supplemental Table S1) have employed either liquid chromatography (LC) combined with photometric detection (Meyer and Wagner, 1985; Dutta et al., 1991; Katahira and Ashihara, 2006), a polymerase assay (Castroviejo et al., 1979; Feller et al., 1980; Wang and Liu, 2006; Garton et al., 2007), or thin-layer chromatography (TLC; Nygaard, 1972). These methods suffer from different drawbacks, for example, low sensitivity (photometric detection), relative quantification of only dNTPs (polymerase), or the need for radiolabelled starting material (TLC). Some metabolome studies quantified rNTs but not dNTs employing ion-exchange chromatography–MS or LC–MS (Rolletschek et al., 2011; Souza et al., 2015); however, recovery rates of these methods were not reported. To improve dNT detection, some protocols remove rNTs by a periodate treatment (Dutta et al., 1991) that likely also affects dNT species such as dGTP, which reacts with dicarbonyl compounds resulting from the addition of periodate (Tanaka et al., 1984; Henneré et al., 2003). Focused rN extraction methods from plant material also exist (Kopecná et al., 2013) but these are not suitable for NT species. Ideally, a method coupling LC with modern MS would be needed.

Plant samples are particularly challenging for metabolomics employing LC–MS because plants contain a plethora of metabolites making extracts (the matrix) very complex.

In fact, we previously quantified dNTPs in embryos of *Drosophila melanogaster* (Liu et al., 2019), but that same method failed to work for Arabidopsis tissues. So-called matrix effects can cause ion suppression greatly reducing the sensitivity of detection in the MS, which for dNTs is critical as they are not abundant in plant cells. Studies in human cells suggest that solid phase extraction (SPE), either employing a silica or an anion exchange resin, is uniquely suited to enrich NTs and reduce matrix effects for sensitive detection (Cohen et al., 2009; Pabst et al., 2010; Kong et al., 2018). Additionally, it has been pointed out that efficient quenching of enzymatic activities is essential to stabilize plant extracts, because they contain phosphatases displaying residual activity in mixed aqueous and organic solutions (Ullrich and Calvin, 1962; Bieleski, 1964; Ikuma and Tetley, 1976).

The LC of charged metabolites as front end for an MS analyzer is also not as straightforward as for moderately polar or hydrophobic compounds. For the LC–MS detection of NTs, various LC column materials have been used such as porous graphitized carbon in reverse-phase mode (Cohen et al., 2009) and resins carrying zwitterionic functional groups for hydrophilic interaction chromatography (Kong et al., 2018). A common problem with these chromatographies is that retention times can be unstable varying up to several minutes (Pabst et al., 2010). With the porous graphitized carbon stationary phase, this variation in chromatographic behavior is aggravated when crude plant extracts are used (our own observation) and might depend on the redox status of the column resin (Pabst et al., 2010). Taken together, these issues suggest that the detection and quantification of NTs in plant samples will require the development of a specially adapted protocol considering that methods of sample preparation and chromatography might interact to some extent.

In this study, we have established a sensitive and robust analytical method to simultaneously determine absolute concentrations of a comprehensive set of NTs and Ns in plants. The method works for a wide range of plant species including a moss (*Physcomitrium* [*Physcomitrella*] *patens*) and three algae (*Chlamydomonas reinhardtii*, *Mougeotia scalaris*, and *Volvox carteri*). We used the new protocol to analyze three Arabidopsis loss-of-function mutants in nucleoside catabolism lacking CDA, GSDA, and NSH1, respectively. We discovered that all three mutants harbor highly unbalanced rNTP and dNTP pools and we present evidence that CDA and NSH1 are not only required for rN but also for dN degradation in vivo.

Results

Optimization of sample preparation

The complex matrix of plant extracts notably complicates MS (Bieleski, 1964; Bieleski and Young, 1963; Nieman et al., 1978) preventing for example the straightforward detection of dNTs, which can be readily detected in non-plant samples (Kuskovsky et al., 2019). Additionally, plant samples must be

efficiently quenched for NT analysis because plants possess stable phosphatases (Bieleski, 1964; Ikuma and Tetley, 1976) that are resistant to harsh conditions like organic solvents. A method (Figure 1) addressing these challenges was devised comprising tissue rupture and acid quenching, liquid/liquid extraction (LLE), weak anion SPE, and LC coupled to MS (LC–MS).

We evaluated different methods for sample preparation. In Table 1, we present the main findings of the optimization while the full detail of tested methods is shown in Supplemental Tables S2, S3. Consistent with previous studies, the NT triphosphates (NTPs) were more efficiently recovered upon extraction with strong acids like trichloroacetic acid (TCA) and perchloric acid (PCA) than with the organic solvent methanol (Bieleski, 1964; Dietmair et al., 2010) since the latter probably failed to inactivate all NT phosphatases. In our hands, recovery with PCA (Ashihara et al., 1987) was not as good as with TCA presumably because the PCA is removed by precipitation creating a bulky pellet, which traps liquid making quantitative sample recovery more difficult.

By contrast, TCA is removed by a LLE step in our method. Such a LLE has been developed over 40 years ago for mammalian systems (Khym, 1975) but has lost popularity, probably because the original protocol requires environmentally hazardous and expensive 1,1,2-trichlor-1,2,2-trifluorethan (Freon-113). We replaced Freon-113 with dichloromethane (DCM) and adopted the technology for plant extracts.
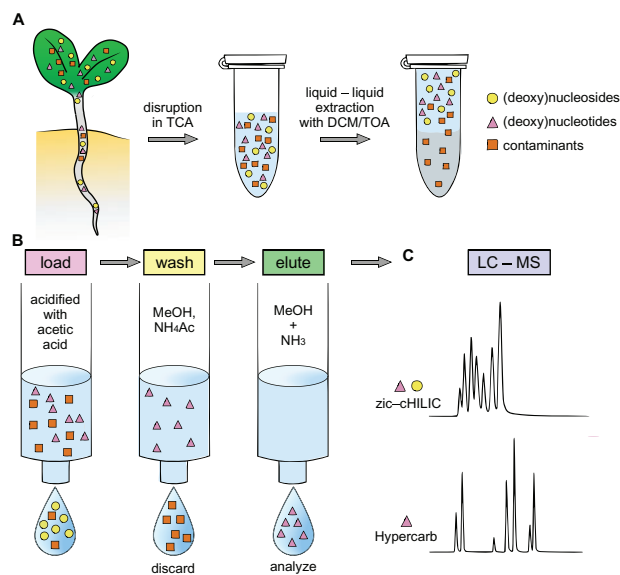
Figure 1 Schematic overview of the method for the extraction and analysis of NTs and Ns. Plant material is disrupted and quenched with TCA, which is removed together with apolar contaminants by LLE with DCM and TOA (A). The extract is loaded onto a weak-anion exchange SPE cartridge, the flow-through contains (deoxy)nucleosides for analysis. Subsequently, contaminants are depleted by washing with methanol (MeOH) and ammonium acetate (NH4Ac), resulting in elution of (deoxy)ribonucleotides with MeOH and ammonia (NH3; B). Isolated fractions are analyzed by LC–MS using a zic-cHILIC column (NTs and Ns) or a Hypercarb column (NTs; C).

Table 1 Selection of tested methods and their combined effect on NT recovery

Method	Quenching	LLE/acid quenching	Dilution	Recovery (%) dTTP	Recovery (%) ATP	Recovery (%) dTMP
3	80/20 MeOH/10 mM NH$_4$Ac pH 4.5 (v/v)	–	–	10.2	10.3	107.8
7	6% PCA	20% KOH	–	92.3	61.6	26.8
8	15% TCA	DCM/TOA (78/22 v/v)	–	89.6	92.7	44.9
9	15% TCA	DCM/TOA (78/22 v/v)	1 mL water	87.8	90.0	88.4

Apart from removing the TCA in an elegant way resulting in little sample loss, the LLE additionally eliminates apolar metabolites from the extract. This is advantageous because it reduces the complexity of the matrix likely contributing to the improved NT recovery in comparison to the PCA method. An additional boost of recovery for NT monophosphates (NMPs) was achieved by diluting the sample with water prior to application to the SPE (Tables 1, 2). The dilution results in a reduced salt concentration in the sample fostering the binding of the NMPs to the anion-exchange matrix of the SPE. One can envisage that by fine-tuning the salt load of the sample, it will be possible to select against NMPs and other less-strongly bound ionic compounds. Such a reduction of matrix complexity could be advantageous when for example only the NTPs are in the analytical focus.

Interestingly, in the flow-through of the SPE loading step, we were able to detect deoxy- and rNs (dNs). Therefore, we included Ns in the evaluation of the final method (Table 2) determining the recovery rate of NMPs and NTPs in the eluate as well as Ns in the flow-through. Recovery rates for NTPs ranged from 41.2% to 104.0%, for NMPs from 85.8% to 113%, and for Ns from 82.6% to 96.4% (Table 2). Dinucleotides were not assessed because we reasoned that they will co-elute with NMPs and NTPs. This was later confirmed for ADP. In summary, the data show that the new method allows the simultaneous extraction and preparation for MS detection of NTs and Ns from the same plant sample.

Comparison of chromatography methods and method validation

The chromatography method has a substantial impact on the sensitivity of the MS analysis because analytes and matrix are differentially separated and focused by distinct chromatographic techniques. Ion chromatography and capillary electrophoresis as well as hydrophilic interaction chromatography (HILIC) and porous graphitized carbon (PGC) chromatography have been used as front ends for NT analysis by MS (Ashihara et al., 1987; Riondet et al., 2005; Pabst et al., 2010; Kong et al., 2018). We used a zic-cHILIC (Merck) for HILIC (in the following called the cHILIC method) and a Hypercarb column (Thermo) for PGC chromatography (in the following called the Hypercarb method) and optimized the respective chromatographic and MS parameters (see the "Materials and methods" section, Supplemental Figures S1–S3, and Supplemental Tables S4–S6). Using the Hypercarb

Table 2 Relative recovery of Ns, NMPs, and NTPs

Relative recovery (%)

NTPs		NMPs		Ns	
dATP 15N	100.2				
dCTP 15N	104.0	dCMP 13C, 15N	113.1		
dGTP 13C, 15N	64.6				
dTTP 13C, 15N	87.8	dTMP 13C, 15N	88.4	Deoxythymidine 13C, 15N	94.8
ATP 2H	90.0	AMP 15N	90.2	Adenosine 13C	82.6
CTP 2H	102.5	CMP 13C, 15N	75.5	Cytidine 15N	96.3
GTP 2H	41.2	GMP 15N	85.8	Guanosine 15N	86.4
UTP 2H	91.0	UMP 15N	103.5	Uridine 15N	82.6
				Inosine 15N	96.4

method, a lower limit of quantitation (LLOQ) of 0.1 pmol on column for NTs was determined, which was 5–50 times lower than the 0.5–5 pmol LLOQ on column measured by the cHILIC method (Table 3). For Ns, the LLOQ was 0.1 pmol on column (Table 4). These sensitivities are similar or even better than those previously described for the analysis of NTPs from mammalian cells (Kong et al., 2018). Thus, the sensitivity of our method is sufficient to allow for the first time the detection and quantification of dNTs in plant material by MS.

The quantification of isotope standards (ISTDs) spiked into matrix derived from SPE and resolved either by the Hypercarb or cHILIC method showed that standard amount and detector signal correlated with good linearity resulting in R^2 values equal or higher than 0.96 for all tested substances (Tables 3, 4). Different physiological conditions leading to changing NT or nucleoside concentrations can therefore be investigated.

Intra- and inter-day precision was evaluated for all used chromatography methods and three different concentrations of ISTDs in plant matrix (Supplemental Tables S7, S8). In general, the coefficient of variation (CV%) was higher for the cHILIC than for the Hypercarb method and inter-day variation was higher than intra-day variation. The values ranged

Table 3 Calibration range, coefficient of determination (R^2), and LLOQ for NTPs using the Hypercarb or the cHILIC method

	Calibration range (pmol)	R^2 Hypercarb	Hypercarb LLOQ (pmol)	R^2 cHILIC	cHILIC LLOQ (pmol)
dATP	2.5–80	0.96	0.1	0.99	1
dCTP		0.98	0.1	0.99	1
dGTP		0.99	0.1	0.97	1
dTTP		0.99	0.1	0.98	0.5
ATP	25–800	0.99	0.1	0.98	1
CTP		0.96	0.1	0.98	1
GTP		0.99	0.1	0.99	5
UTP		0.99	0.1	0.98	1

Table 4 Calibration range, coefficient of determination (R^2), and LLOQ for Ns using the cHILIC method

	Calibration range (nmol)	R^2 cHILIC	cHILIC LLOQ (pmol)
Deoxythymidine	0.125–2	0.99	0.1
Adenosine		0.99	0.1
Cytosine		0.99	0.1
Guanosine		0.99	0.1
Inosine		0.99	0.1
Uridine		0.99	0.1

from 11% to 31% and are in good agreement with the variations determined for other SPE-based methods for NT analysis in non-plant organisms (Kong et al., 2018; Cohen et al., 2009).

In order to assess the effectiveness of the SPE method, we compared the matrix effect factor (MEF) of matrix obtained from the LLE step (before SPE) with the MEF of matrix obtained from the complete method (after SPE). The higher the MEF, the greater is the signal suppression by the respective matrix (see the "Materials and methods" section). Additionally, the comparison was made for both chromatographic techniques used, i.e. depending on whether the sample was separated via the Hypercarb or cHILIC method. Irrespective of the chromatography method, the SPE lowered the MEFs substantially and enabled the detection of several standards that were undetectable without SPE (Supplemental Tables S9, S10). Interestingly, the MEFs were generally lower when the samples were analyzed with the cHILIC method, showing that not only sample preparation but also the type of chromatography is important. For Ns, which are not retained by the SPE, the removal of charged metabolites from the matrix by the SPE nevertheless reduced the MEFs (Supplemental Table S10). The matrix after SPE even exerted a positive effect on sensitivity compared with buffer for some analytes (cytidine, inosine and uridine). Such ion enhancement effects have been reported previously (Zhou et al., 2017).

The robustness of the presented method, reflected in the LLOQs, variations, and MEFs, is equal to published SPE protocols for NT analysis of other organisms (Harmenberg et al., 1987; Cohen et al., 2009; Guo et al., 2013; Kong et al., 2018). Remarkably, for Ns, the method has a superior recovery and a similar MEF compared with SPE methods focusing exclusively on N analysis (Sawert et al., 1987; Farrow and

Emery, 2012; Kopecná et al., 2013). Until now, the parallel quantification of Ns and NTs from plant samples was not possible—this robust method now allows this analysis and thus opens up new possibilities to investigate the NT metabolism of plants in greater depth.

For metabolite analysis, many laboratories have access to an LC system coupled with a photometric detector, whereas a MS detector is less common. Thus, it would be desirable that our sample preparation method also improves the photometric detection of NTs. After treating the samples with LLE alone, we were unable to detect any signals for NTs in the UV trace (at 254 nm) using the cHILIC method, but after SPE several peaks were detected (Supplemental Figure S4). For proof-of-concept, we showed that MS signals for the rNTs ATP, UTP, CTP, and GTP were associated with four of the photometric peaks. This demonstrates that the sample preparation via LLE and SPE improves NT detection also when a photometric detector is employed. However, in this case, it is recommendable to use other (non-MS-compatible) chromatographic separation techniques with increased resolution for NTs (see e.g. Meyer and Wagner, 1985). Our sample preparation method requires little plant material and allows inter alia the quantification of CTP and GTP in Arabidopsis leaves even without MS, which before has been difficult (Ashihara et al., 2020). Therefore, it might generally increase the sensitivity for NTs independent of the chosen detection method. Nonetheless, the analysis of less abundant NTs with photometric detection, for example dNTPs, would require significant upscaling.

dNTP pools in different plants are variable and correlate with the genomic GC content

We first applied the new method for the detection of NTs in the model plant A. thaliana. From as little as 100 mg of fresh material of either 7-day-old seedlings grown in liquid culture or leaves of 33-day-old plants grown on soil, not only all canonical rNTPs but also all of their dNTP counterparts were robustly detected (Figure 2, A and B). The concentrations of the rNTPs were in the nmol g^{-1} range, while dNTPs were 100- to 1000-fold less concentrated (Table 5). Data on the NTP contents in plants are scarce (Ashihara et al., 2020; Witte and Herde, 2020; Supplemental Table S1). However, our results on dNTPs are overall consistent with those from a previous study using radiolabelled plant cells and TLC (Nygaard, 1972). Because ATP and UTP are needed in higher amounts as energy carriers and for cell wall
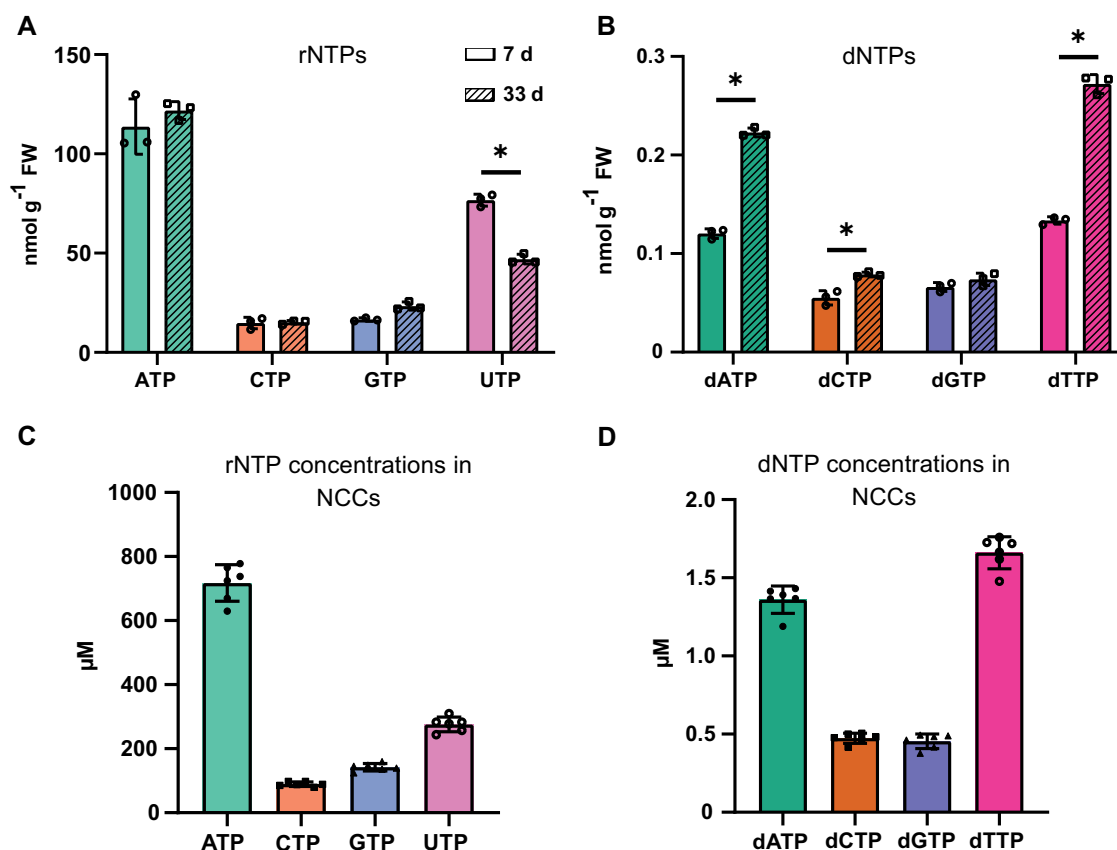
Figure 2 Cellular concentrations and absolute amounts of rNTs and dNTs in Arabidopsis leaves and seedlings. rNTs (rNTPs; A) and dNTs (dNTPs; B) were quantified in 7-day-old seedlings (unhatched bars) and 33-day-old rosette leaves (hatched bars). Each biological replicate represents a pool of several seedlings (A) or two leaves (B) which met the criteria outlined in the "Materials and methods" section from one individual plant. For seedlings, every replicate is a pool of seedlings grown together in one flask. Replicates were grown in parallel under identical conditions. Concentration of rNTPs (C) and dNTPs (D) in NT containing compartments of 33-day-old leaves. The concentration was calculated with the assumptions and formulas given in the "Materials and methods" section. Error bars indicate standard deviations (SD) for $n = 3$ biological replicates (A and B) or $n = 6$ biological replicates (C and D). * means $P < 0.05$ and was determined using a two-way ANOVA with Sidak's post test.

Table 5 Ratios of rNTP/dNTP in 7- and 33-day-old Arabidopsis plants

	ATP/dATP	CTP/dCTP	GTP/dGTP	UTP/dTTP
7 day	945	268	250	574
33 day	528	189	313	166

synthesis, respectively, the ratio of ATP/dATP and UTP/dTTP is higher than for CTP/dCTP and GTP/dGTP in seedlings (Table 5).

Interestingly, the dTTP and the dATP pools are increased and the UTP pool is decreased in older plants in comparison to seedlings (Figure 2, A and B). This may seem counterintuitive because one might expect more dNTPs in young growing tissue undergoing frequent cell divisions, but older tissues might require dNTPs for endoreduplication or for DNA damage repair. However, one needs to bear in mind that the growing conditions of the seedlings and the older plants were quite distinct, which may also account for the differences.

The immediate concentration of metabolites in a cell or cellular compartment influences enzyme activities or regulatory processes. It is therefore interesting to estimate the cellular concentrations of NTs. We used the cell and cell compartment volumes reported by Koffler et al. (2013) for leaves of 33-day-old plants. Based on the simplifying assumptions that NTs are equally distributed in the nucleus, cytoplasm, chloroplasts, and mitochondria, and largely absent from other cellular compartments, the average rNTP concentrations ranged from 100 to 750 μM whereas the concentrations of dNTPs were between 0.5 and 1.6 μM (Figure 2, C and D).

To ensure that the quenching in our method is suitable to preserve the phosphorylation status of the NTs, the adenylates (AMP, ADP, and ATP) in 33-day-old *A. thaliana* leaves were quantified (Table 6). The ratio of ATP/ADP was 15.2, while the ATP/AMP ratio was notably higher, resulting in an adenylate energy charge (AEC) of 0.97. We conclude that the phosphorylation status of these metabolites was maintained during the extraction, which is in line with the high recovery rates for these metabolites (Table 2). The ratios and the AEC are also consistent with results from other studies (Stitt et al., 1982; Guérard et al., 2011), although the

Table 6 Absolute amounts of AMP, ADP, and ATP from 33-day-old Arabidopsis plants grown under long day conditions

Metabolite	nmol g^{-1} FW[a]
AMP	0.4 ± 0.04
ADP	7.6 ± 2.2
ATP	115.6 ± 7.8
Ratios	
ATP/ADP ratio	15.2
ATP/AMP ratio	286.4
AEC[b]	0.97 ± 0.01

[a]$n = 6$ biological replicates, where every replicate represents the oldest leaves of a different plant.
[b]AEC is defined as (ATP + 1/2ADP)/(ATP + ADP + AMP).

ATP/ADP and ATP/AMP ratio in 33 day-old-plants is higher than in other studies for unknown (Stitt et al., 1982; Savitch et al., 2001; Carrari et al., 2005). Additionally, the data show that also ADP as an example for the dinucleotides is co-eluted with the NMPs and NTPs from the SPE and can be quantified with this method.

We also identified NTs using an Orbitrap mass analyzer coupled to the Hypercarb chromatography providing exact masses for precursor and product ions as well as isotope patterns as additional evidence for the correct identification of the metabolites. In samples of 7-day-old seedlings, we identified all canonical NTPs with high confidence by MS/MS (Supplemental Table S11) except dGTP and dCTP for which the detection was not sensitive enough. Several NTPs were also detected in full MS mode suggesting that the samples provided by our method are in principle suitable to perform non-targeted analysis.

Plant species differ greatly in their metabolite composition in part due to variation in secondary metabolism. We wanted to assess if our method established in Arabidopsis is suitable for dNTP quantification in samples from diverse plant backgrounds. We extracted dNTPs from 11 different species encompassing monocotyledons, dicotyledons, a bryophyte, and three algae. In all tested species, the sensitivity was sufficient to detect and quantify dNTPs (Supplemental Figure S5) proving that our method is generally suitable for NT analysis of plants.

Because in DNA G pairs with C and A pairs with T, one might assume that the ratios of the corresponding dNTPs, i.e. the dGTP/dCTP and the dATP/dTTP ratios, are (i) close to unity and (ii) similar in different plants. Consistent with this concept, the dGTP/dCTP ratios were similar in the investigated species (with the exception of *C. reinhardtii*) but were general slightly lower than one, however the dATP/dTTP ratios were more variable, with *Avena sativa*, *Oryza sativa*, and *Solanum lycopersicum* containing notably more dATP than dTTP and *P. patens* having more dTTP than dATP (Supplemental Figure S5 and Figure 3). We additionally asked whether species with a high GC content in the DNA also have proportionally more dGTP and dCTP in their dNTP pool (GC content in dNTPs). Our results show that there is a moderate to strong positive correlation between these two parameters with a R^2 value of 0.67 (Figure 4)

suggesting a link between the genome composition and the dNTP metabolome.

CDA and NSH1 metabolize dNs in vivo

The role of several enzymes in rN degradation is already well established, but to date it remains unclear whether dNs are also substrates of these enzymes in vivo, probably because these compounds are rather difficult to detect and quantify. We chose to investigate null mutants of genes encoding CDA (Vincenzetti et al. 1999; Faivre-Nitschke et al., 1999; Kafer and Thornburg 2000; Chen et al., 2016), GSDA (Dahncke and Witte, 2013; Baccolini and Witte, 2019), and NSH1 (Jung et al., 2009, 2011; Riegler et al., 2011) to address this issue with our newly established method. GSDA deaminates the purine nucleoside guanosine to xanthosine and CDA deaminates the pyrimidine nucleoside cytidine to uridine. NSH1 hydrolyzes the glycosidic bond of uridine generating uracil and ribose and is as well an essential component of a nucleoside hydrolase complex required for the hydrolysis of xanthosine to xanthine and ribose (Figure 5).

In seeds and seedlings of the Arabidopsis wild type, all rNs including the low abundant inosine but also deoxyadenosine in seeds and deoxythymidine in seedlings were reliably detected (Supplemental Figure S6 and Figure 6). Relative and absolute abundances of rNs in the *GSDA*, *CDA*, and *NSH1* loss-of-function mutants and in wild-type plants (Supplemental Figure S6) were consistent with concentrations reported in previous studies using a different extraction approach (Chen et al., 2016; Baccolini and Witte, 2019). In all mutants the abundance of adenosine, a metabolite not investigated previously in the context of these mutants, was elevated in seedlings (but not in seeds) compared with the wild type. With respect to the dNs, seeds contained a pool of deoxyadenosine that was not affected by any of the investigated mutations (Figure 6). By contrast, seeds of the *nsh1* and *cda* mutants showed an accumulation of deoxythymidine not observed in the Col-0 or *gsda* backgrounds (Figure 6, A). In seedlings, deoxythymidine additionally accumulated in the *GSDA* mutant, but was also detected at a lower level in Col-0 (Figure 6, B). However, in both tissues, the *nsh1* mutant accumulated by far the most deoxythymidine, suggesting that NSH1 is directly involved in the turnover of this metabolite. High concentrations of guanosine and cytidine occurring in the *gsda* and *cda* mutants, respectively (Supplemental Figure S3), are probably inhibiting NSH1 partially (Witte and Herde, 2020), leading to the intermediate deoxythymidine build-up observed in these backgrounds. Furthermore, deoxycytidine accumulated exclusively in the *cda* mutant (Figure 6) providing evidence for a role of CDA in the catabolism of this dN.

In summary, we provide evidence that NSH1 is involved in deoxythymidine hydrolysis and that CDA is responsible for deoxycytidine deamination in vivo, consistent with the ability of CDA to deaminate this dN in vitro (Chen et al., 2016). GSDA deaminates deoxyguanosine in vitro (Dahncke and Witte, 2013), but intriguingly deoxyguanosine was not
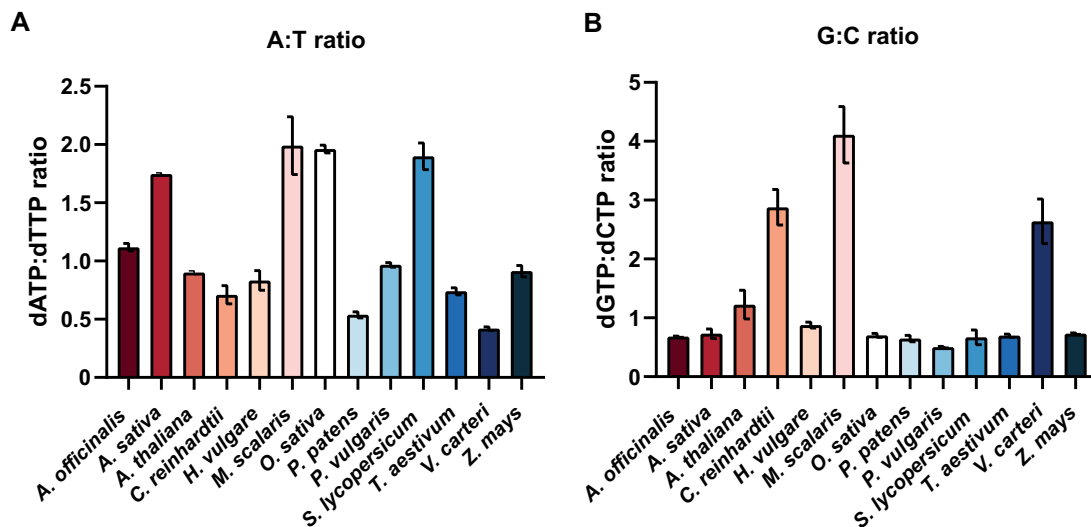
Figure 3 A/T and G/C ratios of dNTs (dNTPs) in different plant species. Ratios of dNTPs in *A. officinalis, A. sativa, A. thaliana, C. reinhardtii, H. vulgare, M. scalaris, O. sativa, P. (P.) patens, P. vulgaris, S. lycopersicum, T. aestivum, V. carteri,* and *Z. mays.* (A) Ratio of dATP to dTTP. (B) Ratio of dGTP to dCTP. Error bars indicate standard deviation (SD) for *n* = 3 biological replicates with each replicate representing a pool of several plants/seedlings all grown in parallel under the same environmental conditions.
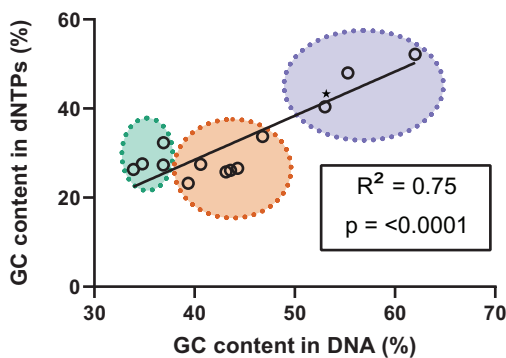

Figure 4 Regression analysis of the GC content in dNTPs and DNA. The GC content in the dNTPs (*y*-axis) is plotted against the genomic GC content (*x*-axis) and the Pearson correlation coefficient and the corresponding *P* value of a linear regression are calculated. Every white circle represents the average of three measurements. The green circle comprise all dicotyledonous species (*A. thaliana, P. vulgaris, S. lycopersicum*) and the bryophyte (*P. patens*), the orange circle encompasses the monocotyledonous species (*A. officinalis, A. sativa, H. vulgare, O. sativa, T. aestivum,* and *Z. mays*) and the purple circle alga (*C. reinhardtii, M. scalaris,* and *V. carteri*). * indicates an estimate of genomic GC content by extrapolation from known codon DNA sequences (R^2 without this datapoint is 0.73).

Figure 5 Scheme of nucleoside and NT metabolism. The (deoxy)nucleoside catabolism is highlighted with a gray background. NK, deoxynucleoside kinase; TK, thymidine kinase; rNK, ribonucleoside kinases; (d)NMK, (deoxy)nucleoside monophosphate kinases; and NDPK, nucleoside diphosphate kinases.

detectable in any of the analyzed genotypes (data not shown), although high concentrations of guanosine were detected in *gsda* seeds and seedlings (Supplemental Figure S6). Maybe for deoxyguanosine there is an additional metabolic escape route which is currently unknown.

CDA, GSDA, and NSH1 influence dNT and rNT pools

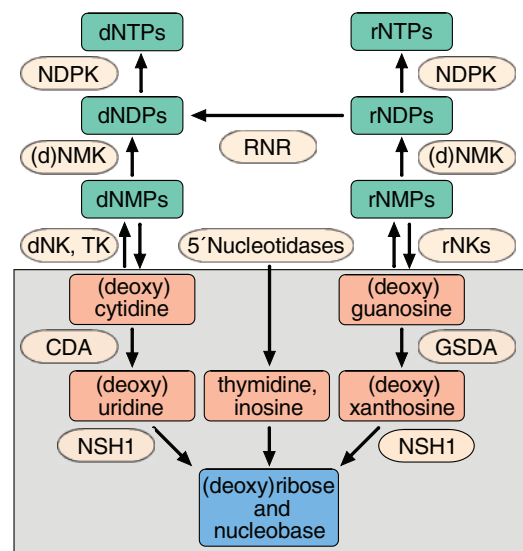It is widely accepted that NT metabolism can be manipulated by the supply of extracellular Ns and nucleoside analogs, a concept exploited in chemotherapy (Galmarini et al., 2003; Robak and Robak, 2013). It is also established in plants that extracellular Ns can be taken up and interfere with plant metabolism (Traub et al., 2007; Chen et al., 2016; Ashihara et al., 2020). In plants, little is known about the impact of imbalanced intracellular N pools on the abundance of NTs. Therefore, we analyzed the NTs in the *cda, gsda,* and *nsh1* mutants, which have altered N pools. In seeds and seedlings lacking CDA and GSDA, the amounts of CTP and GTP are increased, respectively (Figure 7). Consistent with a previous study (Riegler et al., 2011), plants lacking NSH1
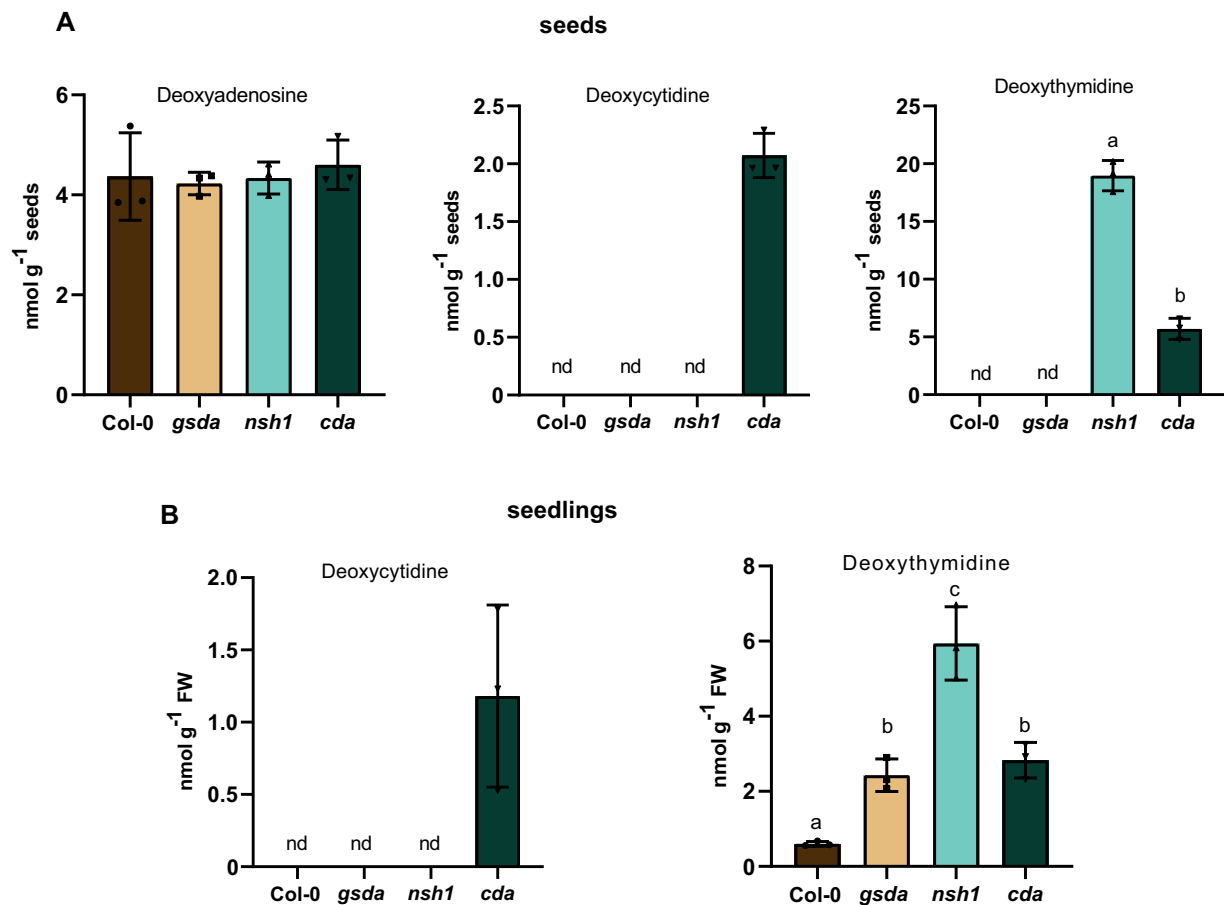
A

seeds

Deoxyadenosine

Deoxycytidine

Deoxythymidine

B

seedlings

Deoxycytidine

Deoxythymidine

Figure 6 Absolute quantification of dNs in mutants impaired in the catabolism of rNs. Concentrations of dNs in seeds (A) and in 7-day-old seedlings (B) of *A. thaliana* wild type, as well as mutants in the degradation of purine and pyrimidine Ns (*gsda*, guanosin deaminase). Error bars are SD, $n = 3$ biological replicates, for seeds, three independent seed pools derived from different mother plants, for seedlings, three pools of seedlings from three independent liquid cultures were used. Statistical analysis was performed using one-way ANOVA with Tukey's post hoc test or in case of deoxythymidine by a two-tailed Student's *t* test. Different letters indicate $P < 0.05$. nd, not detected. FW, fresh weight.

accumulate more UMP and UTP compared with the wild-type (Figures 7, 8).

In general, the effects on NTP levels are also reflected by the respective NMPs (Figure 8). The accumulation of these NTs is probably a direct consequence of the increased pool sizes of the corresponding Ns (cytidine, guanosine, and uridine) in these mutants (Supplemental Figure S6), because the substrate availability for nucleoside and NT kinases (see scheme in Figure 5) is increased. However, some changes in the NT pools of the mutants cannot be explained by direct effects. The decrease of IMP concentrations in plants lacking GSDA or NSH1 (Figure 8, B) for example is puzzling because both contain much more inosine than the wild type (Supplemental Figure 3, B). Seeds and seedlings lacking GSDA have an increased content of ADP and ATP resulting in a higher AEC as well as more UMP and UTP. In *gsda* seedlings, there is also more CTP (Figures 7, 8). Thus, in seedlings, all measured NTP pools are strongly increased in the *gsda* background (Figure 7). The indirect changes in the *cda* and *nsh1* backgrounds are more subtle and are often similar. In seedlings, both mutants contain for example

more GTP and ATP but in tendency less AMP than the wild type resulting in a higher AEC.

Except for the increased UTP level in *cda* seeds, all effects are more pronounced or even exclusively observed in seedlings compared with seeds. This also holds true for the Ns where for example adenosine only accumulates in mutant seedlings, but not in seeds (Supplemental Figure S6). As would be expected, these observations suggest that the flux through the NT metabolism is generally higher in growing seedlings compared with dormant seeds. Consistently, seeds harbor 5–10 times less rNTPs (and no detectable amounts of dNTPs), store adenylates mainly as AMP, display a lower ATP/ADP ratio, and have a lower AEC compared with seedlings which is in accordance with the literature (Raveneau et al., 2017; Figure 7).

Intriguingly, plants lacking enzymes of rN catabolism are not only impaired in the degradation of Ns (Figure 6 and Supplemental Figure S6) but also have elevated concentrations of dNTs, the substrates for DNA synthesis (Figure 9). The *CDA* mutant contains more dCTP similar to a human cancer cell line with reduced CDA activity (Chabosseau
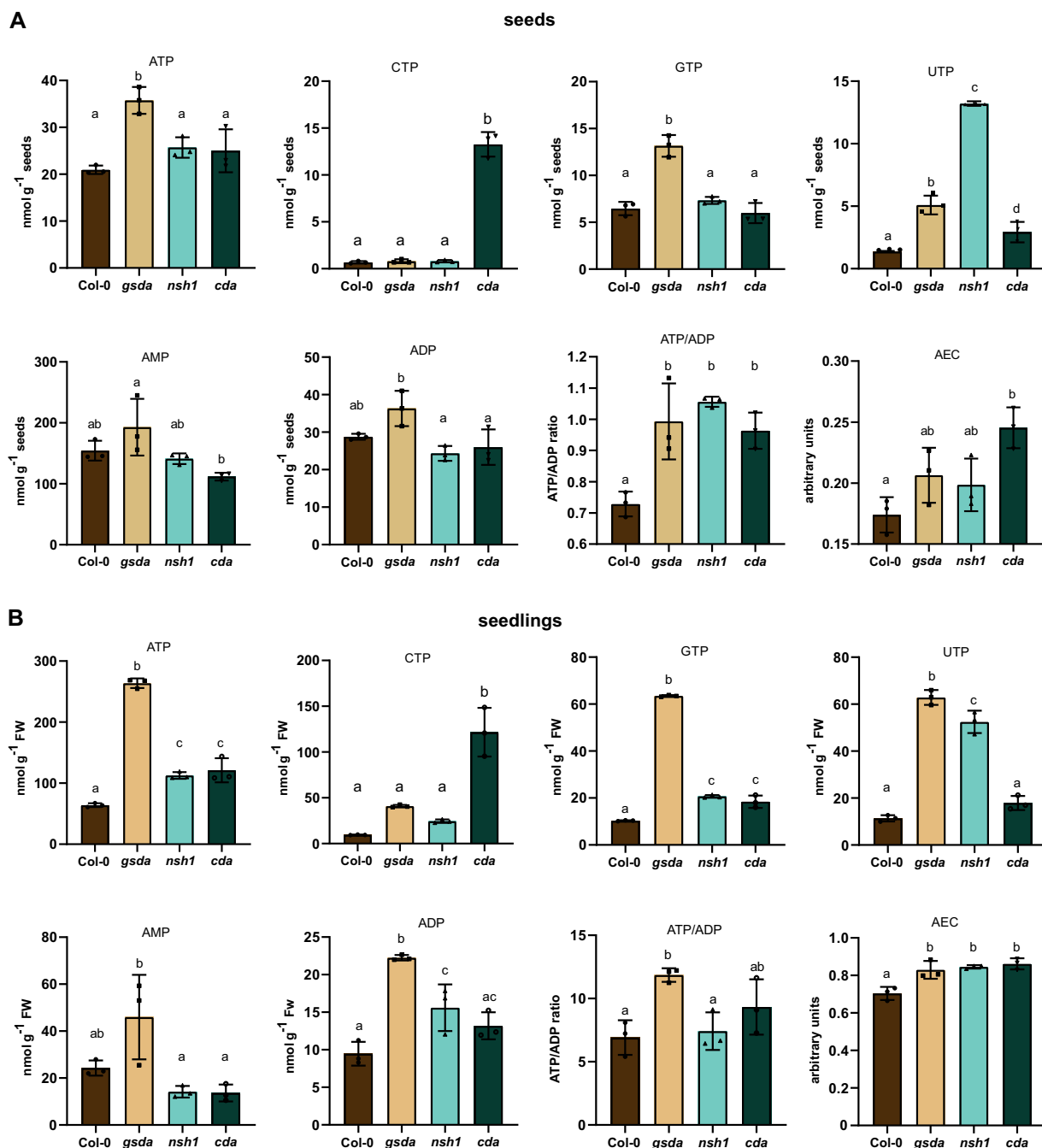
Figure 7 Comparison of the AEC, the ATP/ADP ratio, and the rNT concentrations between seeds and seedlings. Concentrations of rNTs, as well as ATP/ADP ratio and AEC in seeds (A) and in 7-day-old seedlings (B) of *A. thaliana* wild type, as well as mutants in the salvage and degradation of purine and pyrimidine Ns (*gsda*, guanosin deaminase). Error bars are SD, $n = 3$ biological replicates, for seeds, three independent seed pools derived from different mother plants, for seedlings, three pools of seedlings from three independent liquid cultures were used. Statistical analysis was performed using one-way ANOVA with Tukey's post hoc test. Different letters indicate $P < 0.05$. FW, fresh weight.

et al., 2011). The high dCTP concentration may be a consequence of the high CTP concentration in *cda* seedlings. Interestingly, *gsda* seedlings accumulated the highest amounts of dGTP and dATP and also contained more dCTP and dTTP than the wild type. It seems as if the strongly elevated concentrations of ATP, GTP, and CTP in this mutant are mirrored in the corresponding dNTPs indicating that the

pool sizes of these metabolites are directly linked. This also holds true for the higher dATP and dGTP concentrations mirroring the enlarged ATP and GTP pools in the *cda* and *nsh1* seedlings. In none of the mutants, deoxyguanosine was detected and deoxyadenosine concentrations were the same as in the wild type, whereas deoxycytidine was only found in *cda* background. It appears that the dN pool sizes of A,
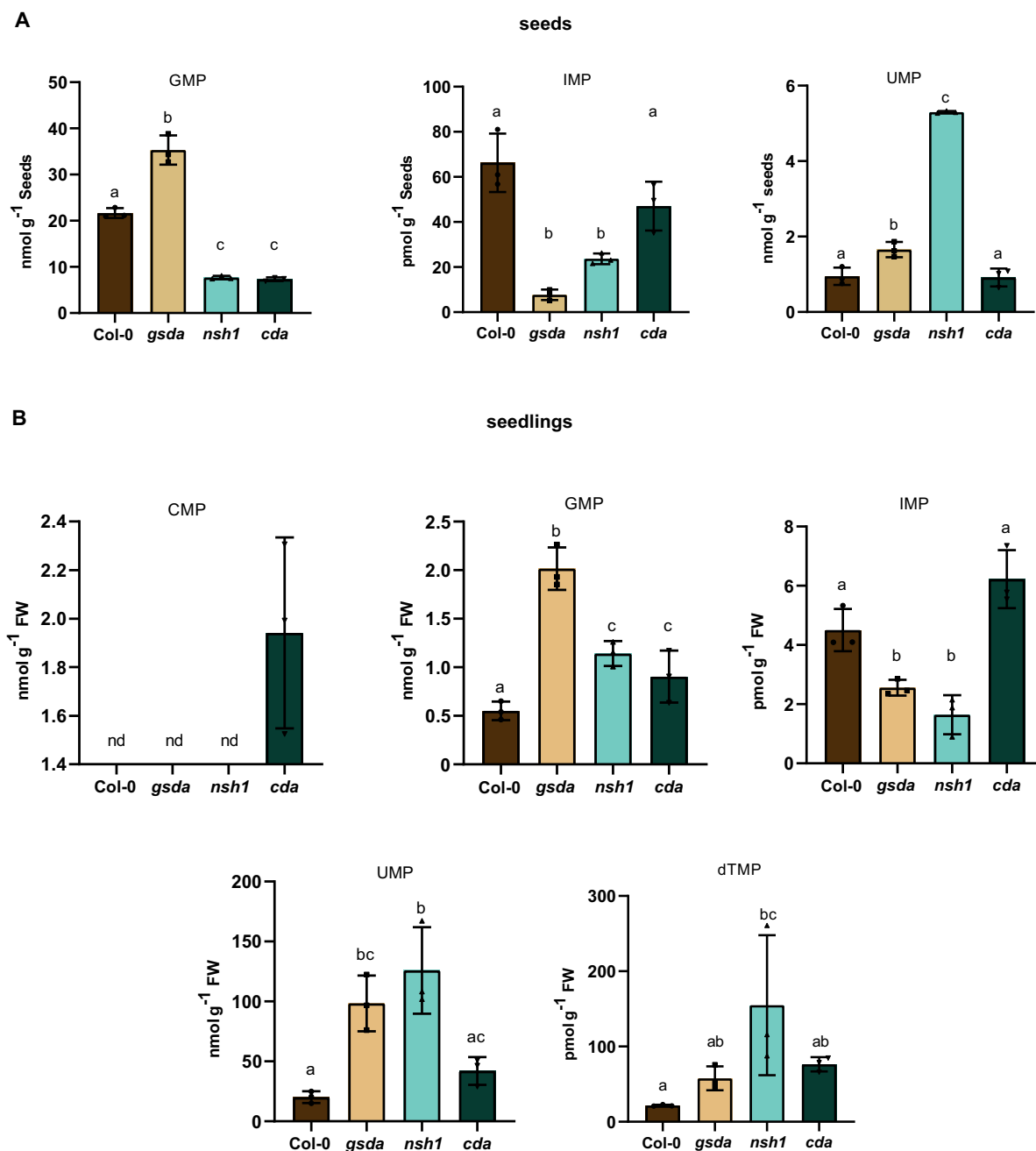
A **seeds**

B **seedlings**

Figure 8 Absolute quantification of NMPs in seeds and seedlings of wild type and mutant plants impaired in nucleoside catabolism. Concentrations of NMPs in seeds (A) and in 7-day-old seedlings (B) of *A. thaliana* wild type, as well as mutants in the degradation of purine and pyrimidine Ns (*gsda*, guanosin deaminase) . Error bars are SD, *n* = 3 biological replicates, for seeds, three independent seed pools derived from different mother plants, for seedlings, three pools of seedlings from three independent liquid cultures were used. Statistical analysis was performed using one-way ANOVA with Tukey's post hoc test. Different letters indicate $P < 0.05$. nd, not detected. FW, fresh weight.

G, and C do not influence the corresponding dNTP amounts. By contrast, the deoxythymidine content seems to have a stronger influence on the dTTP pool size than the UTP concentration. Seedlings of the *cda* and the *gsda* mutants both have more deoxythymidine (Figure 6), more dTMP (Figure 8), and correspondingly more dTTP (Figure 9) than the wild type, but the UTP amounts are only elevated in the *gsda* and not in the *cda* background. However, the

deoxythymidine and dTMP pools are largest in the *NSH1* mutant, but the dTTP concentration is only moderately increased, less than in *gsda* and *cda* seedlings. Although this seems contradictory, one needs to consider that uridine as well as deoxyuridine stemming from deoxycytidine deamination cannot be metabolized in *nsh1* background (uridine and resulting UMP accumulation are shown in Supplemental Figure S6 and Figure 8). The accumulation of
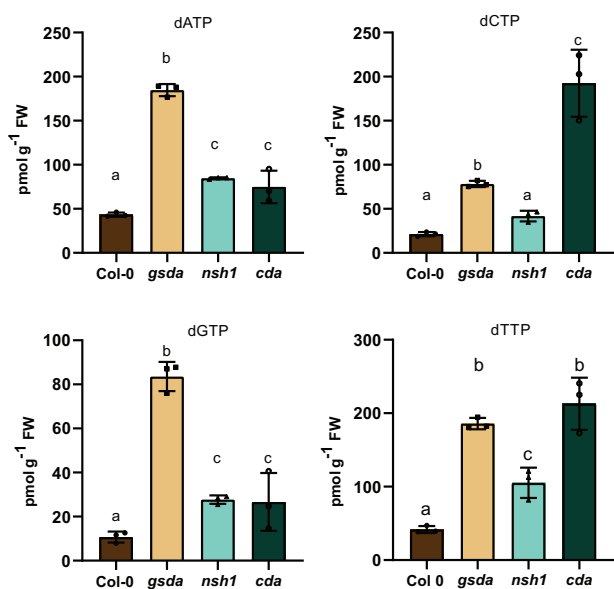
Figure 9 Absolute quantification of deoxynucleotide triphosphates in wild type and mutant seedlings impaired in nucleoside catabolism. Concentrations of dNTPs in 7-day-old seedlings of *A. thaliana* wild type, as well as mutants in the degradation of purine and pyrimidine Ns (*gsda*, guanosin deaminase). Error bars are SD, *n* = 3 biological replicates, for seedlings, three pools of seedlings from three independent liquid cultures were used. Statistical analysis was performed using one-way ANOVA with Tukey's post hoc test. Different letters indicate *P* < 0.05. FW, fresh weight.

these compounds may partially inhibit the thymidine kinases and maybe also thymidylate kinases known to be responsible for deoxythymidine (Clausen et al., 2012) and dTMP phosphorylation, respectively.

GSDA, CDA, and NSH1 were previously only considered to be involved in rN catabolism, but with our new method we can now show that they also play a role in dN degradation. By characterizing these mutants more in depth, we obtained a glimpse of the interconnection of deoxy- and rNT metabolism in plants. The data suggest that the ATP, GTP, and CTP pool sizes influence the amounts of dATP, dGTP, and dCTP, whereas the dTTP concentration seems to be more connected to the deoxythimidine than the UTP pool.

Discussion

A complex interplay of de novo synthesis, degradation, and salvage of rNTPs and dNTPs ensures that appropriate quantities of NTs are available for RNA and DNA synthesis as well as for signaling and energy metabolism (Zrenner et al., 2006; Witte and Herde, 2020). The fidelity of transcription and DNA replication is affected by altered ratios or changed total amounts of dNTPs (Nick McElhinny et al., 2010a, 2010b; Gon et al., 2011; Kumar et al., 2011; Buckland et al., 2014) emphasizing the importance of accurately balancing synthesis, degradation, and salvage of NTs and Ns. To gain a more comprehensive insight into these processes, a method allowing the parallel quantification of NTs and their

respective Ns from one sample is a prerequisite. For the analysis of plants, such a method has not been available so far. Here, we fill this gap showing that the developed protocol is suitable for the analysis of NTs and Ns not only in seedlings and fully grown plants of Arabidopsis but also in a wide phylogenetic range of plants and even algae (Figure 2 and Supplemental Figure S5).

We found that the dGTP/dCTP ratios were quite similar between plant species, whereas the dATP/dTTP ratios (Figure 3) and the absolute contents of the dNTPs (Supplemental Figure S5) were more variable. Although clearly more evidence is needed, it is tempting to speculate that dNTPs produced in excess of their stoichiometric requirement in some plants might have an additional role in the respective species. The variation in dTTP contents for example might be related to the synthesis of dTDP-sugars and enzymes synthesizing dTDP-sugars are known to exist in plants (Neufeld, 1962; Frydman et al., 1963; Katan and Avigad, 1966).

Our data suggest that genome GC contents and the sum of dGTP and dCTP concentrations relative to all dNTPs are correlated (Figure 4). A dNTP-dependent evolution of the genomic GC content has been proposed (Vetsigian and Goldenfeld, 2008; Greilhuber et al., 2012; Šmarda et al., 2014) which may be an example of metabolism-driven evolution (de Lorenzo, 2014), but actually it is unknown if the dNTP abundances influence the DNA composition or vice versa. Certainly, more plant species need to be analyzed to test the validity of this correlation.

By analyzing well-characterized mutants of Arabidopsis impaired in rN catabolism (*nsh1*, *cda*, and *gsda*) with the new method, we obtained novel insights especially regarding the involvement of the corresponding enzymes in dN and dNT metabolism (Figures 6–9 and Supplemental Figure S6). We demonstrate that CDA and NSH1 partake in the catabolism of dNs in vivo (Supplemental Figure S3), showing that the known ability of CDA to deaminate deoxycytidine in vitro (Chen et al., 2016) is of relevance in the plant. Deoxythymidine accumulates in *nsh1* background suggesting that it is a NSH1 substrate in vivo, but so far this has not been investigated with the isolated enzyme in vitro.

The accumulation of dNs in seedlings of wild type and mutant plants indicates that even in this actively growing tissue some DNA or dNT degradation occurs (Figure 6). One needs to bear in mind that the dN accumulation in the mutants was observed despite the ability of the cells to salvage dNs. Salvage can strongly reduce accumulation as was shown for hypoxanthine which only accumulated when degradation and salvage were mutated (Baccolini and Witte, 2019). It is therefore possible that the flux through the dNs is much higher than one would assume from the relatively moderate levels of accumulation in the catabolic mutants. The origin of the dNs is currently unclear. They might be derived from DNA repair or from programmed cell death resulting in DNA degradation and release of dNMPs. Programmed cell death is a widespread process associated

with many aspects of plant development such as endosperm degradation, tracheary element differentiation, senescence, or microbial interactions (Aoyagi et al., 1998; Sakamoto and Takami, 2014). It was even suggested that plants store NTs and phosphate for embryogenesis and germination in the DNA synthesized by endoreduplication within the endosperm (Wang et al., 1998; Leiva-Neto et al., 2004; Lee et al., 2009). Another putative source of dNs might be the continuous house-keeping dephosphorylation of NTs. It was previously hypothesized that the phosphorylation status especially of dNTs may be constantly changed to control dNTP quantity and to improve their quality, i.e. to ensure that the dNTP pools only contain dNTs with canonical bases (Rampazzo et al., 2010; Leija et al., 2016). Because nucleoside and NT kinases best recognize their canonical targets, they may serve to purify the dNT pools if the pools are simultaneously subject to continuous dephosphorylation (Chen et al., 2018; Chen and Witte, 2020). It follows from this idea that dN salvage must be an essential process, because without it the dNTP pools would be depleted. In agreement with this concept, salvage by thymidine kinase 1a and 1b (Pedroza-García et al., 2015, 2019; Xu et al., 2015; Le Ret et al., 2018) is crucial for plant development since plants lacking functional copies of both corresponding genes have etiolated seedlings that need carbohydrate supplementation for survival (Xu et al., 2015).

There is no doubt that salvage of deoxythymidine is important, but here we provide evidence that plants also degrade this compound employing NSH1. This raises the question of how degradation and salvage are coordinated. Partially this might be achieved by spatial separation of both processes because at least Tk1b is located in the mitochondria and the chloroplasts whereas NSH1 and Tk1a are found in the cytosol (Jung et al., 2009; Xu et al., 2015), suggesting a preference for salvage in the organelles and a competition of both processes in the cytosol.

Intriguingly, compromising N catabolism leads not only to N accumulation but results in increased and imbalanced NT pools (Figures 7–9). The N and NT pools are interconnected by kinases (Ashihara et al., 2020; Witte and Herde, 2020; Figure 5), thus the most straightforward way to explain the increases in rNTP abundances is the higher availability of rN substrates for these enzymes. By contrast, higher amounts of dNTPs in the mutants are in most cases probably not a result of phosphorylation of dNs, because their pool sizes are often not increased except for deoxythymidine. Instead increased dNTP levels might result from greater reduction of rN diphosphates by rNT reductase (RNR; Nordlund and Reichard, 2006).

The current literature on plant NT metabolism suggests a temporal or spatial separation of synthesis, salvage, and degradation to avoid futile cycles (Ashihara et al., 2020; Witte and Herde, 2020). However, the concomitant occurrence in the catabolic mutants of increased amounts of Ns and of NTs, indicative of salvage and biosynthetic processes, suggests that a seedling can realize all these processes

simultaneously. The enzymes for synthesis, degradation, and salvage might all be present in the same metabolic space but the flux through the respective pathways might be coordinated by the regulation of key enzymes. Being able to detect and quantify Ns and NTs is a prerequisite to investigate such hypotheses for example by using metabolic flux analysis and conditional mutants.

In the gsda mutant not only the direct substrate (guanosine) and its phosphorylated counterpart (GTP) accumulate but also ATP suggesting a crosstalk between these metabolites. It has been shown biochemically that GTP severely inhibits the activity of plant AMP deaminase, an enzyme that is involved in the regulation of guanylate synthesis, adenylate catabolism, and AEC regulation (Yabuki and Ashihara, 1991; Sabina et al., 2007; Witte and Herde, 2020). Sabina et al. (2007) could not detect an effect on GTP pools when feeding AMP deaminase inhibitors, but detected a two–fivefold increase of all adenylates, similar to the effects we observed in the GSDA mutant. This suggests that GTP accumulation in the gsda background (Figure 7) inhibits AMP deaminase in vivo, thereby raising ATP levels. Because IMP is the product of AMP deaminase, the inhibition of this enzyme might reduce IMP concentrations. That is precisely what we observed in the gsda seeds and seedlings (Figure 8). UTP and CTP levels were also raised in gsda plants probably to balance the elevated ATP and GTP concentrations. How pyrimidine and purine NT concentrations are kept in balance is currently unknown, but with our new method this interesting question could be further investigated.

Several phenotypes have been described for gsda, nsh1, and cda mutants such as delayed germination, compromised growth, and reduced recovery from dark treatment along with chlorosis (Chen et al., 2016; Schroeder et al., 2018; Baccolini and Witte, 2019). It has been suggested that these phenotypes are directly caused by the accumulation of rNs (Schroeder et al., 2018). However, here we show that NT pools are also disturbed in these mutants, which is known or is easily conceivable to cause detrimental effects. An increase in rNTP pool sizes for example leads to the undesired incorporation of rNTPs into human mtDNA (Nick McElhinny et al., 2010a, 2010b; Berglund et al., 2017). An elevated GTP concentration as observed in gsda seedlings might interfere with signaling processes in plants, in which GTP plays an important role (Assmann, 2002; Johnston et al., 2007). The biosynthesis of dNTPs by RNR will be disturbed by changed NTP pool sizes because RNR is stimulated by ATP and inhibited by dATP. Such pool size changes are known to lead to mutations and even to apoptosis in non-plant organisms (Kumar et al., 2011). Also in Arabidopsis, a mutant lacking a subunit of the RNR has disturbed dNTP levels and displays severe defects like sensitivity to UV-C light, DNA damage, developmental abnormalities, and cell death (Wang and Liu, 2006). Furthermore, it was shown that feeding deoxyadenosine to bean roots results in chromosomal breakage which was suggested to be a result of an increased dATP pool (Odmark and Kihlman, 1965).

In summary, it appears possible that the phenotypic alterations observed in the N catabolic mutants are at least partially caused by changed rNTP and dNTP pool sizes. It would be interesting to use these mutants to study the consequences of altered NT levels on transcriptional fidelity, DNA mutation rates, effects on the target of rapamycin (TOR) complex, and the involvement of dATP or ATP in plant immune signaling (Burdett et al., 2019; Nizam et al., 2019; Kazibwe et al., 2020).

NTs and Ns are not only involved in RNA and DNA metabolism but also play diverse roles for example in developmental processes and in plant pathogen interactions. The recently described *venosa4* mutant for example shows severe defects in chloroplast development and has a defect in an enzyme, which is likely involved in the dephosphorylation of dNTPs (Xu et al., 2020). In plant metabolome studies, NTs and Ns are strongly underrepresented hampering the discovery of novel functions for this major class of plant metabolites. The method described here will allow routine NT analysis to be performed in research focused on NT metabolism as well as in broad plant metabolomics surveys used in many plant research disciplines with the prospect of revealing so far unknown connections between the NT metabolome and other biological processes.

Materials and methods

Chemicals

Water, acetonitrile, methanol, ammonium acetate (all LC–MS grade), ethylenediaminetetraacetic acid (EDTA), and magnesium chloride were purchased from AppliChem. PCA and DCM were obtained from Carl Roth. 2′-Deoxyadenosine, 2′-deoxycytidine, 2′-deoxyguanosine, TCA, and trioctylamine (TOA) were purchased from Sigma–Aldrich. Twenty-five percent of ammonia solution and acetic acid (both LC–MS grade) were purchased from Merck. All ISTDs were from Eurisotope. The Strata-X-AW SPE cartridges, 30 mg, 33 μm were bought from Phenomenex.

Preparation of standard solutions

If not stated otherwise, all standards were measured in the matrix obtained after SPE. All stock solutions were stored at −80°C and dilutions were prepared fresh before analysis.

Plant culture

Arabidopsis thaliana plants were grown as previously stated in Niehaus et al. (2020) with slight modifications. Seeds were surface sterilized and cultivated in liquid culture (1.5 mM $MgSO_4 \times 7H_2O$, 1.25 mM KH_2PO_4, 3 mM $CaCl_2$, 18.7 mM KNO_3, 0.1 mM $FeSO_4 \times 7H_2O$, 0.1 mM $Na_2EDTA \times 2H_2O$, 0.13 mM $MnSO_4 \times H_2O$, 0.1 mM H_3BO_3, 30 μM $ZnSO_4 \times 7H_2O$, 1 μM $Na_2MoO_4 \times 2H_2O$, 0.1 μM $CuSO_4 \times 5H_2O$, 0.1 μM $NiCl_2 \times 6H_2O$, 0.125% [w/v] MES, pH 5.7 adjusted with KOH) in a shaker (New Brunswick Innova 42, Eppendorf) with six Sylvania e15t8 tubular fluorescent lamps emitting a photon flux of 45 μmol s^{-1} m^{-2} at 22°C in 100 mL flasks under sterile conditions (10 mg seeds per flask). The shaker

was set to 80 rpm. The seedlings were harvested after 7 days. For the 33-day-old plants, Arabidopsis seeds were sown on soil and grown under long-day conditions (Binder KBFW 720 with Osram Lumilux lights,16-h light/8-h darkness, 22°C day, 20°C night, 100 μmol s^{-1} m^{-2} light, and 70% humidity).

T-DNA insertion mutants were obtained from our in-house collection. Their characterization is described in Dahncke and Witte (2013; *gsda*-2, GK432D08), Chen et al. (2016; *cda*-2, SALK036597), and Baccolini and Witte (2019; *nsh1*-1, SALK083120). The genotypes of all mutant lines were confirmed by PCR as described previously (Dahncke and Witte, 2013; Chen et al., 2016). A uniform seed batch was obtained from mutant and wild-type plants grown in parallel in a randomized fashion. The seeds were analyzed 2 weeks post-harvest. For the comparison of mutant seedlings, plants were grown in liquid culture as described above under constant light.

Seeds of asparagus (*Asparagus officinalis* cv. Ramires), barley (*Hordeum vulgare* cv. Golden Promise), common bean (*Phaseolus vulgaris* cv. Black Jamapa), oat (*A. sativa* cv. Fleuron), rice (*O. sativa* cv. Nipponbare), and wheat (*Triticum aestivum* cv. Thatcher) were surface sterilized and cultivated between sheets of filter paper placed in a container filled with some distilled water (Kirchner et al., 2018) and grown in a growth cabinet (Binder KBFW 720 with Osram Lumilux lights) at 22°C, 100 μmol s^{-1} m^{-2} light, and 70% humidity. Whole seedlings were collected 7 days after germination (dag) for all species except *A. officinalis* and *O. sativa* which needed 14 dag to accumulate enough biomass.

The moss *Physcomitrium* (*Physcomitrella*) *patens*, strain Grandsen 2004 (Kamisugi et al., 2008), was cultivated on Knoop medium (250 mg L^{-1} KH_2PO_4, 250 mg L^{-1} KCl, 250 mg L^{-1} $MgSO_4 \times 7H_2O$, 1 g L^{-1} $Ca(NO_3)$, 12.5 mg L^{-1} $FeSO_4 \times 7H_2O$, pH 5.8, and 1.2% [w/v] agar) in a climate chamber under long-day conditions (16-h light/8-h darkness, 22°C day, 20°C night, 100 μmol s^{-1} m^{-2} light, and 70% humidity). Plants were transferred to new plates every 4 weeks during cultivation. Plant material was collected 7 days after transferring to new plates.

Tomato plants (*S. lycopersicum* cv. Micro-Tom) were surface sterilized and transferred to germination medium (0.5 × Murashige and Skoog, 10 g L^{-1} sucrose, and 8 g L^{-1} phyto-agar). The seeds were placed in darkness at room temperature and moved to long-day conditions after 4 days (16-h light/8-h darkness, 22°C day, 20°C night, 100 μmol s^{-1} m^{-2} light, and 70% humidity). Seedlings were harvested 7 dag. Maize (*Zea mays* cv. Rafinio) was grown hydroponically between two plates in tap water wetted foam under long-day conditions (16-h light/8-h darkness, 22°C day, 20°C night, 100 μmol s^{-1} m^{-2} light, and 70% humidity) and harvested 7 dag. *C. reinhardtii* was cultivated for 7 days under the same conditions as the Arabidopsis seeds in liquid culture. *Volvox carteri* was cultivated for 7 days in Fernbach flasks (28°C, 16-h light/8-h dark and 100 μmol s^{-1} m^{-2} light), as described in Klein et al. (2017). *Mougeotia scalaris*

H. Straube et al.

strain SAG 164.80 was cultivated for 7 days (22–25°C, 16-h light/8-h dark, and 46 μmol s^{-1} m^{-2} light), as described in Regensdorff et al. (2018).

All plant and algae material was collected in the middle of the respective light period. Genomic GC-content of respective plant and algae species was acquired from the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov) or in the case of *M. scalaris* from Regensdorff et al. (2018).

Chromatography and MS parameters

An Agilent 1290 Infinity II LC System coupled with an Agilent 6470 triple quadrupole mass spectrometer was used. Chromatographic separations employed either a 150 × 2.1 mm zic-cHILIC column with 3-μm particle size (Merck) or a 50 × 4.6 mm Hypercarb column with 5-μm particle size (Thermo scientific). The zic-cHILIC column was operated at a flowrate of 0.2 mL min^{-1} and a temperature of 35°C. Mobile phase A was 90% 10 mM ammonium acetate pH 7.7 with 10% acetonitrile and mobile phase B was 10% 2.5 mM ammonium acetate pH 7.7 with 90% acetonitrile. In the following, we refer to chromatographies employing the zic-cHILIC column as cHILIC method. Depending on the analytes, different gradients were used (Tables 7, 8). The Hypercarb column was operated at a flow rate of 0.6 mL min^{-1} and a column temperature of 30°C. Mobile phase A was 5 mM ammonium acetate pH 9.5 and mobile phase B was acetonitrile (for the gradient see Table 9). In the following, we refer to the chromatography using the Hypercarb

Table 7 Gradient for NT chromatography on the zic-cHILIC column

Time (min)	Mobile phase A (%)	Mobile phase B (%)
0.0	20.0	80.0
7.0	20.0	80.0
12.0	40.0	60.0
17.0	40.0	60.0
19.0	20.0	80.0
23.0	20.0	80.0

Table 8 Gradient for nucleoside chromatography on the zic-cHILIC column

Time (min)	Mobile phase A (%)	Mobile phase B (%)
0.0	10.0	90.0
7.0	10.0	90.0
12.0	40.0	60.0
17.0	40.0	60.0
19.0	10.0	90.0
23.0	10.0	90.0

Table 9 Gradient for NT chromatography on the Hypercarb column

Time (min)	Mobile phase A (%)	Mobile phase B (%)
0.0	96.0	4.0
10.0	70.0	30.0
10.10	0.0	100.0
11.50	0.0	100.0
11.60	96.0	4.0
20.0	96.0	4.0

column as the Hypercarb method. The injection volume was 10 μL and analysis was carried out in positive mode for both methods employing the multiple-reaction-monitoring (MRM) mode. Transitions (precursor ions and product ions) as well as collision energies and fragmentor energies are listed in Supplemental Tables S4–S6. The in-source parameters for the cHILIC method were optimized according to Kong et al. (2018) as a starting point. The following adjustments to fit the plant matrix were made: gas temperature 290°C, gas flow 13 L min^{-1}, nebulizer pressure 25 psi, sheath gas temperature 320°C, sheath gas flow 11 L min^{-1}, capillary voltage 2,500 V, and nozzle voltage 2,000 V. The optimized in-source parameters for the Hypercarb method were: gas temperature 250°C, gas flow 12 L min^{-1}, nebulizer pressure 20 psi, sheath gas temperature 395°C, sheath gas flow 12 L min^{-1}, capillary voltage 3,000 V, and nozzle voltage 500 V. For the determination of exact masses, samples were separated with a Vanquish LC (Thermo Fisher) by the Hypercarb method. The metabolites were analyzed with an Orbitrap Q Exactive Plus mass spectrometer (Thermo Fisher) at a resolution of 70,000 when operated in full MS mode or 17,500 for detection of product ions in PRM (parallel reaction monitoring) mode with 35-V normalized collision energy or 140,000 using single-ion monitoring (SIM) in the positive polarity mode. Automatic gain control (AGC) target and maximum injection time were set to 3e^6 and 200 ms, respectively. The heated ESI (electrospray-ionization) source was operated at 0-eV collision-induced dissociation (CID), sheath gas flow 45, auxiliary gas flow 10, sweep gas flow 2, spray voltage 3.5 kV, capillary temperature 250°C, S-lens RF level 45.0, and aux gas heater 400°C. All reported values were obtained with the Freestyle software (ver. 1.5, Thermo Fisher). For UV detection, samples were prepared as described in the respective section and identical to the procedure described for the analysis of the MEF. The detector was a VF-D40-A variable wavelength detector set to 254 nm. Identity of peaks was confirmed by MS as described before with an orbitrap mass analyzer in full MS mode.

Sample preparation and SPE

Plant samples were harvested, briefly washed in tap water, and dried thoroughly with a paper towel. Approximately 100-mg plant material was weighed into a 2-mL safe-lock centrifuge-vial and frozen in liquid nitrogen together with five 5-mm steel beads (one 7-mm steel bead and five 5-mm steel beads in case of seeds). The exact sample weight was noted and used for calculating analyte concentrations. The tissue was disrupted using a MM 400 beadmill (Retsch, Germany) at 28 Hz for 2.30 min. Onto the frozen powder, 1 mL of ice-cold 15% TCA solution was added including the respective isotope standards. The samples were briefly vortexed, ground once more at 28 Hz for 2.30 min, and then centrifuged for 10 min at 4°C at 40,000 × *g*. To the supernatant 1 mL 78/22 DCM/TOA was added. Samples were vortexed for 12 s and centrifuged for 2 min at 4°C and 5,000 × *g*. The upper phase from each sample was transferred to a new tube and 1 mL water as well as 5 μL 0.5%

39

acetic acid were added. This mixture was applied to a 30 mg/1 mL Strata X-AW cartridge that had been equilibrated sequentially with 1 mL methanol, 1 mL 2/25/73 formic acid/methanol/H_2O, and 1 mL 10 mM ammonium acetate pH 4.5. The sample solution was allowed to enter the cartridge for 2 min without suction and then percolated through the solid phase at a flowrate of about 1 mL min^{-1} using a vacuum manifold. The flow-through containing the nucleoside fraction was collected and evaporated in an Alpha 1–2 LDplus freeze dryer (Christ, Germany). The SPE cartridge was washed with 1 mL 1-mM ammonium acetate pH 4.5 and 1 mL methanol, dried shortly, and eluted two times with 0.5 mL 20/80 ammonia/methanol. The eluate was dried in a vacuum concentrator until no liquid was left. For the cHILIC method, the samples were reconstituted in 50 μL 30/9/1 acetonitrile/water/100 mM ammonium acetate pH 7.7. For the Hypercarb method, the samples were reconstituted in 50 μL 95/5 of 5-mM ammonium acetate (pH 9.5)/acetonitrile.

Method validation

We analyzed the developed method in terms of linearity, precision, LLOQ, relative recovery, and matrix effects. For calibration curves, peak area sums of ISTDs were plotted against their concentration. Concentrations were chosen in a range relevant in biological samples. Intra-day precision was calculated using the peak area sums of different concentrations of ISTDs in SPE-matrix, injecting the same sample three times a day. The inter-day precision was calculated using the peak area sums of different concentrations of ISTDs in SPE-matrix that have been injected on 3 days consecutively. Samples were stored at 4°C. The LLOQ was defined as the lowest concentrated standard with acceptable peak shape. The relative recovery was calculated as following:

$$\frac{\text{area of isotope standard added to extraction buffer}}{\text{area of isotope standard added after SPE}} \times 100 = \text{relative recovery } (\%).$$

The matrix effect was determined by comparing ISTDs with the same concentration separated either by the cHILIC method or the Hypercarb method in (i) matrix after extracting the TCA with the DCM/TOA, (ii) in matrix after the SPE procedure, and (iii) in pure buffer (for the cHILIC method 30/9/1 acetonitrile/water/100 mM ammonium acetate pH 7.7 and for the Hypercarb method 95/5 of 5-mM ammonium acetate [pH 9.5]/acetonitrile). The MEF was calculated according to Zhou et al. (2017):

$$\frac{\text{average area ISTD in buffer} - \text{average area of ISTD in matrix}}{\text{average area of ISTD in buffer}} \times 100 = \text{MEF}.$$

Quantification of metabolites

The amount of metabolites in plant samples was calculated either by the isotope dilution technique or with external

calibration curves in SPE-matrix. ISTDs were added to the extraction buffer prior to extraction. Only calibration curves with a coefficient of determination (R^2) > 0.99 were accepted.

Calculation of concentrations of metabolites in plant cell compartments

To determine the concentrations of rNTPs and dNTPs per unit NT-containing cell volume, we used the calculations of Koffler et al. (2013). They determined the total volume of the mesophyll cells in four sections of an Arabidopsis leaf and the volumes of the individual subcellular compartments in these sections. We grew plants until the leaves met the morphological criteria stated by Koffler and colleagues, i.e. to a size of approximately 1.5 × 3.0 cm, and then extracted them. We determined an average (i.e. not in sections but for the whole) leaf mesophyll volume per unit total fresh weight (AMV) of 600.0 μL g^{-1} fresh weight. We assumed that the nucleus, the cytoplasm, the mitochondria, and the plastids contain NTs (these are NT-containing compartments, NCC), whereas other cell compartments are likely devoid of relevant concentrations of NTs. An average of the percentile volumes of the compartments was calculated. Together, the NCCs accounted for 26.86% of the total mesophyll volume. We then calculated the concentrations of dNTPs and rNTPs in the NCC (Y in the formula), while X is the amount of metabolite in pmol g^{-1} fresh weight that was measured:

$$\frac{X \left[\frac{\text{pmol}}{\text{g}}\right]}{\text{AVM} \left[\frac{\mu l}{\text{g}}\right]} \times \frac{100\%}{26.86\%} = Y \, [\mu M].$$

Statistical analysis

Statistical analysis was performed using Prism 8 software. One-way analysis of variance (ANOVA) with Tukey's post test or two-way ANOVA with Sidak's post test were used. Different letters or a star indicate differences at significance level of $P < 0.5$. Statistical analysis results are shown in Supplemental File S1.

Accession numbers

Information regarding used mutants can be found in the GenBank/EMBL data libraries under the following accession numbers: *cda* (At2g19570), *gsda* (At5g28050), *nsh1* (At2g36310).

Supplemental data

Supplemental Figure S1. Chromatograms of all analyzed NTs by hypercarb chromatography.

Supplemental Figure S2. Chromatograms of all analyzed NTs by cHILIC chromatography.

Supplemental Figure S3. Chromatograms of all analyzed Ns by cHILIC chromatography.

Supplemental Figure S4. Analysis of NTs in Arabidopsis leaves by HPLC and UV detection.

Supplemental Figure S5. SPE enables detection of dNTPs in several plant and algae species.

Supplemental Figure S6. SPE enables detection of rNs in Arabidopsis seeds and seedlings.

Supplemental Table S1. Literature survey on NT analysis in plants

Supplemental Table S2. Tested conditions for SPE optimization

Supplemental Table S3. Relative recovery of metabolites extracted by different methods.

Supplemental Table S4. Precursor, transitions (quantifier, qualifier), fragmentor, collision energy and retention times for NTP measurements

Supplemental Table S5. Precursor, transitions (quantifier, qualifier), fragmentor energy, collision energy and retention times for **NMP** measurements

Supplemental Table S6. Precursor, transitions (quantifier, qualifier), fragmentor energy, collision energy and retention times for nucleoside measurements

Supplemental Table S7. Intra-day variation and inter-day variation for different concentrations of NTs seperated by the Hypercarb or cHILIC method

Supplemental Table S8. Intra-day variation and inter-day variation for different concentrations of Ns seperated by the cHILIC method

Supplemental Table S9. Determination of MEF at 5- and 50-pmol dNTP and rNTP ISTDs, respectively, on column

Supplemental Table S10. Determination of MEF at 5-pmol nucleoside ISTD on column

Supplemental Table S11. NTPs analyzed with an Orbitrap mass spectrometer

Supplemental File S1. ANOVA tables.

Acknowledgments

Funding

References

Aoyagi S, Sugiyama M, Fukuda H (1998) BEN1 and ZEN1 cDNAs encoding S1-type DNases that are associated with programmed cell death in plants. FEBS Lett **429**: 134–138

Ashihara H, Crozier A, Ludwig IA (2020) Plant Nucleotide Metabolism: Biosynthesis, Degradation, and Alkaloid Formation. Wiley Blackwell, Chichester

Ashihara H, Mitsui K, Ukaji T (1987) A simple analysis of purine and pyrimidine nucleotides in plant cells by high-performance liquid chromatography. Z Naturforsch C **42**: 297–299

Assmann SM (2002) Heterotrimeric and unconventional GTP binding proteins in plant cell signaling. Plant Cell **14**: S355–S373

Baccolini C, Witte C-P (2019) AMP and GMP catabolism in arabidopsis converge on Xanthosine, which is degraded by a nucleoside hydrolase heterocomplex. Plant Cell **31**: 734–751

Berglund A-K, Navarrete C, Engqvist MKM, Hoberg E, Szilagyi Z, Taylor RW, Gustafsson CM, Falkenberg M, Clausen AR (2017) Nucleotide pools dictate the identity and frequency of ribonucleotide incorporation in mitochondrial DNA. PLoS Genet **13**: e1006628

Bieleski RL (1964) The problem of halting enzyme action when extracting plant tissues. Anal Biochem **9**: 431–442

Bieleski RL, Young RE (1963) Extraction and separation of phosphate esters from plant tissues. Anal Biochem **6**: 54–68

Buckland RJ, Watt DL, Chittoor B, Nilsson AK, Kunkel TA, Chabes A (2014) Increased and imbalanced dNTP pools symmetrically promote both leading and lagging strand replication infidelity. PLoS Genet **10**: e1004846

Burdett H, Bentham AR, Williams SJ, Dodds PN, Anderson PA, Banfield MJ, Kobe B (2019) The plant "resistosome": structural insights into immune signaling. Cell Host Microbe **26**: 193–201

Carrari F, Coll-Garcia D, Schauer N, Lytovchenko A, Palacios-Rojas N, Balbo I, Rosso M, Fernie AR (2005) Deficiency of a plastidial adenylate kinase in Arabidopsis results in elevated photosynthetic amino acid biosynthesis and enhanced growth. Plant Physiol **137**: 70–82

Castroviejo M, Tharaud D, Mocquot B, Litvak S (1979) Factors affecting the onset of deoxyribonucleic acid synthesis during wheat embryo germination: study of the changes in DNA polymerases A, B and C and the pool of DNA precursors. Biochem J **181**: 193–199

Chabosseau P, Buhagiar-Labarchède G, Onclercq-Delic R, Lambert S, Debatisse M, Brison O, Amor-Guéret M (2011) Pyrimidine pool imbalance induced by BLM helicase deficiency contributes to genetic instability in Bloom syndrome. Nat Commun **2**: 368

Chen M, Herde M, Witte C-P (2016) Of the nine cytidine deaminase-like genes in arabidopsis, eight are pseudogenes and only one is required to maintain pyrimidine homeostasis in vivo. Plant Physiol **171**: 799–809

Chen M, Urs MJ, Sánchez-González I, Olayioye MA, Herde M, Witte C-P (2018) m6A RNA degradation products are catabolized by an evolutionarily conserved N6-methyl-AMP deaminase in plant and mammalian cells. Plant Cell **30**: 1511–1522

Chen M, Witte C-P (2020) A kinase and a glycosylase catabolize pseudouridine in the peroxisome to prevent toxic pseudouridine monophosphate accumulation. Plant Cell **32**: 722–739

Clausen AR, Girandon L, Ali A, Knecht W, Rozpedowska E, Sandrini MPB, Andreasson E, Munch-Petersen B, Piškur J (2012) Two thymidine kinases and one multisubstrate deoxyribonucleoside kinase salvage DNA precursors in *Arabidopsis thaliana*. FEBS J **279**: 3889–3897

Cohen S, Megherbi M, Jordheim LP, Lefebvre I, Perigaud C, Dumontet C, Guitton J (2009) Simultaneous analysis of eight nucleoside triphosphates in cell lines by liquid chromatography coupled with tandem mass spectrometry. J Chromatogr B **877**: 3831–3840

Dahncke K, Witte C-P (2013) Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in Arabidopsis. Plant Cell **25**: 4101–4109

Dietmair S, Timmins NE, Gray PP, Nielsen LK, Krömer JO (2010) Towards quantitative metabolomics of mammalian cells:

development of a metabolite extraction protocol. Anal Biochem **404**: 155–164

Dubois E, Córdoba-Cañero D, Massot S, Siaud N, Gakière B, Domenichini S, Guérard F, Roldan-Arjona T, Doutriaux M-P (2011) Homologous recombination is stimulated by a decrease in dUTPase in Arabidopsis. PLoS ONE **6**: e18658

Dutta I, Dutta PK, Smith DW, O'Donovan GA (1991) High-performance liquid chromatography of deoxyribonucleoside di- and triphosphates in tomato roots. J Chromatogr A **536**: 237–243

Faivre-Nitschke SE, Grienenberger JM, Gualberto JM (1999) A prokaryotic-type cytidine deaminase from *Arabidopsis thaliana* gene expression and functional characterization. Eur J Biochem **263**: 896–903

Farrow SC, Emery RN (2012) Concurrent profiling of indole-3-acetic acid, abscisic acid, and cytokinins and structurally related purines by high-performance-liquid-chromatography tandem electrospray mass spectrometry. Plant Methods **8**: 42

Feller W, Schimpf-Weiland G, Follmann H (1980) Deoxyribonucleotide biosynthesis in synchronous algae cells. Eur J Biochem **110**: 85–92

Frydman RB, Neufeld EF, Hassid WZ (1963) Thymidine diphosphate d-galactose pyrophosphorylase of *Phaseolus aureus*. Biochim Biophys Acta **77**: 332–334

Galmarini CM, Jordheim L, Dumontet C (2003) Pyrimidine nucleoside analogs in cancer treatment. Exp Rev Anticancer Ther **3**: 717–728

Garton S, Knight H, Warren GJ, Knight MR, Thorlby GJ (2007) Crinkled leaves 8—a mutation in the large subunit of ribonucleotide reductase—leads to defects in leaf development and chloroplast division in *Arabidopsis thaliana*. Plant J Cell Mol Biol **50**: 118–127

Gon S, Napolitano R, Rocha W, Coulon S, Fuchs RP (2011) Increase in dNTP pool size during the DNA damage response plays a key role in spontaneous and induced-mutagenesis in *Escherichia coli*. Proc Natl Acad Sci U S A **108**: 19311–19316

Greilhuber J, Dolezel J, Leitch IJ (2012) Plant Genomes, Their Residents, and Their Evolutionary Dynamics, Springer, Wien

Guérard F, Pétriacq P, Gakière B, Tcherkez G (2011) Liquid chromatography/time-of-flight mass spectrometry for the analysis of plant samples: a method for simultaneous screening of common cofactors or nucleotides and application to an engineered plant line. Plant Physiol Biochem **49**: 1117–1125

Guo S, Duan J-A, Qian D, Wang H, Tang Y, Qian Y, Wu D, Su S, Shang E (2013) Hydrophilic interaction ultra-high performance liquid chromatography coupled with triple quadrupole mass spectrometry for determination of nucleotides, nucleosides and nucleobases in *Ziziphus* plants. J Chromatogr A, **1301**: 147–155

Harmenberg J, Karlsson AHJ, Gilljam G (1987) Comparison of sample preparation methods for the high-performance liquid chromatographic analysis of cell culture extracts for triphosphate ribonucleosides and deoxyribonucleosides. Anal Biochem **161**: 26–31

Henneré G, Becher F, Pruvost A, Goujard C, Grassi J, Benech H (2003) Liquid chromatography–tandem mass spectrometry assays for intracellular deoxyribonucleotide triphosphate competitors of nucleoside antiretrovirals. J Chromatogr B **789**: 273–281

Ikuma H, Tetley RM (1976) Possible interference by an acid-stable enzyme during the extraction of nucleoside di- and triphosphates from higher plant tissues. Plant Physiol **58**: 320–323

Johnston CA, Taylor JP, Gao Y, Kimple AJ, Grigston JC, Chen J-G, Siderovski DP, Jones AM, Willard FS (2007) GTPase acceleration as the rate-limiting step in Arabidopsis G protein-coupled sugar signaling. Proc Natl Acad Sci U S A **104**: 17317–17322

Jung B, Flörchinger M, Kunz H-H, Traub M, Wartenberg R, Jeblick W, Neuhaus HE, Möhlmann T (2009) Uridine-ribohydrolase is a key regulator in the uridine degradation pathway of Arabidopsis. Plant Cell **21**: 876–891

Jung B, Hoffmann C, Möhlmann T (2011) Arabidopsis nucleoside hydrolases involved in intracellular and extracellular degradation of purines. Plant J Cell Mol Biol **65**: 703–711

Kafer C, Thornburg RW (2000) *Arabidopsis thaliana* cytidine deaminase 1 shows more similarity to prokaryotic enzymes than to eukaryotic enzymes. J Plant Biol **43**: 162–170

Kamisugi Y, von Stackelberg M, Lang D, Care M, Reski R, Rensing SA, Cuming AC (2008) A sequence-anchored genetic linkage map for the moss, *Physcomitrella patens*. Plant J Cell Mol Biol **56**: 855–866

Katahira R, Ashihara H (2006) Role of adenosine salvage in wound-induced adenylate biosynthesis in potato tuber slices. Plant Physiol Biochem **44**: 551–555

Katan R, Avigad G (1966) NADP dependent oxidation of TDP-glucose by an enzyme system from sugar beets. Biochem Biophys Res Commun **24**: 18–24

Kazibwe Z, Soto-Burgos J, MacIntosh GC, Bassham DC (2020) TOR mediates the autophagy response to altered nucleotide homeostasis in a ribonuclease mutant. J Exp Bot eraa410

Khym JX (1975) An analytical system for rapid separation of tissue nucleotides at low pressures. Clin Chem **21**: 1245–1252

Kirchner TW, Niehaus M, Rössig KL, Lauterbach T, Herde M, Küster H, Schenk MK (2018) Molecular background of pi deficiency-induced root hair growth in *Brassica carinata*—a fasciclin-like arabinogalactan protein is involved. Front Plant Sci **9**: 1372

Klein B, Wibberg D, Hallmann A (2017) Whole transcriptome RNA-Seq analysis reveals extensive cell type-specific compartmentalization in *Volvox carteri*. BMC Biol **15**: 1–22

Koffler BE, Bloem E, Zellnig G, Zechmann B (2013) High resolution imaging of subcellular glutathione concentrations by quantitative immunoelectron microscopy in different leaf areas of Arabidopsis. Micron **45**: 119–128

Kong Z, Jia S, Chabes AL, Appelblad P, Lundmark R, Moritz T, Chabes A (2018) Simultaneous determination of ribonucleoside and deoxyribonucleoside triphosphates in biological samples by hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry. Nucl Acids Res **46**: e66

Kopecná M, Blaschke H, Kopecny D, Vigouroux A, Koncitíková R, Novák O, Kotland O, Strnad M, Moréra S, von Schwartzenberg K (2013) Structure and function of nucleoside hydrolases from *Physcomitrella patens* and maize catalyzing the hydrolysis of purine, pyrimidine, and cytokinin ribosides. Plant Physiol **163**: 1568–1583

Kumar D, Abdulovic AL, Viberg J, Nilsson AK, Kunkel TA, Chabes A (2011) Mechanisms of mutagenesis in vivo due to imbalanced dNTP pools. Nucl Acids Res **39**: 1360–1371

Kuskovsky R, Buj R, Xu P, Hofbauer S, Doan MT, Jiang H, Bostwick A, Mesaros C, Aird KM, Snyder NW (2019) Simultaneous isotope dilution quantification and metabolic tracing of deoxyribonucleotides by liquid chromatography high resolution mass spectrometry. Anal Biochem **568**: 65–72

Le Ret M, Belcher S, Graindorge S, Wallet C, Koechler S, Erhardt M, Williams-Carrier R, Barkan A, Gualberto JM (2018) Efficient replication of the plastid genome requires an organellar thymidine kinase. Plant Physiol **178**: 1643–1656

Lee HO, Davidson JM, Duronio RJ (2009) Endoreplication: polyploidy with purpose. Genes Dev **23**: 2461–2477

Leija C, Rijo-Ferreira F, Kinch LN, Grishin NV, Nischan N, Kohler JJ, Hu Z, Phillips MA (2016) Pyrimidine salvage enzymes are essential for de novo biosynthesis of deoxypyrimidine nucleotides in *Trypanosoma brucei*. PLoS Pathogens **12**: e1006010

Leiva-Neto JT, Grafi G, Sabelli PA, Dante RA, Woo Y-M, Maddock S, Gordon-Kamm WJ, Larkins BA (2004) A dominant negative mutant of cyclin-dependent kinase A reduces endoreduplication but not cell size or gene expression in maize endosperm. Plant Cell **16**: 1854–1869

42

Liu B, Winkler F, Herde M, Witte C-P, Großhans J (2019) A Link between deoxyribonucleotide metabolites and embryonic cell-cycle control. Curr Biol 29: 1187–1192.e3

de Lorenzo V (2014) From the selfish gene to selfish metabolism: revisiting the central dogma. BioEssays 36: 226–235

Meyer R, Wagner KG (1985) Determination of nucleotide pools in plant tissue by high-performance liquid chromatography. Anal Biochem 148: 269–276

Neufeld EF (1962) Formation and epimerization of dTDP-D-galactose catalyzed by plant enzymes. Biochem Biophys Res Commun 7: 461–466

Nick McElhinny SA, Kumar D, Clark AB, Watt DL, Watts BE, Lundström E-B, Johansson E, Chabes A, Kunkel TA (2010a) Genome instability due to ribonucleotide incorporation into DNA. Nat Chem Biol 6: 774–781

Nick McElhinny SA, Watts BE, Kumar D, Watt DL, Lundström E-B, Burgers PMJ, Johansson E, Chabes A, Kunkel TA (2010b) Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. Proc Natl Acad Sci U S A 107: 4949–4954

Niehaus M, Straube H, Künzler P, Rugen N, Hegermann J, Giavalisco P, Eubel H, Witte C-P, Herde M (2020) Rapid affinity purification of tagged plant mitochondria (Mito-AP) for metabolome and proteome analyses. Plant Physiol 182: 1194–1210

Nieman RH, Pap DL, Clark RA (1978) Rapid purification of plant nucleotide extracts with xad-2, polyvinylpolypyrrolidone and charcoal. J Chromatogr A 161: 137–146

Nizam S, Qiang X, Wawra S, Nostadt R, Getzke F, Schwanke F, Dreyer I, Langen G, Zuccaro A (2019) Serendipita indica E5′NT modulates extracellular nucleotide levels in the plant apoplast and affects fungal colonization. EMBO Rep 20

Nordlund P, Reichard P (2006) Ribonucleotide reductases. Annu Rev Biochem 75: 681–706

Nygaard P (1972) Deoxyribonucleotide pools in plant tissue cultures. Physiol Plant 26: 29–33

Odmark G, Kihlman BA (1965) Effects of chromosome-breaking purine derivatives on nucleic acid synthesis and on the levels of adenosine 5′-triphosphate and deoxyadenosine 5′-triphosphate in bean root tips. Mutat Res 2: 274–286

Pabst M, Grass J, Fischl R, Léonard R, Jin C, Hinterkörner G, Borth N, Altmann F (2010) Nucleotide and nucleotide sugar analysis by liquid chromatography–electrospray ionization–mass spectrometry on surface-conditioned porous graphitic carbon. Anal Chem 82: 9782–9788

Pedroza-García JA, Nájera-Martínez M, de La Paz Sanchez M, Plasencia J (2015) Arabidopsis thaliana thymidine kinase 1a is ubiquitously expressed during development and contributes to confer tolerance to genotoxic stress. Plant Mol Biol 87: 303–315

Pedroza-García J-A, Nájera-Martínez M, Mazubert C, Aguilera-Alvarado P, Drouin-Wahbi J, Sánchez-Nieto S, Gualberto JM, Raynaud C, Plasencia J (2019) Role of pyrimidine salvage pathway in the maintenance of organellar and nuclear genome integrity. Plant J Cell Mol Biol 97: 430–446

Rampazzo C, Miazzi C, Franzolin E, Pontarin G, Ferraro P, Frangini M, Reichard P, Bianchi V (2010) Regulation by degradation, a cellular defense against deoxyribonucleotide pool imbalances. Mutat Res 703: 2–10

Raveneau M-P, Benamar A, Macherel D (2017) Water content, adenylate kinase, and mitochondria drive adenylate balance in dehydrating and imbibing seeds. J Exp Bot 68: 3501–3512

Regensdorff M, Deckena M, Stein M, Borchers A, Scherer G, Lammers M, Hänsch R, Zachgo S, Buschmann H (2018) Transient genetic transformation of Mougeotia scalaris (Zygnematophyceae) mediated by the endogenous α-tubulin1 promoter. J Phycol 54: 840–849

Riegler H, Geserick C, Zrenner R (2011) Arabidopsis thaliana nucleosidase mutants provide new insights into nucleoside degradation. New Phytol 191: 349–359

Riondet C, Morel S, Alcaraz G (2005) Determination of total ribonucleotide pool in plant materials by high-pH anion-exchange high-performance liquid chromatography following extraction with potassium hydroxide. J Chromatogr A 1077: 120–127

Robak P, Robak T (2013) Older and new purine nucleoside analogs for patients with acute leukemias. Cancer Treat Rev 39: 851–861

Rolletschek H, Melkus G, Grafahrend-Belau E, Fuchs J, Heinzel N, Schreiber F, Jakob PM, Borisjuk L (2011) Combined noninvasive imaging and modeling approaches reveal metabolic compartmentation in the barley endosperm. Plant Cell 23: 3041–3054

Sabina RL, Paul A-L, Ferl RJ, Laber B, Lindell SD (2007) Adenine nucleotide pool perturbation is a metabolic trigger for AMP deaminase inhibitor-based herbicide toxicity. Plant Physiol 143: 1752–1760

Sakamoto W, Takami T (2014) Nucleases in higher plants and their possible involvement in DNA degradation during leaf senescence. J Exp Bot 65: 3835–3843

Salem SA, Yoshida T, Perez de Souza L, Alseekh S, Bajdzienko K, Fernie AR, Giavalisco P (2020) An improved extraction method enables the comprehensive analysis of lipids, proteins, metabolites and phytohormones from a single sample of leaf tissue under water-deficit stress. Plant J 103: 1614–1632

Savitch LV, Barker-Åstrom J, Ivanov AG, Hurry V, Öquist G, Huner NP, Gardeström P (2001) Cold acclimation of Arabidopsis thaliana results in incomplete recovery of photosynthetic capacity, associated with an increased reduction of the chloroplast stroma. Planta 214: 295–303

Sawert A, Backer A, Plank-Schumacher K-H, Wagner KG (1987) Determination of nucleotides and nucleosides in cereal leaves by high performance liquid chromatography. J Plant Physiol 127: 183–186

Schroeder RY, Zhu A, Eubel H, Dahncke K, Witte C-P (2018) The ribokinases of Arabidopsis thaliana and Saccharomyces cerevisiae are required for ribose recycling from nucleotide catabolism, which in plants is not essential to survive prolonged dark stress. New Phytol 217: 233–244

Šmarda P, Bureš P, Horová L, Leitch IJ, Mucina L, Pacini E, Tichý L, Grulich V, Rotreklová O (2014) Ecological and evolutionary significance of genomic GC content diversity in monocots. Proc Natl Acad Sci U S A 111: E4096–E4102

de Souza AP, Cocuron J-C, Garcia AC, Alonso AP, Buckeridge MS (2015) Changes in whole-plant metabolism during the grain-filling stage in sorghum grown under elevated CO₂ and drought. Plant Physiol 169: 1755–1765

Stasolla C, Katahira R, Thorpe TA, Ashihara H (2003) Purine and pyrimidine nucleotide metabolism in higher plants. J Plant Physiol 160: 1271–1295

Stitt M, Lilley RM, Heldt HW (1982) Adenine nucleotide levels in the cytosol, chloroplasts, and mitochondria of wheat leaf protoplasts. Plant Physiol 70: 971–977

Tanaka K, Yoshioka A, Tanaka S, Wataya Y (1984) An improved method for the quantitative determination of deoxyribonucleoside triphosphates in cell extracts. Anal Biochem 139: 35–41

Traub M, Flörchinger M, Piecuch J, Kunz H-H, Weise-Steinmetz A, Deitmer JW, Ekkehard Neuhaus H, Möhlmann T (2007). The fluorouridine insensitive 1 (fur1) mutant is defective in equilibrative nucleoside transporter 3 (ENT3), and thus represents an important pyrimidine nucleoside uptake system in Arabidopsis thaliana. Plant J Cell Mol Biol 49: 855–864

Ullrich J, Calvin M (1962) Alcohol-resistant phosphatase activity in chloroplasts. Biochim Biophys Acta 63: 1–10

Vetsigian K, Goldenfeld N (2008) Genome rhetoric and the emergence of compositional bias. Proc Natl Acad Sci U S A 106: 215–220

Vincenzetti S, Cambi A, Neuhard J, Schnorr K, Grelloni M, Vita A (1999) Cloning, expression, and purification of cytidine deaminase from Arabidopsis thaliana. Protein Expr Purif 15: 8–15

43

Wang C, Liu Z (2006) Arabidopsis ribonucleotide reductases are critical for cell cycle progression, DNA damage repair, and plant development. Plant Cell **18**: 350–365

Wang M, Oppedijk BJ, Caspers MPM, Lamers GEM, Boot MJ, Geerlings DNG, Bakhuizen B, Meijer AH, Duijn BV (1998) Spatial and temporal regulation of DNA fragmentation in the aleurone of germinating barley. J Exp Bot **49**: 1293–1301

Werner AK, Romeis T, Witte C-P (2010) Ureide catabolism in *Arabidopsis thaliana* and Escherichia coli. Nat Chem Biol **6**: 19–21

Werner AK, Witte C-P (2011) The biochemistry of nitrogen mobilization: purine ring catabolism. Trends Plant Sci **16**: 381–387

Witte C-P, Herde M (2020) Nucleotide metabolism in plants. Plant Physiol **182**: 63–78

Xu D, Leister D, Kleine T (2020) VENOSA4, a human dNTPase SAMHD1 homolog, contributes to chloroplast development and abiotic stress tolerance. Plant Physiol **182**: 721–729

Xu J, Zhang L, Yang D-L, Li Q, He Z (2015) Thymidine kinases share a conserved function for nucleotide salvage and play an essential role in *Arabidopsis thaliana* growth and development. New Phytol **208**: 1089–1103

Yabuki N, Ashihara H (1991) Catabolism of adenine nucleotides in suspension-cultured plant cells. Biochim Biophys Acta **1073**: 474–480

Yoo S-C, Cho S-H, Sugimoto H, Li J, Kusumi K, Koh H-J, Iba K, Paek N-C (2009) Rice virescent3 and stripe1 encoding the large and small subunits of ribonucleotide reductase are required for chloroplast biogenesis during early leaf development. Plant Physiol **150**: 388–401

Zhou W, Yang S, Wang PG (2017) Matrix effects and application of matrix effect factor. Bioanalysis **9**: 1839–1844

Zrenner R, Stitt M, Sonnewald U, Boldt R (2006) Pyrimidine and purine biosynthesis and degradation in plants. Annu Rev Plant Biol **57**: 805–836
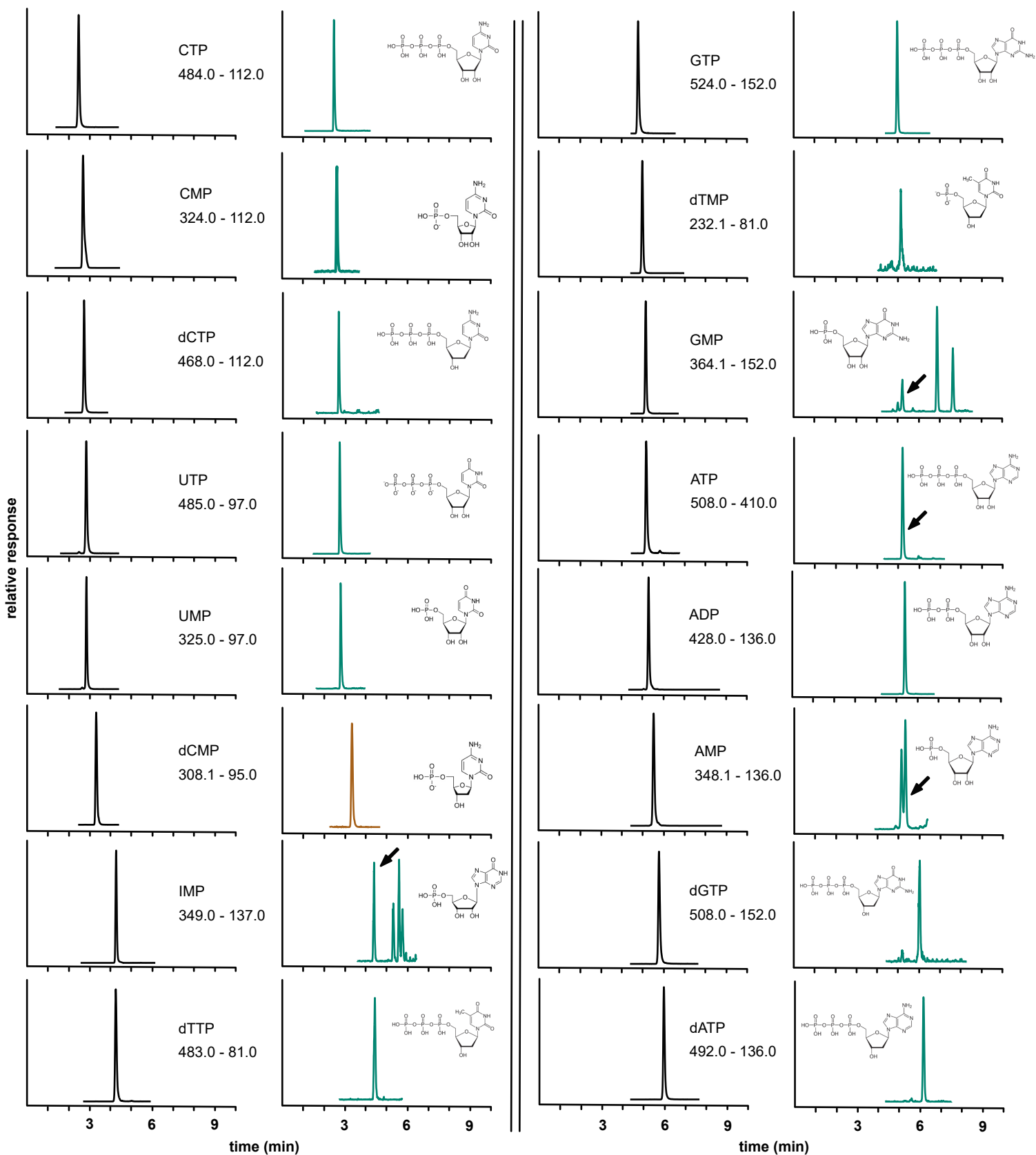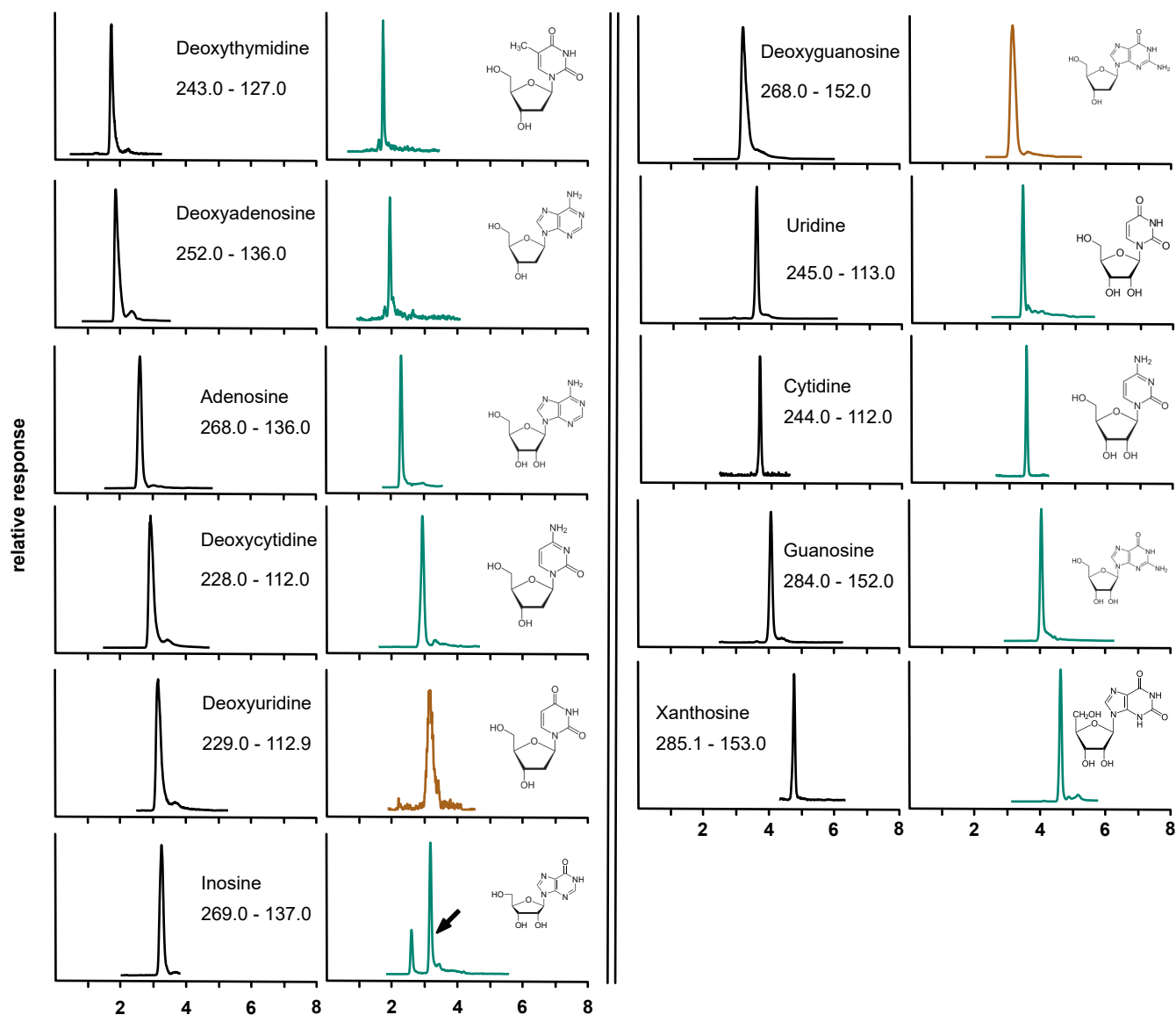
Supplemental Figure 1. Chromatograms of all analyzed nucleotides by hypercarb chromatography.

Representative depiction of Hypercarb multiple reaction monitoring (MRM) chromatograms of all analyzed (deoxy)nucleotides in buffer (left column) or in plant matrix (right column). Chromatograms shown in turquoise represent naturally occuring compounds, while chromatograms shown in brown indicates spiking of the respective analyte in plant matrix. MRM transitions used for quantification are given for the shown analytes. (Supports Figure 2,3,4,7,8 and 9)
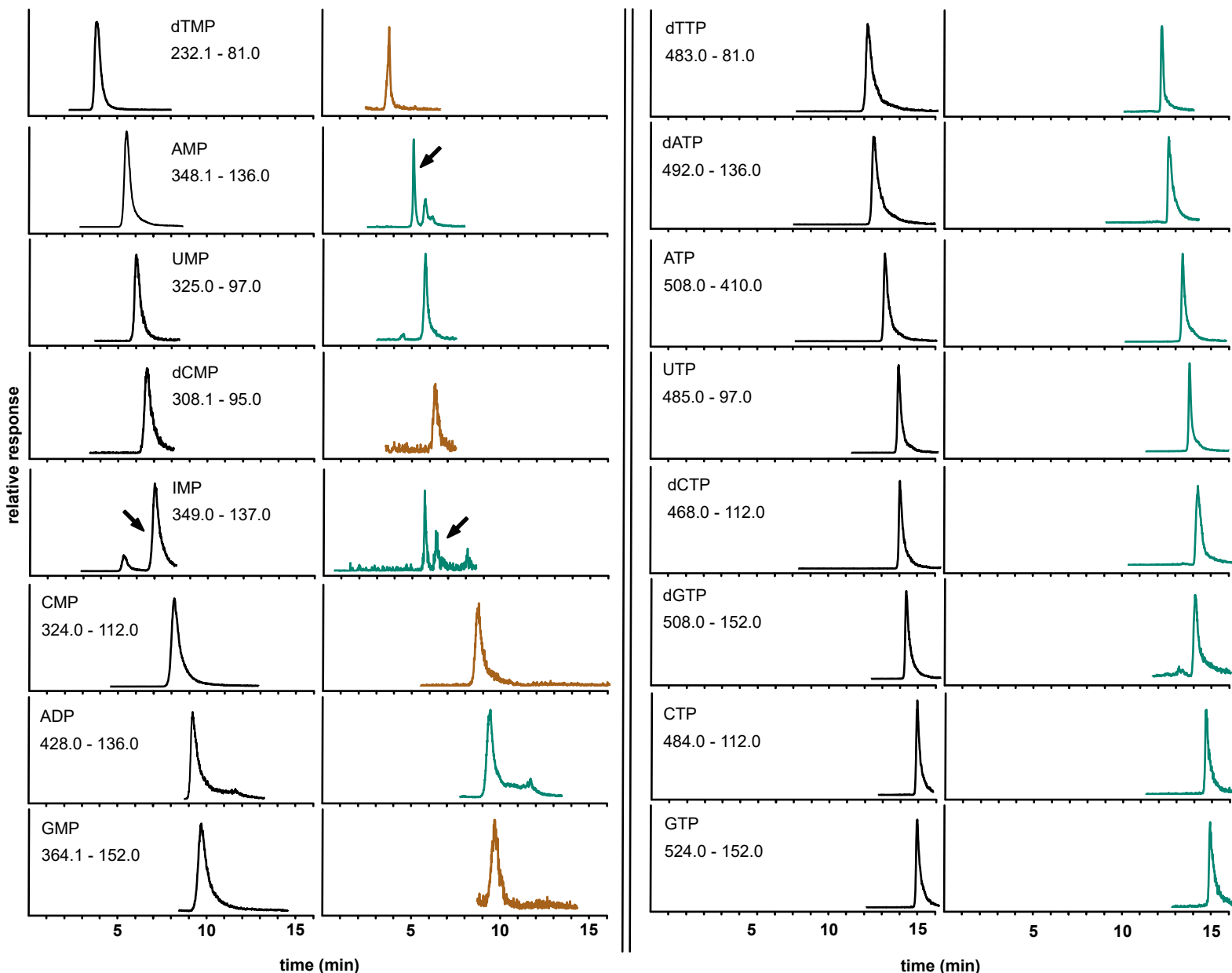
Supplemental Figure 2. Chromatograms of all analyzed nucleotides by cHILIC chromatography.

Representative depiction of cHILIC multiple reaction monitoring (MRM) chromatograms of all analyzed (deoxy)nucleotides in buffer (left column) or in plant matrix (right column). Chromatograms shown in turquoise represent naturally occuring compounds, while chromatograms shown in brown indicates spiking of the respective analyte in plant matrix. MRM transitions used for quantification are given for the shown analytes. (Supports Figure 2,3,4,7,8 and 9)

Supplemental Figure 3. Chromatograms of all analyzed nucleosides by cHILIC chromatography.

Representative depiction of cHILIC multiple reaction monitoring (MRM) chromatograms of all analyzed (deoxy)nucleosides in buffer (left column) or in plant matrix (right column). Chromatograms shown in turquoise represent naturally occuring compounds, while chromatograms shown in brown indicates spiking of the respective analyte in plant matrix. MRM transitions used for quantification are given for the shown analytes. (Supports Figure 6)
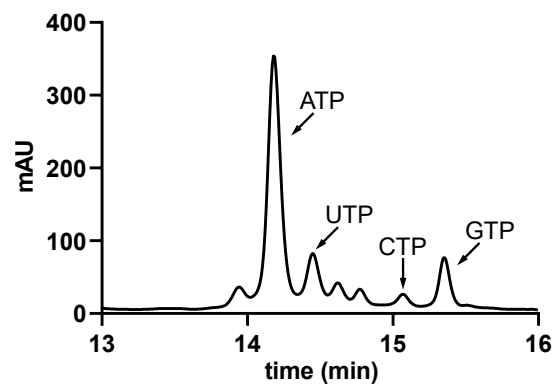
**Supplemental Figure 4. Analysis of nucleotides in Arabidopsis leaves by HPLC and UV detection.**

Analysis of extracts treated with LLE and SPE from 6-week-old Arabidopsis leaves by cHILIC LC using photometric detection at 254 nm. Major photometric peaks were associated with MS signals of NTs using an Orbitrap MS analyzer considering the delay volume between the photometric detector and the MS. mAU, milliabsorption
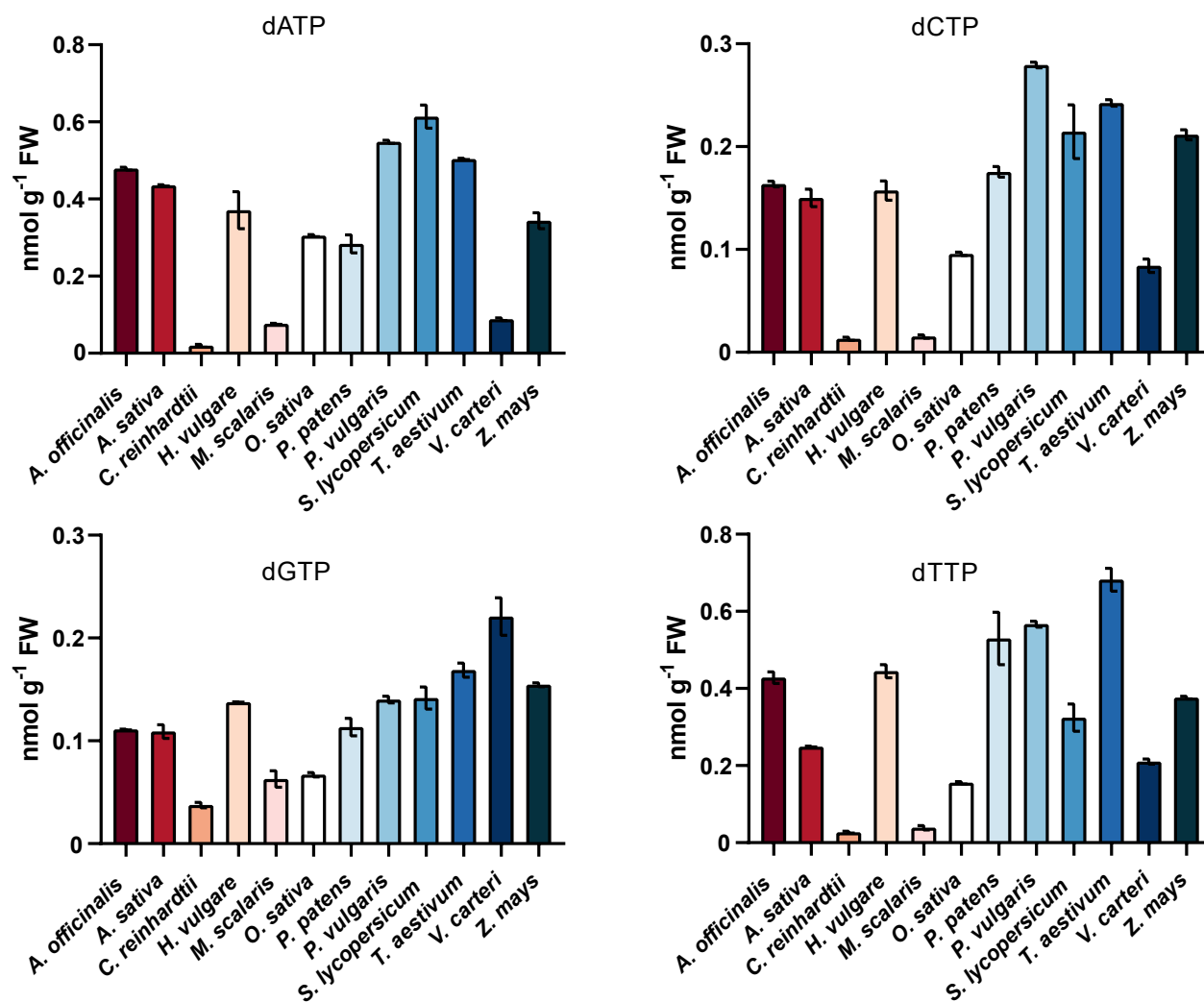
Supplemental Figure 5. SPE enables detection of deoxyribonucleotide triphosphates in several plant and algae species.

Determination of dNTP concentrations in *Asparagus officinalis, Avena sativa, Arabidopsis thaliana, Chlamydomonas reinhardtii, Hordeum vulgare, Mougeotia scalaris, Oryza sativa, Physcomitrella patens, Phaseolus vulgaris, Solanum lycopersicum, Triticum aestivum, Volvox carteri and Zea mays*. Error bars are SD. n = 3 biological replicates, consisting of several plants each. (Supports Figure 3)

A

seeds

B
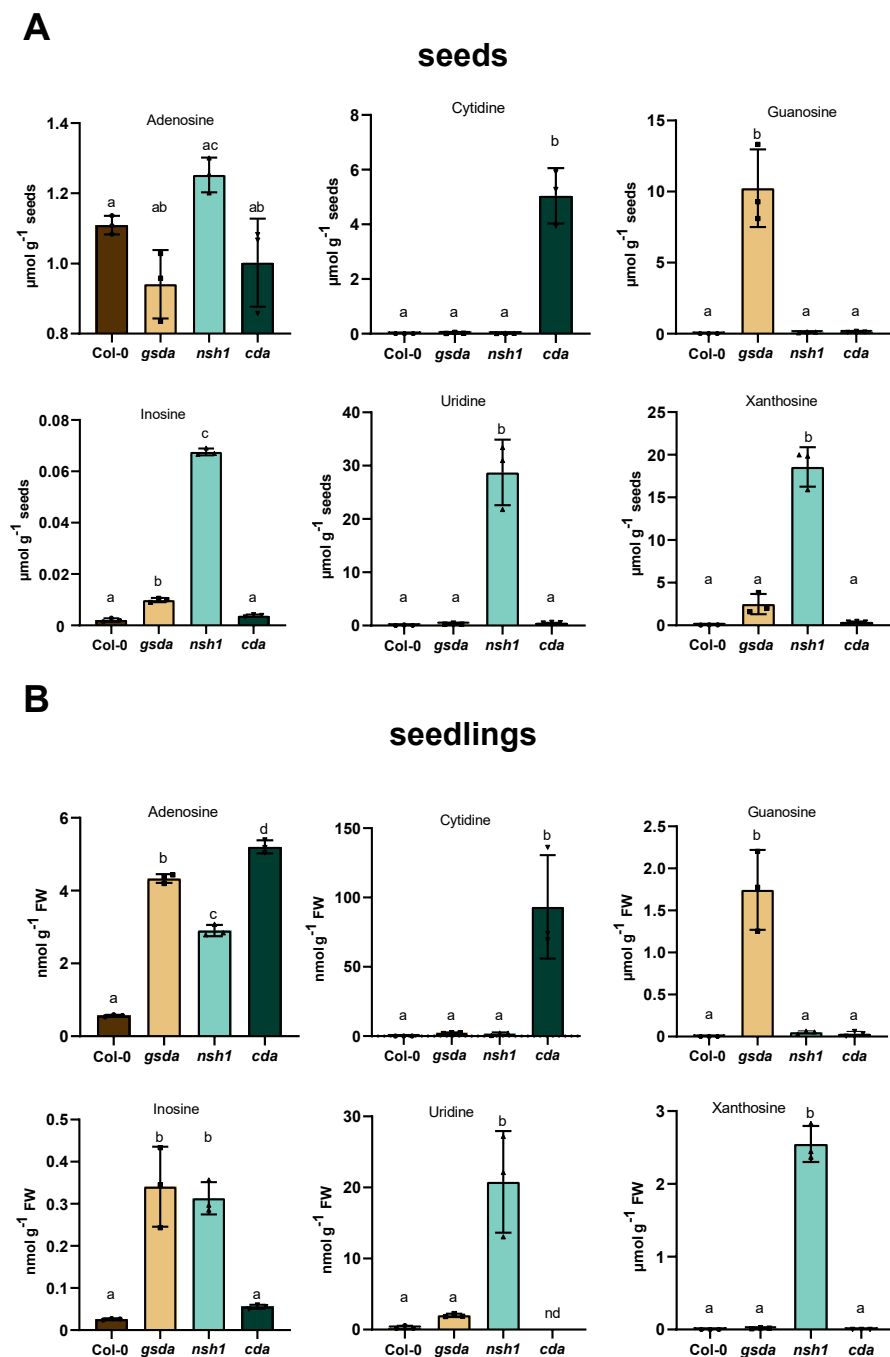
seedlings

Supplemental Figure 6. SPE enables detection of ribonucleosides in Arabidopsis seeds and seedlings.

Concentrations of ribonucleosides in seeds (A) and seedlings (B) of *A. thaliana* wild type, as well as mutants in the salvage and degradation of purine and pyrimidine nucleosides (*gsda*, guanosin deaminase; *cda*, cytidine deaminase; *nsh1* nucleoside hydrolase 1). Error bars are SD, n = 3 biological replicates, for seeds, three independent seed pools derived from different mother plants were used; for seedlings, pools of seedlings from three independent liquid cultures were used. Statistical analysis was performed using one-way ANOVA with Tukey´s post test. Different letters indicate P < 0.05. nd, not detected. FW, fresh weight. (Supports Figure 6)

Supplemental Table 1: *Literature survey on nucleotide analysis in plants*

sample material	measured metabolites	absolute/ relative	method	reference
Vicia faba root tips	dATP ATP	relative (ATP) and absolute (dATP)	acidic extraction/periodate treatment, spectroscopy	Odmark and Kihlman, 1965
Acer pseudoplatanus callus culture, *Pinus mugo* pollen	dNTPs rNTPs	absolute	PCA extraction, TLC	Nygaard, 1972
Gossypium barbadense radicle tips	dTTP	relative	polymerase based assay	Katterman *et al.*, 1975
Triticum aestivum embryos	dNTPs	absolute	TCA extraction, Polymerase based assay	Castroviejo *et al.*, 1978
Scenedesmus ohliquus	dNTPs	absolute	60% MeOH extraction, polymerase based assay	Feller *et al.*, 1980
Datura innoxia and Nicotiana tabacum cell culture, leaves and roots of *Nicotiana tabacum*	ribonucleosides ribonucleotides nucleotide-sugars NAD/NADP	absolute	PCA extraction, Phenylsilane-silica gel, LC-UV	Meyer and Wagner, 1985
Solanum lycopersicum root tips	dNTPs dNDPs	absolute	TCA extraction with PVP[a]/freon-amin/periodate treatment, LC-UV	Dutta *et al.*, 1991
plants and cell cultures of *Nicotiana tabacum*, *Arabidopsis thaliana*, *Solanum lycopersicum*	ribonucleotides cNMPs[c]	absolute	KOH[b] extraction, LC-UV	Riondet *et al.*, 2005
Arabidopsis thaliana	dNTPs	relative	polymerase based assay	Wang and Liu, 2006
Solanum tuberosum tubers	ribonucleotides	absolute	TCA extraction with EGTA[d]/diethylether-amin, LC-UV	Katahira and Ashihara, 2006
leaves of *Nicotiana benthamiana* and *Arabidopsis*	ribonucleotides nucleotide- sugars		sodium fluoride extraction, SPE, LC-MS	Pabst *et al.*, 2010

[a] polyvinylpyrrolidone; [b] potassium hydroxide; [c] cyclic nucleotide monophosphates, [d] ethylenglycol-bis(aminoethylether)-N,N,N,N-tetraaceticacid, [e]ion-exchange chromatography

Supplemental Table 2: *Tested conditions for SPE optimization*

	method 1	method 2	method 3	method 4	method 5	method 6	method 7	method 8	method 9
quenching	5% TCA, 5 mM EDTA	5% TCA, 5 mM EDTA	80/20 MeOH/ 10 mM NH$_4$AC pH 4.5 (v/v)	15% TCA, 30 mM MgCl$_2$	15% TCA, 30 mM MgCl$_2$	15% TCA, 30 mM MgCl$_2$	6 % PCA	15 % TCA	15 % TCA
LLE/ acid quenching	-	-	-	DCM/TOA (78/22 v/v)	DCM/TOA (78/22 v/v)	DCM/TOA (78/22 v/v)	20% KOH	DCM/TOA (78/22 v/v)	DCM/TOA (78/22 v/v)
dilution	1 ml water	1 ml 10 mM NH$_4$AC pH 4.5	-	-	1 ml 10 mM NH$_4$AC pH 4.5	1 ml MeOH	-	-	1 ml water
washing	1 ml MeOH, 1 ml 10 mM NH$_4$AC pH 4.5	1 ml MeOH, 1 ml 10 mM NH$_4$AC pH 4.5	1 ml MeOH, 1 ml 10 mM NH$_4$AC pH 4.5	1 ml MeOH, 1 ml 10 mM NH$_4$AC pH 4.5	1 ml MeOH, 1 ml 10 mM NH$_4$AC pH 4.5	1 ml MeOH, 1 ml 10 mM NH$_4$AC pH 4.5	1 ml MeOH, 1 ml 10 mM NH$_4$AC pH 4.5	1 ml MeOH, 1 ml 1 mM NH$_4$AC pH 4.5	1 ml MeOH, 1 ml 1 mM NH$_4$AC pH 4.5
elution					2 x 500 µl 5% NH$_3$ in MeOH				

Supplemental Table 3: *Relative recovery of metabolites extracted by different methods*

	relative recovery (%)								
	method 1	method 2	method 3	method 4	method 5	method 6	method 7	method 8	method 9
dATP; $_{15N}$	42.2	48.0	10.4	86.9	66.5	87.0	69.0	93.2	100.2
dCTP; $_{15N}$	55.7	54.0	14.1	97.7	49.9	69.6	84.0	106.4	104.0
dGTP; $_{13C, 15N}$	46.4	48.0	3.8	77.5	56.4	66.3	47.8	63.7	64.6
dTTP; $_{13C, 15N}$	57.9	57.8	10.2	96.7	53.4	90.5	92.3	89.6	87.8
ATP; $_{Deut.}$	45.6	49.7	10.3	81.2	69.4	92.0	61.6	92.7	90.0
CTP; $_{Deut.}$	55.7	54.8	19.4	78.57	66.5	96.1	68.0	84.7	102.5
GTP; $_{Deut.}$	42.5	53.4	16.4	59.0	28.4	36.8	23.6	51.8	41.2
UTP; $_{Deut.}$	58.9	49.0	16.1	47.0	47.6	95.2	72.6	78.0	91.0
dTMP; $_{13C, 15N}$	1.2	0.7	107.8	12.8	15.2	15.5	26.8	44.9	88.4

Supplemental Table 4: *Prescursor, transitions (quantifier, qualifier), fragmentor, collision energy and retention times for nucleotide triphosphate measurements*

	precursor ion [M+H]$^+$ (m/z)	product ions* (m/z)	fragmentor	collision energy (V)	retention time cHILIC (min)	retention time Hypercarb (min)
ATP	508	410	30	15	13.8	5.1
		136	30	45		
ATP; Deut.	512	414	30	15	13.8	5.1
		136	30	45		
CTP	483.99	112	25	21	15.0	2.4
		97	25	40		
CTP; Deut.	488.99	113	25	21	15.0	2.4
		100	25	40		
dATP	492.01	81.1	55	40	13.2	6.0
		136	55	29		
dATP; 15N	497	81.1	55	40	13.2	6.0
		141	55	29		
dCTP	468	111.9	85	15	14.6	2.7
		81.1	85	15		
dCTP; 15N	471	114.9	85	15	14.6	2.7
		81.1	85	15		
dGTP	508	152	122	18	14.9	5.8
		81.1	122	18		
dGTP; 13C, 15N	523	162	122	18	14.9	5.8
		81.1	122	18		
dTTP	483	207	80	15	12.7	4.2
		81.1	80	15		
dTTP; 13C,15N	495	219	80	15	12.7	4.2
		86	80	15		
GTP	524	151.9	104	25	15.3	4.7
		97.1	104	25		
GTP; Deut.	528	151.9	104	25	15.3	4.7
		100	104	25		
UTP	485	227	135	25	14.2	2.7
		97	135	25		
UTP; Deut.	490	232	135	25	14.2	2.7
		100	135	25		

* The first product ion was used as the quantifier, while the second product ion (if stated) was used as the qualifier.

Supplemental Table 5: *Prescursor, transitions (quantifier, qualifier), fragmentor energy, collision energy and retention times for nucleotide monophosphate measurements*

	precursor ion [M+H]$^+$ (m/z)	product ion* (m/z)	fragmentor	collision energy (V)	retention time cHILIC (min)	retention time Hypercarb (min)
AMP	348	136	132	25	5.4	5.5
AMP; 15N	353.1	141	132	25	5.4	5.5
CMP	324	112	90	12	8.7	2.6
		97	90	28		
CMP; 13C 15N	336	119	90	28	8.7	2.6
dCMP	308.1	112	60	9	6.9	3.3
		95	60	40		
		81	60	33		
dCMP; 15N	311.1	115	60	9	6.9	3.3
		97	60	40		
dTMP	323.1	126.9	60	25	4.2	5.0
		81	60	17		
dTMP; 13C 15N	335	134	60	25	4.2	5.0
		86.1	60	17		
GMP	364	152	80	13	10.0	5.1
		135	80	45		
GMP; 15N	369	157	80	13	10.0	5.1
		139	80	15		
UMP	325	212.9	98	3	6.0	2.8
		97	98	10		
UMP; 15N	327	212.9	98	3	6.0	2.8
		97	98	10		

* The first product ion was used as the quantifier, while the second product ion (if stated) was used as the qualifier.

Supplemental Table 6: *Prescursor, transitions (quantifier, qualifier), fragmentor energy, collision energy and retention times for nucleoside measurements*

	precursor ion [M+H]+ (m/z)	product ion* (m/z)	fragmentor	collision energy (V)	retention time cHILIC (min)
Adenosine	268	136	86	15	2.6
		119	86	15	
Adenosine; 13C	273	136	86	15	2.6
		119	86	15	
Cytidine	244	133	20	150	3.7
		112	20	150	
Cytidine; 15N	247	133	20	150	3.7
		115	20	150	
Deoxyadenosine	252	136	59	19	2.0
		119	59	51	
Deoxycytidine	228	112	55	9	3.0
		95	55	40	
Deoxyguanosine	268	152	59	19	3.4
		135	59	39	
		110	59	43	
Deoxythymidine	243	127	60	9	1.9
		81	60	17	
Deoxythymidine; 13C 15N			60		1.9
	255	134		9	
	255	86	60	17	
Deoxyuridine	229	117	50	2	3.2
		113	50	10	
Guanosine	284	152	90	10	4.0
		135	90	45	
Guanosine; 15N	289	157	90	10	4.0
		139	90	45	
Inosine	269	137	55	14	3.3
		119	55	40	
Inosine, 15N	273	141	55	14	3.3
Uridine	245	133	85	14	3.6
		113	85	14	
Uridine; 15N	247	115	85	14	3.6

* The first product ion was used as the quantifier, while the second product ion (if stated) was used as the qualifier.

Supplemental Table 7. *Intra-day variation and inter-day variation for different concentrations of nucleotides seperated by the Hypercarb or cHILIC method*

amount (pmol on column)	Hypercarb method[a]		cHILIC method[a]	
	intra-day precision (CV%)	inter-day precision (CV%)	intra-day precision (CV%)	inter-day precision (CV%)
dATP				
5	6.4	10.4	14.1	14.9
10	8.7	14.7	10.3	12.5
20	10.8	17.4	9.5	21.5
dCTP				
5	3.0	2.5	19.7	26.3
10	1.6	1.3	11.0	22.1
20	12.7	3.7	10.7	18.2
dGTP				
5	4.5	5.1	22.3	26.2
10	1.1	5.8	16.5	12.4
20	1.3	6.2	12.3	30.8
dTTP				
5	6.7	5.1	5.8	2.3
10	2.5	1.4	14.0	3.5
20	1.7	0.6	10.0	10.4
ATP				
100	1.1	5.6	11.0	7.5
400	3.5	10.2	8.1	3.8
800	2.5	10.0	5.8	11.8
CTP				
50	13.4	10.9	26.6	11.5
100	10.2	8.6	27.5	14.5
200	8.7	6.5	14.0	23.3
GTP				
50	0.4	7.7	24.1	17.0
100	2.2	9.3	19.9	15.0
400	2.3	8.3	8.9	2.7
UTP				
50	7.0	9.0	17.8	13.3
100	1.9	9.6	13.3	6.3
400	2.3	10.8	6.7	5.8

[a] n = 3 replicates

Supplemental Table 8: *Intra-day variation and inter-day variation for different concentrations of nucleosides seperated by the cHILIC method*

	cHILIC method[a]	
amount (nmol on column)	intra-day precision (CV%)	inter-day precision (CV%)
Adenosine		
0.125	0.8	11.7
0.250	3.0	8.0
1.000	0.7	15.6
Cytdine		
0.125	6.6	9.8
0.250	6.6	4.6
1.000	0.9	16.1
Guanosine		
0.125	6.0	12.8
0.250	7.6	6.0
1.000	0.6	19.0
Inosine		
0.125	10.8	16.8
0.250	6.1	6.5
1.000	0.8	21.4
Uridine		
0.125	5.6	9.8
0.250	5.6	4.6
1.000	0.9	16.9
Deoxythymidine		
0.125	1.6	16.8
0.250	2.1	8.1
0.500	1.7	25.5

[a] n = 3 replicates

Supplemental *Table 9.* *Determination of matrix effect factor (MEF) at 5 and 50 pmol dNTP and rNTP ISTDs, respectively, on column*

metabolite	Hypercarb		cHILIC	
	MEF before SPE (%)	MEF after SPE (%)	MEF before SPE (%)	MEF after SPE (%)
dATP	97.0	43.9	96.0	3.7
dCTP	nd[a]	56.3	nd	13.7
dGTP	97.8	63.9	nd	14.0
dTTP	nd	65.3	93.3	10.3
ATP	99.8	65.1	91.1	12.2
CTP	nd	74.8	nd	16.8
GTP	98.0	72.7	nd	30.3
UTP	nd	73.1	nd	17.2

[a] *nd,* metabolite was not detected

Supplemental Table 10. *Determination of matrix effect factor (MEF) at 5 pmol nucleoside ISTD on column*

metabolite	cHILIC	
	MEF before SPE (%)	MEF after SPE (%)
Deoxythymidine	89.2	13.8
Adenosine	95.9	81.2
Cytidine	52.1	-111.2
Guanosine	58.8	36.5
Inosine	63.3	-122.0
Uridine	50.8	-111.2

Table 11: *NTPs analyzed with an Orbitrap mass spectrometer*

nucleotides	elemental composition	measured exact mass	Δppm	masses of product ions	identified by full MS mode	matched isotopes/expected isotopes	RT
ATP	C10H16N5O13P3	508.0027	-0.72	136.0617 348.0702	yes	4/4	7.12
GTP	C10H16N5O14P3	523.9974	-0.95	97.0288 152.0566	yes	4/4	6.68
CTP	C9H16N3O14P3	483.9914	-0.81	97.0288 112.0508	no	2/4	4.31
UTP	C9H15N2O15P3	484.9742	-3.41	97.0288 227.0661	yes	2/3	4.4
dATP	C10H16N5O12P3	492.0078	-0.66	136.0617 103.0393	yes	3/3	9.1
dGTP	C10H16N5O13P3	*nd[a]*	*nd*	*nd*	no	*nd*	*nd*
dCTP	C9H16N3O13P3	*nd*	*nd*	*nd*	no	*nd*	*nd*
dTTP	C10H17N2O14P3	482.9965	-0.1	81.0340 112.0508	yes	3/3	6.75

[a] nd, not detected

2.2 Purification and analysis of Nucleotides and Nucleosides from Plants

Henryk Straube[1] and Marco Herde[1]

[1] Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany

Chapter 11

Purification and Analysis of Nucleotides and Nucleosides from Plants

Henryk Straube and Marco Herde

Abstract

This protocol describes necessary steps to isolate and quantify nucleotides and nucleosides from plant samples. Proper sample preparation in combination with liquid chromatography coupled to mass spectrometry enables the sensitive detection and quantification of metabolites of low abundance. Utilizing a liquid–liquid extraction in combination with a weak anion-exchange solid phase extraction enables the separation of negatively charged molecules from uncharged metabolites or cations.

Key words Plant nucleotide metabolism, Weak anion-exchange solid phase extraction, Metabolomics, Liquid chromatography, Mass spectrometry, Polar metabolomics

1 Introduction

Metabolomic data reporting the *in planta* concentration of nucleotides and nucleosides are scarce, since their detection and quantification is impeded in part by the low abundance of, for example, deoxyribonucleotides and the complex plant matrix which complicates chromatographic separation and detection with mass spectrometry [1]. Hitherto, deoxyribonucleotides, the building blocks of DNA, were assayed by a semiquantitative polymerase chain reaction assay [2–4], thin-layer chromatography [5] or a ribonucleotide-elimination protocol [6] rather than a direct liquid chromatography–mass spectrometry (LC-MS) approach. While there are also dedicated methods for the analysis of nucleosides [7–9] these methods are not suitable for the analysis of any phosphorylated compounds due to the lack of appropriate strategies to eliminate enzymatic activities (quenching), in the plant extracts that would consume nucleotides. Thus, we developed a method that is capable of quenching, extracting, and quantifying not only deoxyribonucleotides and ribonucleotides but also nucleosides in one workflow [10]. The key steps in this protocol are (I) a quenching and extraction procedure (Fig. 1a), utilizing a strong acid, halting
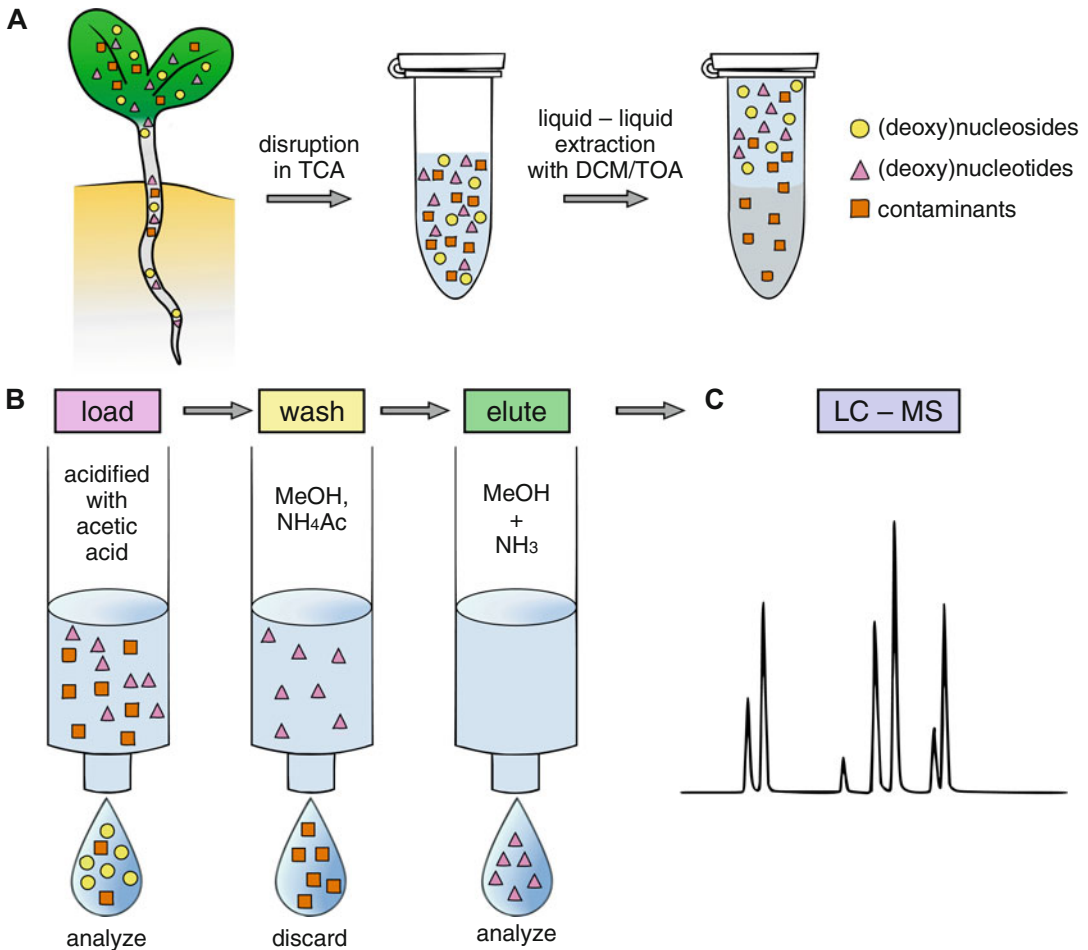
Fig. 1 Schematic overview of the method for the extraction and analysis of nucleotides and nucleosides. (**a**) Plant material is disrupted and quenched with trichloroacetic acid (TCA), which is removed together with apolar contaminants by liquid–liquid extraction with dichloromethane (DCM) and trioctylamine (TOA). (**b**) The extract is loaded onto a weak-anion exchange SPE cartridge, the flow-through contains nucleosides for the analysis. Subsequently, contaminants are depleted by washing with methanol (MeOH) and ammonium acetate (NH_4AC), resulting in elution of nucleotides with MeOH and ammonia (NH_3). (**c**) Isolated fractions are analyzed by LC-MS using a porous graphitized black column or a polar end-capped C18 column (reproduced from [10] (www.plantcell.org), Copyright American Society of Plant Biologists)

all phosphatase activity in the extract, resulting in high recovery rates for phosphorylated compounds [10, 11]; (II) a liquid–liquid extraction step with an organic solvent and a tertiary amine that removes lipids and neutralizes the extractant and (III) a solid phase extraction using a weak anion-exchange resin in which the polar nucleotides are separated from the less polar nucleosides (Fig. 1b). Technical variance during the extraction is corrected with the iso- tope dilution technique that also allows for absolute quantification in the plant. The extracts are then analyzed by liquid

chromatography coupled to mass spectrometry (Fig. 1c). This method provides unprecedented sensitivity, is scalable and the use of dichloromethane constitutes an environmentally friendly and cost-effective alternative to previously used solvents, like 1,1,2-trichloro-1,2,2-trifluorethane.

2 Materials

Prepare all buffers using MS-grade water and MS-grade reagents (*see* **Note 1**). Store all reagents at 4 °C, prepare buffers freshly (buffers may be used up to 2 weeks). Steps including solvents like methanol or dichloromethane, as well as strong acids should be performed under a fume hood. Follow waste disposal regulations for toxic chemicals.

- 15% trichloroacetic acid solution: Weigh 15 g trichloroacetic acid (*see* **Note 2**) and dissolve it in 70 mL water. Transfer the solution to a graduated 100 mL glass cylinder and add water to a final volume of 100 mL. Store at 4 °C.

- 78%/22% dichloromethane–trioctylamine solution: Add dichloromethane up to a volume of 78 mL into a graduated glass cylinder. Then add trioctylamine up to a final volume of 100 mL. Store at 4 °C (*see* **Note 3**).

- 10 mM Ammonium acetate pH 4.5: Weigh 77.09 mg ammonium acetate and solve it in 70 mL water in a 200 mL glass beaker by mixing. Adjust the pH with pure acetic acid. Make up to a final volume of 100 mL with water in a 100 mL graduated glass cylinder. Store at 4 °C (*see* **Note 4**).

- 1 mM Ammonium acetate pH 4.5: Add 10 mL 10 mM ammonium acetate pH 4.5 to a 100 mL graduated glass cylinder and make up to a final volume of 100 mL with water. Store at 4 °C (*see* **Note 4**).

- 0.5% acetic acid: Add 5 mL water to a graduated 10 mL glass cylinder, add 50 µL pure acetic acid and fill up to final volume of 10 mL with water. Store at 4 °C.

- Conditioning solution: Add 50 mL water to a graduated 100 mL glass cylinder. Add 2 mL pure formic acid by pipetting, make up to volume of 75 mL with water. Then add pure methanol to a final volume of 100 mL. Store at 4 °C.

- Elution solution: Add 12 mL pure methanol to a graduated 20 mL glass cylinder. Add 3 mL 25% ammonia solution by pipetting. Store at room temperature until use (*see* **Note 5**).

- Buffer A, 5 mM Ammonium acetate pH 9.5: Add 900 mL water to a graduated 1 L glass cylinder. Add 300 µL pure acetic acid and 1.4 mL 25% ammonia solution to the cylinder. Check the

pH, it should be 9.5. Otherwise, adjust pH with pure acetic acid or 25% ammonia solution. Make up to a final volume of 1 L with water. Store at 4 °C (*see* **Note 6**).

- Buffer B: 1 L 100% acetonitrile.

- Buffer C, 0.1% (v/v) formic acid in water: Add 950 mL water to a graduated 1 L glass cylinder. Add 1 mL formic acid and make up to a final volume of 1 L with water. Store at 4 °C.

- Buffer D, 0.1% (v/v) formic acid in methanol: Add 950 mL methanol to a graduated 1 L glass cylinder. Add 1 mL formic acid and make up to a final volume of 1 L with methanol. Store at room temperature in a bottle.

- Buffer E, 0.1% (v/v) formic acid in water/methanol: Add 96 mL of buffer C to a graduated 100 mL glass cylinder. Make up to a final volume of 100 mL with buffer D.

- Prepare 2 mM stocks for 2′-deoxyadenosine 5′-triphosphate (dATP, 49.12 mg/ 20 mL), 2′-deoxycytidine 5′-triphosphate (dCTP, 46.72 mg/ 20 mL), 2′-deoxyguanosine 5′-triphosphate (dGTP, 50.72 mg/ 20 mL), 2′-deoxythymidine 5′-triphosphate (dTTP, 48.22 mg/ 20 mL), adenosine 5′-triphosphate (ATP, 50.72 mg/ 20 mL), cytidine 5′-triphosphate (CTP, 48.32 mg/ 20 mL), guanosine 5′-triphosphate (GTP, 52.32 mg/20 mL), uridine 5′-triphosphate (UTP, 48.14 mg/20 mL), adenosine 5′-monophosphate (AMP, 34.72 mg/20 mL) and guanosine 5′-monophosphate (GMP, 36.32 mg/20 mL) using buffer A as solvent. Store in 15 mL centrifugation tubes at −80 °C.

- For creating the nucleotide test sample add 1 μL of each 2 mM stock to 190 μL buffer A in a reaction tube and vortex briefly. Transfer into a vial and store at 4 °C until use.

- Create 2 mM stocks for adenosine (26.72 mg/20 mL), cytidine (24.32 mg/20 mL), guanosine (28.32 mg/ 20 mL) and uridine (24.42 mg/20 mL).

- For creating the nucleoside test sample add 1 μL of each 2 mM stock to 196 μL 96% buffer C and 4% buffer D in a reaction tube and vortex briefly. Transfer into a vial and store at 4 °C until use.

- Create a nucleoside standard curve with the concentrations 200 μM, 100 μM, 50 μM, 25 μM, and 12.5 μM. Add 20 μL of each of the 2 mM stock solutions to 120 μL buffer E to create the 200 μM standard. Now serially dilute 100 μL of the 200 μM standard with 100 μL buffer E to create the 100 μM standard. Repeat this process to create the rest of the standards.

3   Methods

The described method is for 12 samples. All centrifugation steps have to be performed at 4 °C in precooled centrifuges. All parts of the mixer mill in direct contact with the sample tubes have to be precooled in liquid nitrogen. Before starting the extraction add 1.3 μL of 4 mM isotope-standards (^{15}N dATP, ^{15}N dCTP, ^{13}C^{15}N dGTP, ^{13}C^{15}N dTTP, $^{deut.}$ rNTPs, ^{15}N AMP, ^{15}N GMP) per ml to 12.983 mL extraction buffer. Store the extraction buffer on ice until use (*see* **Note 7**).

3.1   Tissue Disruption and Metabolite Extraction

1. Harvest approximately 100 mg of any plant material into a 2 mL safe seal reaction tube containing five 5 mm steal beads (*see* **Note 8**). Depending on the material consider a brief washing step with tap-water. Note the weight and quickly freeze the material in liquid nitrogen. Samples can be stored for at least 6 months at −80 °C (*see* **Note 9**).

2. Disrupt the tissue using an oscillating mixer mill at 28 Hz for 2:30 min.

3. Add 1 mL of the isotope-standard containing extraction buffer to every sample and vortex briefly.

4. Disrupt the tissue again in a oscillating mixer mill at 28 Hz for 2:30 min.

5. Centrifuge the samples for 10 min at $40,000 \times g$ (*see* **Note 10**).

3.2   Neutralization and Liquid–Liquid Extraction

1. Transfer the supernatant into a new 2 mL safe-lock reaction tube containing 1 mL of 78%/22% dichloromethane–trioctylamine solution.

2. Vortex thoroughly for 12 s.

3. Centrifuge for 3 min at $5000 \times g$.

4. Transfer the upper phase into a new 2 ml reaction tube containing 995 μL water and 5 μL 0.5% acetic acid. Vortex briefly.

3.3   Solid Phase Extraction Using a Weak-Anion Exchange Resin

All steps should be performed with a flow rate of less than 1 mL min^{-1} (*see* **Note 11**). Make sure that the resin is never completely dried out during the procedure as this will reduce recovery.

1. Add 1 mL methanol to a 1 mL cartridge containing 30 mg weak anion-exchange resin and percolate it through the cartridge using a vacuum-manifold and a vacuum-pump or a positive pressure manifold.

2. Repeat this process with 1 mL conditioning solution (*see* **Note 12**) and 1 mL 10 mM ammonium acetate pH 4.5 (*see* **Note 13**).

3. Load 1 mL of sample and let it enter the resin without suction for 2 min (*see* **Note 14**). Then percolate the sample through the cartridge. Collect the flow-through in a 2 mL reaction tube, this is the nucleoside fraction.

4. Repeat this step with the rest of the sample without the waiting step, collect the flow-through in the same reaction tube.

5. Wash the cartridge with 1 mL 1 mM ammonium acetate pH 4.5. Discard the flow-through.

6. Wash the cartridge with 1 mL pure methanol. Discard the flow-through.

7. Dry the outlets of the manifold with a paper towel to reduce contamination with the washing fraction.

8. Add 0.5 mL elution solution and let it enter the resin without suction for 2 min (*see* **Note 14**). Then percolate the sample through the cartridge. Collect the flow-through in a 2 mL reaction tube, this is the nucleotides containing fraction.

9. Repeat this step and collect the flow-through in the same reaction tube.

10. The nucleosides containing fraction is then evaporated in a freeze-dryer, the nucleotide containing fraction in a vacuum concentrator until no liquid is left. Samples prepared this way can be stored for at least 6 months at $-80\ ^{\circ}$C (*see* **Note 15**).

3.4 Analysis of Nucleotides by Liquid Chromatography Coupled to Mass Spectrometry

1. Set up a 50×4.6 mm graphitized carbon black column with a particle size of 5 μm for chromatography. Flush this column with 100% buffer A, prior to use, for at least 60 min with a flowrate of 0.6 mL/min (*see* **Note 16**).

2. Use a mass spectrometer with an electrospray-ionization source, source parameters have to be optimized on every individual device. For example, when using an Agilent Jet Stream Source, following parameters can be considered a good starting point: positive mode, gas temperature 250 °C, gas flow 12 L min^{-1}, nebulizer pressure 20 psi, sheath gas temperature 395 °C, sheath gas flow 12 L min^{-1}, capillary voltage 3.000 V, and nozzle voltage 500 V. Parameters specific for each nucleotide are listed in Table 1.

3. Run no-injection runs twice using the following gradient: 0.6 mL min^{-1}; 0.0–3.0 min, 0% buffer B; 3.0–18.0 min, 30% buffer B; 18.0–19.0 min, 100% buffer B; 19.0–22.0 min, 100% buffer B; 22.0–22.5, 0% buffer B until the end of the run (30 min) (*see* **Note 17**). The first 5 min and the last 14 min of the method the valve is switched to waste to reduce contamination of the mass spectrometer source.

4. Start injecting (injection volume of 10 μL, *see* **Note 18**) test samples containing 10 μM dATP, dGTP, dTTP, ATP, CTP,

Table 1
Precursor, transitions (quantifier, qualifier), fragmentor and collision energy for nucleotide measurements

	Precursor ion [M + H]$^+$ (m/z)	Product ions[a] (m/z)	Fragmentor	Collision energy (V)
ATP	508	410	30	15
		136	30	45
ATP; Deut.	512	414	30	15
		136	30	45
CTP	483.99	112	25	21
		97	25	40
CTP; Deut.	488.99	113	25	21
		100	25	40
dATP	492.01	81.1	55	40
		136	55	29
dATP; 15 N	497	81.1	55	40
		141	55	29
dCTP	468	111.9	85	15
		81.1	85	15
dCTP; 15 N	471	114.9	85	15
		81.1	85	15
dGTP	508	152	122	18
		81.1	122	18
dGTP; 13C, 15 N	523	162	122	18
		81.1	122	18
dTTP	483	207	80	15
		81.1	80	15
dTTP; 13C,15 N	495	219		15
		86	80	15
GTP	524	151.9	104	25
		97.1	104	25
GTP; Deut.	528	151.9	104	25
		100	104	25
UTP	485	227	135	25
		97	135	25
UTP; Deut.	490	232	135	25
		100	135	25
AMP	348	136	132	25

(continued)

Table 1
(continued)

	Precursor ion [M + H]⁺ (m/z)	Product ions[a] (m/z)	Fragmentor	Collision energy (V)
AMP; 15 N	353.1	141	132	25
GMP	364	152 135	80 80	13 45
GMP; 15 N	369	157 139	80 80	13 15

[a]The first product ion was used as the quantifier, while the second product ion (if stated) was used as the qualifier

GTP, UTP, AMP, and GMP three consecutive times. Note the response as point of reference for instrument performance.

5. Before running plant samples one should always run a sample consisting of the pure buffer A to check for eventual contaminations of the column or sample carryover.

6. If peak-shapes, retention times (±0.5 min) and blank runs are appropriate, resuspend your samples in 50 μL buffer A and start measuring your samples.

7. When finished, wash the column with 80% acetonitrile and 20% water for 60 min. Store the column with these buffers.

8. The concentration of nucleotides in your sample is determined as follows.

$$\left(\frac{\frac{0.08}{\text{area isotope standard}} \times \text{area natural compound}}{[\text{mg}] \text{ fresh weight}} \times 5 \right)$$

$$\times \left(\frac{1}{\text{molecular weight}} \right) \times 1000$$

$$= \text{natural compound} \left[\text{nmol mg}^{-1} \right]$$

In addition, samples with different amounts of standard in the plant matrix (as described for nucleosides) are used to demonstrate the absence of saturation.

3.5 Analysis of Nucleosides by Liquid Chromatography Coupled to Mass Spectrometry

1. Set up your chromatography using a 50×4.6 mm polar end-capped C18 column with a particle size of 5 μm.

2. Use a mass spectrometer with an electrospray-ionization source; a good starting point for an Agilent Jet Stream Source is the following parameters: positive mode, gas temperature 80 °C, gas flow 13 L min⁻¹, nebulizer pressure 30 psi, sheath gas temperature 275 °C, sheath gas flow 11 L min⁻¹, capillary voltage 2.500 V, and nozzle voltage 500 V. Parameters specific for each nucleoside are given in Table 2.

Table 2
Precursor, transitions (quantifier, qualifier) and fragmentor energy for nucleoside measurements

	Precursor ion [M + H]+ (m/z)	Product iona (m/z)	Fragmentor	Collision energy (V)
Adenosine	268	136	86	15
		119	86	15
Cytidine	244	133	20	150
		112	20	150
Guanosine	284	152	90	10
		135	90	45
Uridine	245	133	85	14
		113	85	14

aThe first product ion was used as the quantifier, while the second product ion (if stated) was used as the qualifier

3. Run no-injection runs twice using the following gradient: 0.6 mL min^{-1}; start with 96% buffer C and 4% buffer D; 0.0–8.0 min, 65% buffer D; 8.0–8.2 min, 100% buffer D; 8.2–10.0 min, 100% buffer D; 10.0–10.1 min, 4% buffer D; until the end of the run (12.5 min). The first 2 min of the method the valve is switched to the waste which avoids salt contamination of the mass spectrometer source.

4. Start injecting (injection volume of 10 μL, *see* **Note 18**) standard samples containing 10 μM adenosine, cytidine, guanosine, and uridine three consecutive times.

5. If peak-shapes and retention times (±0.5 min) are appropriate, resuspend your samples in 50 μL buffer E and start measuring your samples.

6. Subsequently measure the five concentrations of your metabolites of interest with three technical replicates each (*see* **Note 19**).

7. Create a standard curve using *Microsoft Excel* or *GraphPad Prism* and create a linear regression. Be aware that in some instances a nonlinear graph is a superior fit for the data. Take the function of the regression ($y = mx \pm c$) and rearrange the equation to solve it for x ($x = (y \pm c)/m$). Fill in your values for your natural component to get the concentration in μM on column. Make sure the values of the natural compound is in the scope of the regression analysis. Now apply the following calculation.

$$\frac{\mu M \text{ on colum} \times 5}{mg \text{ fresh weight}} = \frac{\mu M}{mg \text{ fresh weight}}$$

4   Notes

1. Alternatively use deionized water (resistivity at 25 °C of 18 MΩ·cm); however, nontargeted mass spectrometry approaches require pure solvents.

2. Trichloroacetic acid is a very strong acid, consider weighing under a fume hood.

3. The 78%/22% dichloromethane–trioctylamine solution has a reduced shelf life and can be used up to 3 months.

4. The ammonium acetate buffers are especially prone to contamination by algae or bacteria, check always for contaminations if stored buffer is used.

5. The elution solution has to be prepared fresh each time due to the highly volatile character of the solution.

6. Filtration of buffers for liquid chromatography is recommended for the protection of in-line filters, guard columns or analytical columns. Here, no solids are used for the preparation of buffers thus contamination with particles is unlikely.

7. Other isotopes may be used, one has to adjust the respective source-parameters like precursor and product ions accordingly. Dilution of isotopes also may vary depending on the concentration delivered by the vendor.

8. The amount of beads and the frequency of the oscillating mixer mill may have to be adapted depending on the plant material.

9. This protocol also works using freeze-dried material (Straube, unpublished).

10. Not all reaction tubes are able to withstand this force. If unsure, test the tubes before using them with real samples. The stability of the reaction tubes is also dependent on the size and amount of beads used during tissue disruption.

11. The flowrate is of importance, as a high flow rate leads to tunneling and results in less interaction of metabolites with the resin.

12. This step conditions the resin and increases retention of negatively charged compounds.

13. The ion-strength of the buffer is of importance, as a high ionic strength leads to a low recovery of nucleotide monophosphates.

14. This step allows the resin to be wetted appropriately and reduces tunneling effects leading to a breakthrough of metabolites.

15. We analyzed even older samples and never saw any change in metabolite content.

16. This step helps to increase retention stability and peak shape.

17. The relatively long equilibration time leads to a massive increase in retention stability and overall column performance.

18. The injection volume can be increased, depending on concentration and sample. In our hands an injection of 20 μL increases reliability of the detection and quantification of low-abundant metabolites.

19. If isotope standards for respective nucleosides are available one can also use those analogous to the method described for the nucleotide quantification.

References

1. Straube H, Witte C-P, Herde M (2021) Analysis of nucleosides and nucleotides in plants: an update on sample preparation and LC-MS techniques. Cell 10:689. https://doi.org/10.3390/cells10030689

2. Yoo S-C, Cho S-H, Sugimoto H et al (2009) Rice virescent3 and stripe1 encoding the large and small subunits of ribonucleotide reductase are required for chloroplast biogenesis during early leaf development. Plant Physiol 150:388–401. https://doi.org/10.1104/pp.109.136648

3. Garton S, Knight H, Warren GJ et al (2007) Crinkled leaves 8 - a mutation in the large subunit of ribonucleotide reductase leads to defects in leaf development and chloroplast division in Arabidopsis thaliana. Plant J 50:118–127. https://doi.org/10.1111/j.1365-313X.2007.03035.x

4. Wang C, Liu Z (2006) Arabidopsis ribonucleotide reductases are critical for cell cycle progression, DNA damage repair, and plant development. Plant Cell 18:350–365. https://doi.org/10.1105/tpc.105.037044

5. Nygaard P (1972) Deoxyribonucleotide pools in plant tissue cultures. Physiol Plant 26:29–33. https://doi.org/10.1111/j.1399-3054.1972.tb03541.x

6. Dutta I, Dutta PK, Smith DW et al (1991) High-performance liquid chromatography of deoxyribonucleoside di- and triphosphates in tomato roots. J Chromatogr A 536:237–243. https://doi.org/10.1016/S0021-9673(01)89255-4

7. Chen M, Herde M, Witte C-P (2016) Of the nine cytidine deaminase-like genes in Arabidopsis, eight are pseudogenes and only one is required to maintain pyrimidine homeostasis in vivo. Plant Physiol 171:799–809. https://doi.org/10.1104/pp.15.02031

8. Dahncke K, Witte C-P (2013) Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in Arabidopsis. Plant Cell 25:4101–4109. https://doi.org/10.1105/tpc.113.117184

9. Kopecná M, Blaschke H, Kopecny D et al (2013) Structure and function of nucleoside hydrolases from Physcomitrella patens and maize catalyzing the hydrolysis of purine, pyrimidine, and cytokinin ribosides. Plant Physiol 163:1568–1583. https://doi.org/10.1104/pp.113.228775

10. Straube H, Niehaus M, Zwittian S et al (2021) Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. Plant Cell 33:270–289. https://doi.org/10.1093/plcell/koaa028

11. Bieleski R-L (1964) The problem of halting enzyme action when extracting plant tissues. Anal Biochem 9:431–442. https://doi.org/10.1016/0003-2697(64)90204-0

2.3 Analysis of nucleosides and nucleotides in plants: An update on sample preparation and LC–MS techniques

Henryk Straube[1], Claus-Peter Witte[1] and Marco Herde[1]

[1] Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany

Analysis of Nucleosides and Nucleotides in Plants: An Update on Sample Preparation and LC–MS Techniques

Henryk Straube [ID], Claus-Peter Witte *[ID] and Marco Herde *[ID]

Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, 30419 Hannover, Germany; straube@pflern.uni-hannover.de
* Correspondence: cpwitte@pflern.uni-hannover.de (C.-P.W.); mherde@pflern.uni-hannover.de (M.H.); Tel.: +49-511-762-4578 (C.-P.W.); +49-511-762-4760 (M.H.)

Abstract: Nucleotides fulfill many essential functions in plants. Compared to non-plant systems, these hydrophilic metabolites have not been adequately investigated in plants, especially the less abundant nucleotide species such as deoxyribonucleotides and modified or damaged nucleotides. Until recently, this was mainly due to a lack of adequate methods for in-depth analysis of nucleotides and nucleosides in plants. In this review, we focus on the current state-of-the-art of nucleotide analysis in plants with liquid chromatography coupled to mass spectrometry and describe recent major advances. Tissue disruption, quenching, liquid–liquid and solid-phase extraction, chromatographic strategies, and peculiarities of nucleotides and nucleosides in mass spectrometry are covered. We describe how the different steps of the analytical workflow influence each other, highlight the specific challenges of nucleotide analysis, and outline promising future developments. The metabolite matrix of plants is particularly complex. Therefore, it is likely that nucleotide analysis methods that work for plants can be applied to other organisms as well. Although this review focuses on plants, we also discuss advances in nucleotide analysis from non-plant systems to provide an overview of the analytical techniques available for this challenging class of metabolites.

1. Introduction

Nucleotides (NTs), nucleosides (Ns), nucleobases (Nbs), and many derived compounds, for example, nucleotide sugars and nucleotide-containing cofactors, are central metabolites in all organisms (Figure 1). The metabolism of nucleotides in plants and their physiological functions, including those beyond being building blocks of nucleic acids, were recently reviewed [1]. For the investigation of nucleotides in plants, the availability of methods for their comprehensive analysis and robust quantification in plant extracts is pivotal. In this review, we summarize the state of the art for NT and N analysis in plants and also discuss methods used in other organisms that may be applicable for plants. Arguably, the most powerful equipment for metabolite analysis is a chromatographic separation system coupled to an electrospray ionization (ESI) mass spectrometer (MS). This technology is in the focus here. However, our recent study [2] also emphasized the importance of sample preparation for comprehensive NT and N analysis; therefore, we discuss the entire workflow of NT and N analysis in all its aspects—from sample preparation to mass spectrometry—and we comment on how the different steps of the workflow influence each other.

The choice of analytical strategies is naturally determined by the physicochemical properties of NTs and Ns. These are polar compounds and the phosphate group(s) of the NTs deprotonate and carry negative charges (pKa values of about 2) [3] within a broad range of pH values. NTs and Ns contain primary, secondary, and tertiary amines,

which, especially under more acid conditions, can protonate and acquire a positive charge especially under more acidic conditions [3]. Due to these characteristics, Ns and especially NTs are better soluble in polar (aqueous) solvents.
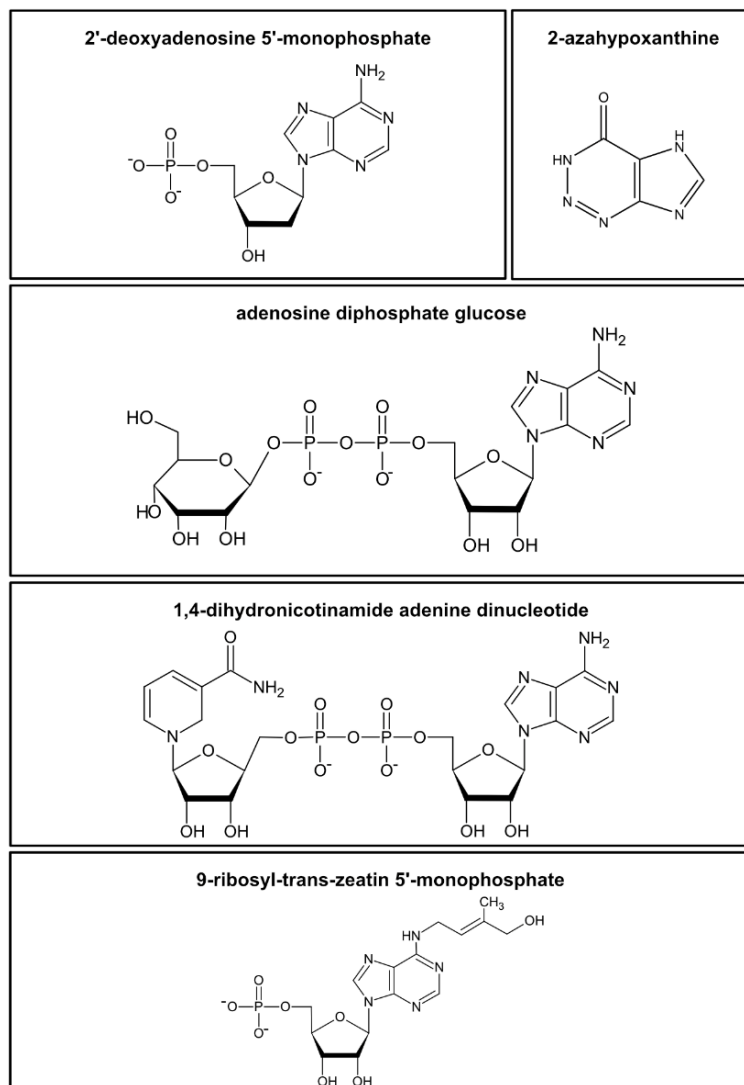

Figure 1. Examples of the diversity of nucleotides and structurally related metabolites in plants. Examples are shown for a nucleotide monophosphate (2′-deoxyadenosine 5′-monophosphate), a "fairy chemical" (2-azahypoxanthine), a nucleotide sugar (adenosine diphosphate glucose), a nucleotide cofactor (1,4-dihydronicotinamide adenine dinucleotide, NAD), and a cytokinine-ribotide (9-ribosyl-trans-zeatin 5′-monophosphate).

The concentrations of different NTs vary significantly in plant cells. For example, ribonucleotide triphosphates are about 1000-fold more abundant than deoxyribonucleotide triphosphates [2]. To detect nucleotides of low abundance, additional techniques for enrichment are usually required.

We noticed that methods for nucleotide analysis that are well established in some non-plant systems are not suitable for plants. For example, the simple workflow we applied for the analysis of deoxyribonucleotides (dNTs) in *Drosophila* [4] failed to work for samples from *Arabidopsis thaliana*. For the detection of low-abundance nucleotides such as dNTs in phylogenetically diverse plant species, the development of a more complex method was necessary [2]. We can only speculate why the analysis of plants is more complicated, but one contributing factor may be the complexity of the plant matrix which contains a plethora

of secondary metabolites [5]. These are probably the reason for the strong ion suppression effects in the ESI source, called matrix effects (ME), which decrease the sensitivity and are usually observed when analyzing plant extracts [2]. Thus, a highly sensitive method for the analysis of the plant NT and N metabolome must include steps to separate these from other interfering metabolites before the chromatographic step. However, if only the more abundant ribonucleosides (rNs) and ribonucleotides (rNTs) are to be monitored, methods for the analysis of the polar metabolome that do not require such extensive sample preparation are sufficient [6–8]. For many years, distinct analysis methods for the NT and N metabolome have been evaluated and improved. For a brief overview of the milestones in plant nucleotide analysis, see Table 1. In the following sections, we discuss the workflow for modern NT/N analysis step by step.

Table 1. Milestones in the development of a workflow for the comprehensive analysis of nucleotides, nucleosides, and nucleobases in plants.

Year	Technique	Achievement	Reference
1964	Different extraction methods compared by enzymatic assays	Identification of plant-specific problems during the extraction of phosphorylated metabolites	[5]
1972	Acid quenching combined with thin-layer chromatography	Comprehensive identification of radiolabeled nucleotide triphosphates in plants	[9]
1975	Acid quenching combined with ion-exchange LC [a] coupled to a UV [b]-detector	Establishment of a workflow that allows the proper quenching of enzyme activity by acidic conditions coupled with a liquid–liquid extraction removing the acid from the extractant, allowing for chromatographic separation by ion-exchange LC	[10]
1991	Acid quenching combined with β-elimination and LC–UV	Analysis of ribo- and deoxyribonucleotides in plants by LC–UV	[11]
2010	Graphitized carbon SPE [c] combined with porous graphitized carbon chromatography MS [d]	Comprehensive analysis of ribonucleotides and nucleotide sugars in plants utilizing porous graphitized carbon chromatography	[12]
2011	SPE combined with liquid chromatography and time-of-flight MS	Simultaneous analysis of 23 nucleotides and cofactors	[13]
2013	SPE combined with LC–MS	Comprehensive analysis of nucleosides and nucleobases	[14]
2020	SPE combined with LC–MS	Comprehensive analysis of nucleotides and nucleosides	[2]

[a] liquid chromatography; [b] ultraviolet; [c] solid-phase extraction; [d] mass spectrometry.

2. Disruption of the Tissue

Because plant cells have a cell wall, efficient physical disruption is required, often using a mortar and pestle or an oscillating mixer mill (Figure 2A). The NTs and Ns are best quantified by the isotope dilution technique [2,15,16]. This accounts for analyte losses in the workflow after tissue disruption, as well as the matrix effect (see later). The efficiency of the tissue disruption itself cannot be assessed because the isotope standards are added to the extraction buffer. Incomplete tissue disruption can, therefore, cause an underestimation of the actual analyte concentration in vivo and result in an increased standard deviation because the degree of disruption is usually not reproducible.

Tissue disruption mixes metabolites and enzymes from all cellular compartments, which can result in metabolite instability, for example, because cytosolic NTs are dephosphorylated by vacuolar phosphatases. It is, therefore, crucial to stop all enzymatic activities when the cellular integrity is lost. This is usually achieved by deep-freezing the sample during tissue disruption with liquid nitrogen. Alternatively, the sample may first be freeze-dried before grinding it to a fine powder in a dry state at room temperature. In our hands, the NT extraction with an aqueous solvent is similarly efficient for freeze-dried

material of *Arabidopsis thaliana* compared to frozen fresh material (our data, unpublished), which is consistent with data from yeast [17]. Enzymes may be reactivated when frozen or freeze-dried samples thaw or come in contact with an aqueous extraction buffer.
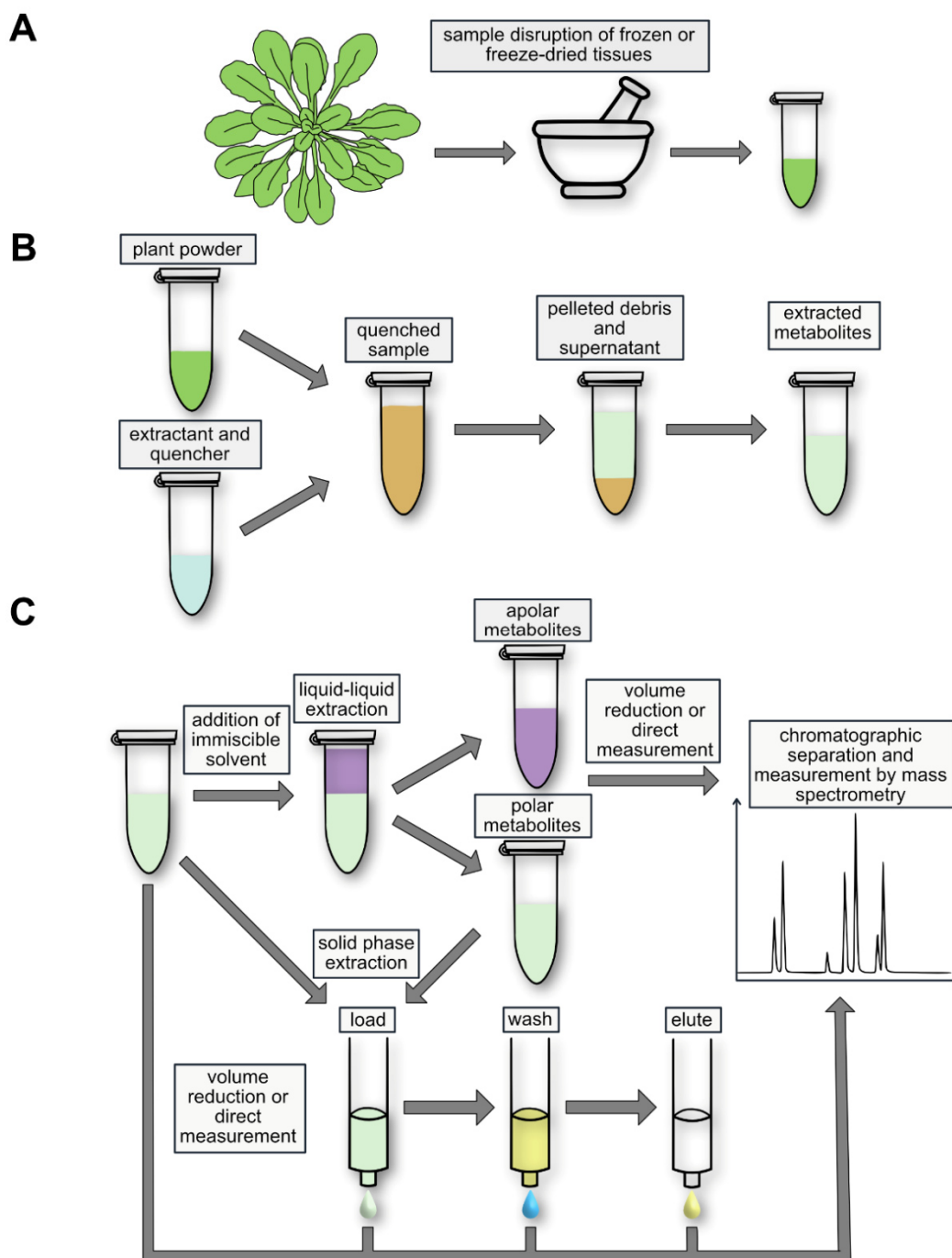

Figure 2. Schematic representation of a typical workflow for the analysis of nucleotides and nucleosides in plants. (**A**) Tissue disruption, here symbolized by mortar and pestle, is followed by (**B**) quenching of the sample and extraction of the metabolites. (**C**) Subsequently, non-polar and polar metabolites are separated by liquid–liquid extraction (LLE). The fraction containing polar metabolites can be further purified to enrich for nucleotides by solid-phase extraction (SPE) with a weak anion-exchange column. In this procedure, nucleotides are separated from nucleosides that are present in the flow-through. Extracts containing nucleosides and nucleotides can be separately analyzed using liquid chromatography and mass spectrometry.

3. Quenching of the Sample

Metabolomic measurements often attempt to estimate the pool sizes of metabolites in vivo. This is only possible if enzymatic processes that can alter metabolite concentrations are irreversibly suppressed (quenched) during and after extraction (Figure 2B). Non-polar metabolites are frequently extracted with mixtures of a high proportion of an organic solvent and some water. Such extractants efficiently quench the sample by precipitation of proteins. However, for NT extraction, organic–aqueous mixtures are problematic because it has been noted that phosphatases, which can dephosphorylate NTs, are not fully quenched in such solvents [5,18–20]. In addition, methanol or ethanol applied cold or boiling can lead to the formation of methyl or ethyl phosphate, which might be derived from nucleophilic attack of the alcohol on the anhydride bonds of the NTs [5,18,21]. Rapid boiling of the ground sample in aqueous buffers without alcohols can be used for Ns [22], but this method is also not suited for NT analysis [2,23]. However, if organic–aqueous mixtures or water are to be used, phosphatases can be inhibited by sodium fluoride [12]. Nonetheless, the effectiveness of the inhibitor should be tested.

A very efficient method for quenching is protein precipitation by addition of a strong acid [5,9,23–25]. Frequently, perchloric acid (PCA) or trichloroacetic acid (TCA) is used. A disadvantage of PCA is that it precipitates upon neutralization which may cause some loss of analytes, especially in samples with high protein content [2,26]. TCA can be removed and neutralized by liquid–liquid extraction (LLE) using a trioctylamine-containing organic phase that is immiscible with water [10]. Drawbacks of this method are the additional hands-on time and potential chemical alterations of metabolites when kept at extreme pH [5]. However, swift handling and cold temperature lead to good recovery rates using TCA coupled to an LLE. The LLE additionally removes undesired non-polar metabolites, further reducing matrix complexity [2]. For all quenching methods, especially those using harsh conditions (such as heat or low pH), it is mandatory to determine whether NTs and Ns are quantitatively recovered and not chemically modified or degraded by the procedure.

4. Extraction of NTs and Ns from Plant Samples

Good solubility of the metabolites in the extraction solvent is crucial for efficient extraction. NTs and Ns are polar and, at least in the case of NTs, also charged, requiring a polar solvent for efficient extraction (Figure 2B). Often the extractant is aqueous [2,22], but water/acetonitrile mixtures are in use as well [27–29]. The solubility of the analytes in the solvent should be considered when choosing the extractant volume in relation to the sample amount, especially when less soluble Nbs [3] are extracted. If solubility is limiting, re-extraction of a pellet will result again in high analyte concentration in the supernatant. The solubility is also dependent on the pH of the extractant. For the extraction of polar metabolites, a 1:10 ratio of fresh material to solvent is recommended, whereas a ratio of 1:100 is recommended for lyophilized material [30].

Extraction and all following steps should be designed to stabilize the analytes. This is essential for NTs, since they may lose a phosphate moiety, for example, during chromatography [31]. One study found that acidification of the extractant can substantially improve the stability of NTs [27]. Our recent work confirmed that, at least for the duration of the extraction, NTs and Ns are acid-stable with good recoveries [2]. Although, more studies have used acidic conditions for NT extraction, there are protocols for the determination of rNTs in *Nicotiana tabacum*, *Arabidopsis thaliana*, and *Solanum lycopersicum* using quenching with a high pH, which also led to excellent recoveries [32]. At high pH, NTs were reported to be stable even when stored for 3 days at room temperature [32]. In general, frozen nucleotide solutions are stable for a prolonged time if the acid from the quenching step is removed or neutralized [33]. With methods that directly use clarified crude extracts for MS analysis, extraction may result in the release of metabolites that cause ion suppression and, thus, decrease the sensitivity in mass spectrometry (matrix effect) [27]. A further concern is that the concentrations of NTs and Ns may be overestimated if the extraction procedure can potentially release them from other metabolites. The most obvious example

is the hydrolytic release of phosphate groups from nucleotide triphosphates (NTPs), reducing NTP concentrations and leading to an overestimation of nucleotide diphosphates (NDPs), nucleotide monophosphates (NMPs), or Ns. Another example is the release of uridine diphosphate (UDP) from UDP-glucose, since the hydrolysis of this compound was observed under mild acidic conditions and at an elevated temperature in nucleotide sugar analysis [34].

Adding a non-polar solvent immiscible with water to the extraction will result in the partitioning of Ns and NTs to the polar/aqueous phase and of non-polar metabolites to the non-polar phase (Figure 2C). Separation and individual processing of both phases results in a coarse fractionation of metabolites which is used in studies simultaneously reporting polar and non-polar metabolites [35]. As a non-polar solvent immiscible with water, methy-*tert*-butyl-ether (MTBE) is a less toxic alternative to fluorocarbons such as 1,1,2-trichlor-1,2,2-trifluorethan used previously [10] and is, therefore, more environmentally friendly and safer to handle [35]. Moreover, dichloromethane and water-saturated diethyl ether have been used as less toxic alternatives [2,33]. Hexane or chloroform was employed in some studies for the removal of phospholipids by LLE [36,37]. The addition of polyvinylpolypyrrolidone can remove phenolic compounds that cannot be eliminated by other means and can interfere with the detection of NTs [11,38].

The rNTs and rNs, but not their dNT and deoxyribonucleoside (dN) counterparts, have two neighboring hydroxyl groups (*cis*-diol) in the sugar moiety. This feature can be used either for the extraction of rNTs and rNs (see next chapter) or for their selective degradation by periodate and methylamine-driven β-elimination [11,39–41]. Selective degradation of the more abundant rNTs and rNs can enhance the detection of dNs and dNTs by reducing the specific NT/N background. This can be explained, for example, by less competition for binding to the resin during solid-phase extraction (next chapter) or simply by less ion suppression in the ESI source. Selective *cis*-diol degradation may also be an improvement for detecting NTs and Ns with modified sugars such as 2'-*O*-methylguanosine, because these will be protected from degradation. When using the selective degradation protocol, it has been recommended to first determine the recovery rates of the metabolites in focus, because, even though dNTs are generally not degraded, deoxyguanosine triphosphate (dGTP), for example, can react with periodate [42].

5. Solid-Phase Extraction

Solid-phase extraction (SPE) is a cartridge chromatographic technique where the interplay between a solid phase of certain binding characteristics and a mobile liquid phase for washing and elution is important to obtain the desired separation results (Figure 2C). The advantage of using SPE is that the analytes of interest are separated from unwanted matrix and are potentially concentrated on the resin. Both effects increase the sensitivity for the analytes in the mass spectrometer. A disadvantage of SPE is that it is laborious to set up. Different solid and mobile phases must usually be tested, and careful adjustment of all parameters such as sample volume and composition of the mobile phases for loading, washing, and elution is required.

For the analysis of NTs and Ns from plants and algae, we recently described a protocol employing a mixed-mode SPE, which is a weak anion-exchange resin together with a hydrophobic binding component [2]. The NTs were retained, but the uncharged Ns did not bind to the resin and were recovered in the flow-through. Choosing the pH of the mobile phase is important because it can influence the charge of metabolites and, therefore, their retention on the SPE cartridge. To facilitate the choice of the pH, a database of pKa values [43] (http://ibond.nankai.edu.cn, accessed on 18 March 2021) for a variety of metabolites is available. On a weak ion-exchange resin, changing the mobile phase pH is used to elute metabolites, because the pH shift causes a loss of ionization of the stationary phase [44]. The elution in our protocol is based on this principle. The ionic strength of the mobile phase is also an important determinant for retention on an ion-exchange SPE cartridge. We observed that the less charged NMPs were not fully retained when a mobile

phase of higher ionic strength was used, whereas NTPs were still bound quantitatively [2]. This result indicates that further sub-fractionation of the NT metabolome by fine-tuning the ionic strength of the mobile phase may be possible.

Ion exchange has also been employed in other protocols for sample preparation in NT analyses. For example, 23 NTs and cofactors from *Arabidopsis thaliana* samples were isolated by a combination of cation exchange and weak anion exchange [13]. An anion-exchange resin was also applied for magnetic dispersive solid-phase micro extraction of NTs from medicinal plants (*Anoectochilus roxburghii*) in a batch procedure. The surface area of this resin is very high potentially increasing the metabolite yield [44]. Furthermore, a mixed-mode anion-exchange sorbent was used as stationary phase to extract Ns, Nbs, and cytokinins from *Physcomitrella patens* [14].

Apart from ion exchange, other separation principles were also applied. Graphitized carbon as a stationary phase for the SPE of polar compounds has gained popularity [45]. In particular, nucleotide sugars can be successfully extracted with this method using an ion-pairing reagent as eluent [46–48]. Stationary phases with phenyl-boronate groups are uniquely suited to retain sugars with a *cis*-diol group (see Section 4), which can be used not only for the removal or enrichment of rNTs and rNs [49], but also for the extraction of brassinosteroids from *Arabidopsis thaliana* [50]. A recently developed strategy uses molecularly imprinted polymers to provide tailor-made binding sites for binding specific metabolites to a stationary matrix [51]. By this approach, modified Ns and NTs have been extracted from urine to serve as biomarkers [52,53]. The discovery of so-called fairy chemicals in several plants (*Arabidopsis thaliana, Oryza sativa, Solanum lycopersicum*; 2-azahypoxanthine, imidazole-4-carboxamide, and 2-aza-8-oxohypoxanthine [54,55]), which are structurally closely related to Nbs and Ns, involved a fractionation by flash chromatography [56] that is based on a silica stationary phase also known from thin-layer chromatography [57,58]. Once identified, it was shown that these metabolites can also be enriched from a complex matrix by a combination of two different SPE procedures. A so-called hydrophilic–lipophilic balance SPE was followed by a cation-exchange SPE that removed undesired compounds and fractionated the fairy chemicals by using basic and acidic eluents [59].

6. Derivatization

Chemical derivatization uses chemical agents that react with the targeted metabolites and form a predictable product that is quantified instead of the original metabolite (Figure 3). Often, the aim of derivatization is to mask the polar features of a metabolite, which facilitates chromatographic separation and/or detection by mass spectrometry [60]. For NTs, a derivatization protocol with 8-(diazomethyl)quinoline was reported recently. This chemical reacts with the phosphate groups of NTs, forming a product that is less polar (Figure 3A). Such derivatives can be detected with higher sensitivity, even allowing the quantification of rare modified NTs in mammalian cell extracts [61]. In *Medicago truncatula*, this derivatization strategy was used for the sensitive detection of sugar phosphates [62]. Since the derivatization reagent is highly reactive and unstable, the success of this strategy depends in part on the concentration of other phosphorylated metabolites that consume the reagent and, thus, relies on the efficiency of other clean-up steps in the workflow [62].

Another useful derivatization strategy is the esterification of the hydroxyl groups of the sugar moiety with propionic anhydride (Figure 3B), which was used to detect cytokinins and NTs in *Arabidopsis thaliana* with high sensitivity [60]. A very similar approach employing acetic anhydride was used to monitor cytokinins by gas chromatography [63]. Silylation, often used in the context of gas chromatography, can also foster the detection of NTs by LC–MS [64].
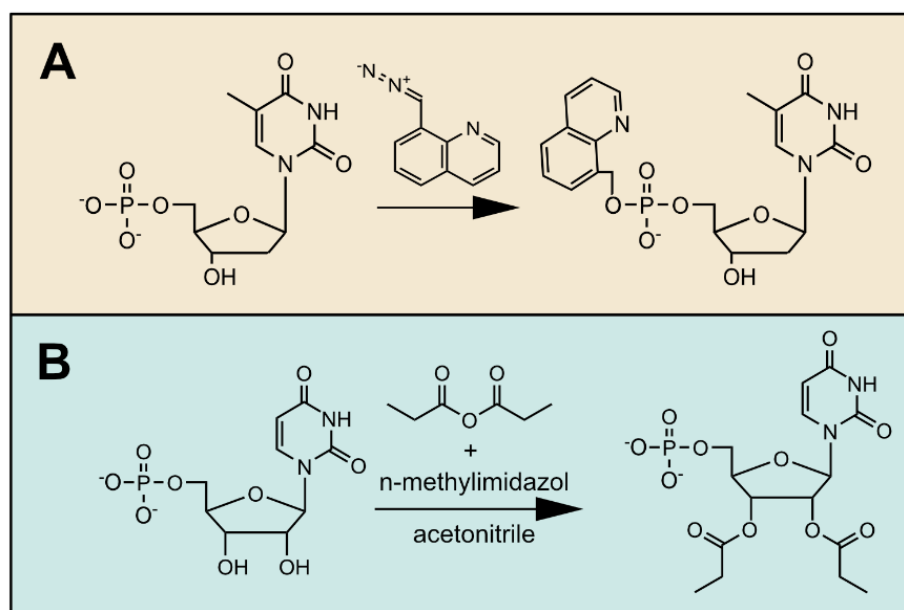
Figure 3. Examples of different derivatization strategies for nucleotides. (**A**) 8-(diazomethyl) quinoline reacts with the phosphate group neutralizing a negative charge; (**B**) propionic anhydride is used for the derivatization of hydroxyl groups. Both strategies reduce the polarity of the nucleotide and facilitate chromatographic separation, which fosters a more sensitive detection by mass spectrometry.

7. Reduction in Sample Volume

The solvent is often removed to concentrate the sample or to change from a polar to a non-polar solvent prior to the injection in a chromatographic system. If polar metabolites have been extracted, this usually involves the evaporation of an aqueous phase, for example, in a speed vacuum centrifuge. Because of the high boiling point of aqueous solvents, solvent evaporation is a comparatively long procedure that exposes the sample to ambient temperatures for an extended period of time. Degradation of labile compounds might, therefore, be a problem, which can also be accelerated by possible pH changes during the evaporation of the solvent.

For NT analysis, we tested freeze-drying and vacuum centrifugation of samples. Freeze-drying has the advantage that the samples are cooled during the procedure; however, with vacuum centrifugation, which requires about 2 h at room temperature, we also obtained satisfying recoveries of NTs [2]. Organic–aqueous mixtures such as water–methanol for polar metabolites (used, for example, in [35]) are often incompatible with freeze-drying equipment. Such extracts must be dried down in a nitrogen flow evaporator or a speed vacuum centrifuge. The dried pellets are then dissolved in polar or non-polar solvents depending on the type of chromatography that follows. The sample volume and the composition of the injection solvent are critical for the chromatographic separation. For an HILIC (hydrophilic interaction liquid chromatography) stationary phase, the proportion of organic solvent in the injection solution has been optimized to ensure maximal solubility of NTs without compromising the chromatography [31].

8. Chromatographic Separation

Sample preparation techniques such as LLE and SPE greatly reduce the complexity of a sample. The chromatography then separates the sample analytes from each other so that they ideally arrive at different times at the ESI source of the MS. Together, these techniques ensure that the MS processes as few metabolites/ions as possible at any given time. This is important because it greatly improves the sensitivity (see next section). Furthermore, the unequivocal identification of analytes benefits from sample preparation and especially chromatography. For example, naturally occurring heavy isotopes of adenosine

monophosphate (AMP) cannot be distinguished from inosine monophosphate (IMP) with a triple-quadrupole mass spectrometer. This is because they have nearly the same mass (i.e., they are isobaric), and such compounds often have even identical fragmentation patterns [65]. It is, therefore, essential that they are chromatographically separated and reach the MS at different times.

Basic concepts of liquid chromatography were reviewed elsewhere [66]. Here, we would like to focus on systems that are especially useful for the analysis of NTs and Ns. The well-established reverse-phase chromatography uses non-polar stationary phases, for example, C8 or C18 columns, which are not able to retain Ns and NTs. Modifications of such hydrophobic stationary phases by adding polar groups—either at the silanol groups (end-capping) or directly embedded as ligands—result in the increased retention of polar and decreased retention of non-polar metabolites [67–69]. For the separation of Ns, for example, the non-canonical dNs [70], such matrices have been used successfully, but the retention of NTs is usually poor. This problem can be addressed by adding an ion-pairing agent to the mobile phase, which promote the retention of NTs on a hydrophobic stationary phase. However, the classical ion-pairing agents are not volatile and contaminate the source of the MS. They are also associated with suppression of the signal [71,72]. Interestingly, the development of volatile ion-pairing agents with favorable characteristics for mass spectrometry recently increased the interest in reverse-phase chromatography for the separation of charged compounds [73], and applications for NTs also exist [74].

A stationary phase, especially useful for the analysis of polar and charged metabolites, is porous graphitized carbon (PGC). Such a matrix separates polar and charged substances in a reverse-phase mode fully compatible with mass spectrometry. The main disadvantage is a certain instability of retention times which may hamper the identification of compounds, especially in non-targeted metabolome studies. It was reported that the redox status of the column can interfere with retention times [12]. In our recent study, we used PGC chromatography for NT and N analysis [2] and did not observe marked retention time instability. Although we did not address this formally, we found that samples processed by SPE ran more reproducibly on PGC chromatography than unprocessed samples. Generally, thorough column equilibration seems to improve retention time stability. PGC chromatography is a powerful tool for NT and N analysis [2], which can also separate isobaric 3'- and 5'-nucleotide monophosphates (our data, unpublished). As mentioned above, the chromatographic separation of isobaric NTs is very important to avoid that metabolites of nearly identical mass are wrongly assigned and quantified, which could lead, for example, to confounding dGTP with adenosine triphosphate (ATP) [12] or IMP with AMP isotopes. Furthermore, several publications have shown that PCG-chromatography is able to separate nucleotide sugars [12,46,47].

An alternative separation technique for the polar metabolome is hydrophilic interaction liquid chromatography (HILIC). Many polar and charged metabolite classes such as Ns, NTs, and amino acids can be analyzed with this stationary phase. Recently, we were also able to separate Nbs from *Arabidopsis thaliana* extracts with HILIC, as well as with a modified C18 column (our data, unpublished). Different HILIC phases and their respective potential for separating plant metabolites were excellently reviewed elsewhere [30]. The advantages of HILIC are the high tolerance to salts, the excellent reproducibility of retention times, the even peak shape, and the stability over a wide pH range, allowing the development of low-pH and high-pH methods [75]. While the low-pH methods are well suited for amino acids, vitamins, and polyamines, the high-pH protocols are preferable for carbohydrates, carbohydrate phosphates, organic acids, and NTs [76–78]. Additionally, nucleotide sugars can be analyzed with HILIC [48,79]. Samples for HILIC must contain a relatively high proportion of organic solvent, which is usually achieved by drying the sample and redissolving the pellet. In this process, it is important to consider the limit of solubility for the polar and charged analytes in a less polar solvent. A noisy baseline and peak tailing are sometimes observed in HILIC, especially for phosphorylated and carboxylated metabolites, due to their interaction with metallic surfaces. These problems can be

addressed with deactivator additives, high-pH buffers, PEEK (polyetheretherketone)-lined columns, and PEEK-lined instruments [75,80].

Another option for the MS-compatible chromatography of polar and charged metabolites is the use of a silica-hydride stationary phase in aqueous normal-phase chromatography. These provide a high degree of polar selectivity and are also able to retain non-polar metabolites. With this approach, organic acids, NTs, and amino acids were successfully analyzed [81–83], but applications in plants are still scarce. However, recently, Ns from *Asparagus officinalis* extracts were successfully separated from other polar metabolites using this technique [6].

An emerging technology for the separation and detection of charged metabolites is capillary electrophoresis mass spectrometry (CE–MS) which is not yet widely in use. This technology is especially advantageous for NTs when limited sample material is available [84]. A recent study reported that the reproducibility of CE–MS among several laboratories is very good, suggesting a high potential for this technique in the future [85].

Lastly, ion chromatography (IC) using an ion exchange stationary phase is highly suited to separate cations and anions. Elution of bound analytes from ion exchange matrices usually requires a rising concentration of competing ions. Phosphate is a very popular anion classically used in anion exchange chromatography; but phosphate-based buffers are not compatible with MS. By producing potassium hydroxide as an eluent in situ via electrolysis and suppressing the eluent ions after the chromatographic separation, IC has become highly reproducible and compatible with MS. However, a comprehensive analysis of the NT/N metabolome is not straightforward with IC, because for NTs different columns and running parameters are required than for Ns. Nonetheless, NTs from plants and other organisms were successfully separated and comprehensively analyzed with IC–MS [7,86–88]. One challenge is that the analytes leave the instrument in a pure aqueous solution, but ionization in an electrospray ionization (ESI) source benefits from the presence of a polar organic solvent because the surface tension of pure water is too high for establishing a stable ESI spray [89]. A significant volume of organic solvent [89] or other modifiers are usually added to the analyte stream prior to the MS with a T-liquid junction. This reduces the chromatographic resolution and potentially decreases the sensitivity. Instead of a T-liquid junction, a sheath liquid interface can be used to attenuate these effects. Such an interface was employed for the analysis of sugars [90]. An alternative IC works with a weak anion exchange column and an inverse acetate gradient that is fully compatible with MS [91]. However, the stationary phase required for this technique is less durable and, currently, the availability of this matrix from commercial suppliers is limited.

9. Mass Spectrometry

The signal strength and, thus, the sensitivity of the measurement not only depend on the abundance of the metabolite but also on (i) the ionization efficiency (IE) in the ESI source (Figure 4A), (ii) the matrix effect (ME), which is the suppression of the ionization of the target analyte by other metabolites that co-elute in the chromatography, and (iii) losses due to undesired degradation in the source (in-source decay) or inefficient fragmentation in the collision cell [89]. Other factors such as the ion transmission through the MS are also crucial for sensitivity, but they depend mostly on the technical specifications of the instrument and cannot be influenced by the operator. Polar metabolites such as NTs and Ns generally have a poor surface activity (i.e., the tendency of an analyte to be present at the surface of ESI droplets (Figure 4B)) and, therefore, have a lower ionization efficiency [89]. This results in strong ion suppression of polar metabolites by co-eluting less polar compounds [92], emphasizing the importance of selectivity in the sample preparation [93] and of resolution in the chromatography. In our analysis of NTs and Ns, we determined the ME [94] with isotope standards and concluded that the SPE significantly reduced the ME [2]. In another study, the impact of the chromatography on the ME was determined, suggesting that a separation with an HILIC stationary phase is suitable to lower the ME for polar metabolites [95].
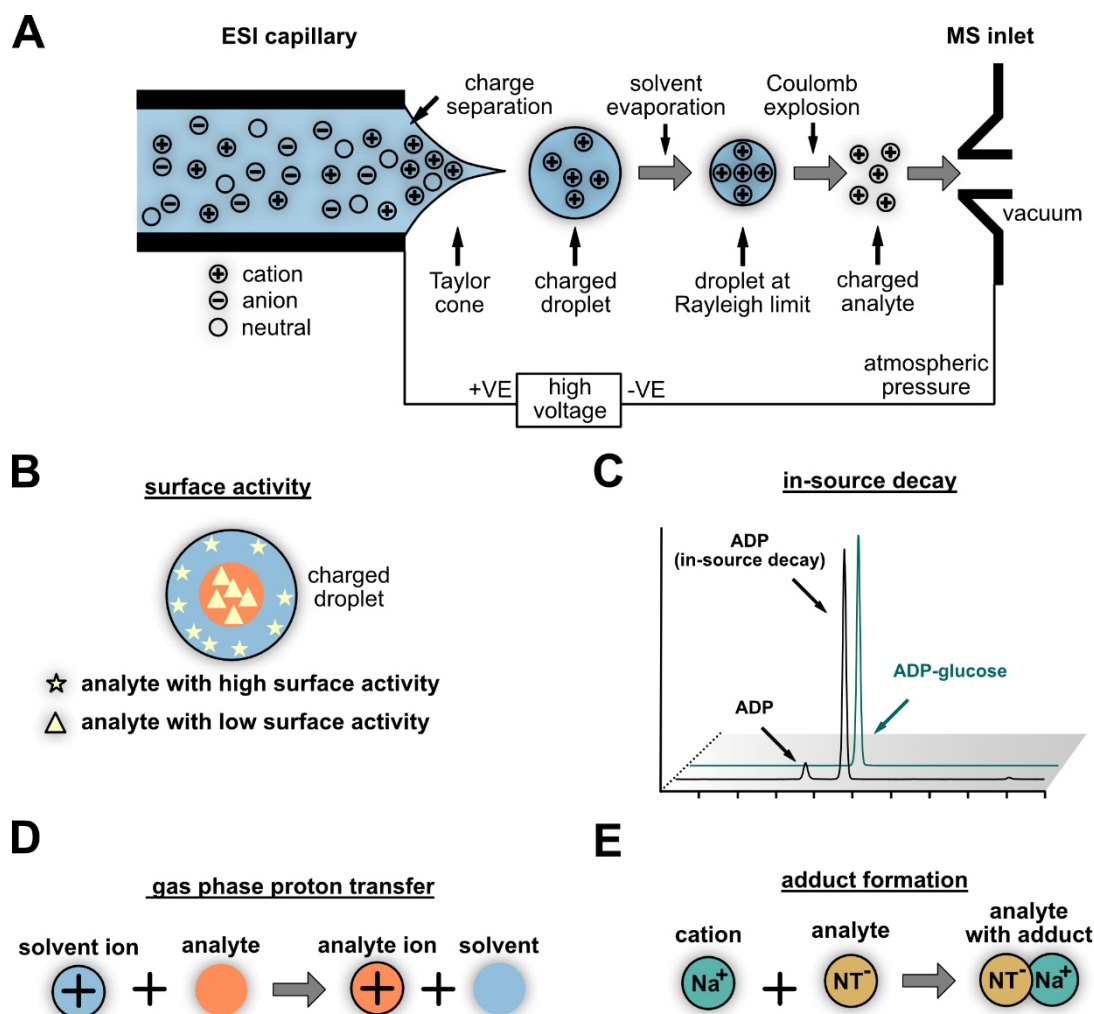
Figure 4. Processes in the electrospray ionization source that result in the formation of ions or influence the sensitivity of detection. (**A**) Charge separation during droplet formation and subsequent solvent evaporation results in charged analytes which can be detected by the mass spectrometer. (**B**) Analytes are not evenly distributed in the droplets. Depending on their surface activity, they locate more in the center (analytes with higher polarity; triangles) or nearer to the surface (analytes with lower polarity; stars) of the droplet. Less polar analytes with strong surface activity have a higher ionization efficiency than more polar analytes in the center of the droplet. (**C**) Decay of the analyte in the source results in signal loss. Adenosine diphosphate (ADP) is chromatographically separated from ADP-glucose in this example, but the fragmentation of ADP-glucose in the source results in the formation of ADP that appears to co-elute with ADP-glucose. If this is not noted, the amount of ADP-glucose can be underestimated (**D**). An additional process leading to ionization is the gas phase proton transfer that is determined by the proton affinities of the solvent, the modifiers, and the analytes. (**E**) The formation of adducts with, e.g., sodium or potassium, often results in the formation of two or more species from the analyte with different masses. Splitting the signal of the analyte usually leads to decreased sensitivity.

In general, NTs can be effectively measured in the negative and positive polarity mode, but the tandem MS fragmentation spectra are more informative in the negative mode [96]. Depending on the pH of the solvent and the pKa of the metabolite, cations or anions are formed in solution. Upon charge separation (Figure 4A), these are measured in the MS in positive or negative mode, respectively [89]. Over a wide range of pH values, NTs are anions, which suggests that a detection in negative mode is the most straightforward choice. However, the negative mode requires a charge carrier in the solvent, especially when the analyte concentration is low. A mixture of hexafluoroisopropanol and methanol was used successfully as a charge carrier for the sensitive detection of oligonucleotides [97–99]. Additionally, 2,2,2-trifluoroethanol and formaldehyde were shown to reduce noise and

increase ionization fostering a very sensitive detection of oligonucleotides [100]. One would expect that a low pH facilitates the protonation of metabolites, thus promoting the detection in the positive polarity mode, whereas a high pH results in deprotonation, boosting the charge separation of preformed anions in the negative polarity mode [89]. Interestingly, NTs are often measured at a neutral or high pH in the positive polarity mode [2,76,101,102]. This "wrong-way-round ionization" [103] also occurs for other metabolites. Ammonia, used in many mobile phases for nucleotide analysis, is hypothesized to be a gas-phase proton donor that facilitates ionization in the positive mode [89,103]. The gas-phase proton affinities, crucial for effective ionization [104], were determined for NMPs and are lower for pyrimidine NTs compared to purine NTs [105], which is consistent with our observation that, in positive polarity mode, pyrimidine NTs are detected with less sensitivity than purine NTs (our data, unpublished).

In addition to charge separation of preformed ions (Figure 4A) and gas-phase proton transfer (Figure 4D), ionization by reduction and oxidation processes, as well as adduct formation (Figure 4E), is possible. Adduct formation usually results in an altered precursor mass and a distribution of the signal between the parent ion and its adducts. This generally reduces sensitivity and complicates the interpretation of data from non-targeted approaches [106,107], although adduct formation can increase sensitivity for some carbohydrates [108]. NTs are electrostatically attracted to, e.g., sodium and potassium, and strategies to attenuate the formation of cation adducts by preconditioning the chromatographic system were shown to increase the sensitivity for the detection of oligonucleotides [106]. Moreover, the elimination of sodium originating from glassware can reduce the formation of sodium adducts [89], and adjusting the ESI source parameters (cone voltage) might disfavor adduct formation [107]. A significant amount of alkali metals can originate from the sample, and their concentration may vary depending on the environmental conditions in which the sampled plants have grown [109]. Sample preparation using an SPE approach (e.g., with a weak anion-exchange stationary phase) can potentially reduce the abundance of alkali metals in the matrix. This might be particularly important for HILIC, where NTs and Ns can co-elute with alkali metal ions. Adduct formation has been demonstrated for guanosine in HILIC [110].

Oxidation and reduction reactions in the ESI source resulting in radical formation have been described for plant metabolites [111], but it is unknown if NTs and Ns are affected by such processes. However, the oxidation of NTs can be provoked by an on-line electrochemical cell prior to ESI to simulate the occurrence of damaged or modified NTs [112].

In addition to adduct formation and potentially oxidation/reduction processes, the fragmentation of the NTs and Ns in the ESI source (in-source decay) can result in a loss of analyte (Figure 4C). Often, the glycosidic bond is affected [96,113], which creates the false impression that an Nb co-elutes with the corresponding N or NT (our data, unpublished, Figure 4C). A high nozzle voltage can foster in-source decay, but if this source parameter is set too low, the ionization of the analyte might be reduced [89]. In-source fragmentation can also lead to over- or underestimation of metabolite abundance, for example, when AMP, adenosine diphosphate (ADP), and ATP are not baseline-separated in the chromatography and source parameters are used that cause a decay of ATP to ADP or AMP within the source [114].

In tandem mass spectrometry, single reaction monitoring (SRM) is used, in which a precursor ion is first selected for fragmentation in the collision cell, and then the resulting product ions are quantified [115]. The value of the collision energy voltage is an important parameter for this process and needs to be optimized in the method development [116]. Interestingly, NTs exhibit characteristic fragmentation pathways that follow general rules depending on the nucleotide structure [96]. This makes the occurrence of product ions for NTs predictable, which can guide the development of an SRM assay. In addition, data on experimentally observed product ions for NTs, including many modified NT species, can be obtained from a public database [96] (www.msTide-db.com, accessed on 18 March 2021).

How do different mass spectrometers compare in performance for NT analysis? Triple-quadrupole (QqQ) instruments are traditionally considered the gold standard for the sensitive detection and quantification of metabolites. Because the mass resolution of such instruments is limited, they strongly rely on fragmentation patterns for the confident identification of Ns and NTs in a plant matrix [2,7,8,14]. Alternatively, high-resolution mass spectrometers (HRMS) provide an exact mass not only for the targeted analyte but also for all metabolites within a specified m/z (mass to charge) window, which is a tremendous advantage for non-targeted analyses and troubleshooting. In contrast to QqQ, HRMS spectra can later be reanalyzed for metabolites that were not initially in the focus of the study. With the Orbitrap and time-of-flight (TOF) mass spectrometers, two different instrument types for HRMS are available that show a similar performance for plant metabolomics [117]. Both systems can be combined with a quadrupole and a collision cell (Q-TOF and Q-Orbitrap) to select precursors and create fragment ions. For the analysis of NTs in plants, both instrument types have been used [2,12,13]. One drawback of HRMS is a reduced sensitivity compared to QqQ instruments [118]. In our recent study, we were able to detect all low-abundance dNTs in plant matrices with a QqQ instrument, but with a Q-Orbitrap some were not detectable at all or only when using the quadrupole of the Orbitrap as a narrow mass filter focusing directly on the ion mass of the dNT in question [2]. The sensitivity of an HRMS is negatively affected by the matrix. Interestingly, this not only depends mainly on ESI processes as for QqQ instruments, but also on the total ion load entering the machine at any time [118]. Reducing the ion load by selective sample preparation and good chromatographic separation is, therefore, of additional importance for sensitivity when using HR instruments. In addition to reducing the ion load, increasing the ion mass by derivatization of Ns and NTs might increase the sensitivity, because most matrix components have comparatively low m/z values, and these could be excluded by appropriate setting of the quadrupole mass filter in Q-TOF or Q-Orbitrap machines.

10. Outlook

Although the comprehensive quantification of Ns and NTs is now possible in samples from a variety of plants and algae, many challenges still remain. While current workflows can detect the relatively low abundant dNs and dNTs in plants [2,11], reports about the detection of damaged or modified NTs which are probably even less abundant are still scarce [119,120]. Since such rare NTs are more frequently observed in non-plant systems [61,121], it is tempting to speculate that it is the complexity of the plant matrix which hampers their detection in plants. It will be necessary to boost the sensitivity to also detect these rare NTs, which may be achieved in part by further reducing the complexity of the matrix, for example, by eliminating the more abundant rNTs [11,41]. Additionally, up-scaling of the sample amount in combination with enrichment techniques and derivatization protocols might be part of a solution.

Robust sample preparation coupled with a high sensitivity of detection will improve the analysis of NTs and Ns in cases where only little starting material is available, for example, when only a certain plant tissue is investigated. It will also be feasible to analyze plants with a particularly complex metabolome and correspondingly complex matrix (e.g., in *Viscum album* [122]).

Some rNTs and rNs but not the dNs and dNTs or the modified NTs were previously reported in studies aiming at the description of the whole (polar) metabolome [6,7]. Currently, there is a tradeoff between the depth of NT analysis on one side and the ability to comprehensively describe the entire metabolome on the other side. Coupling our sample preparation protocol for NT and N analysis [2] with protocols for the simultaneous analysis of amino acids, phytohormones, and lipids [35] may help to overcome this problem. Although the sample preparation would be more time-consuming, the preparation of a single sample may be sufficient for a more comprehensive in-depth analysis of the metabolome. Potentially, several groups of non-polar but also polar and less abundant metabolites such as phytohormones, (d)NTs, and (d)Ns could be quantified from such samples. An

alternative approach uses a combination of reverse-phase and HILIC chromatography to monitor lipids and polar metabolites in one analytical run [123].

A major drawback of LC–MS techniques is that they are unable to monitor NTs in living cells and cannot assess NT concentrations in particular cells within complex tissues or in subcellular compartments. For some abundant NTs such as ATP and nicotinamide adenine dinucleotide, molecular probes were developed to monitor the in vivo concentrations within *Arabidopsis thaliana* cells and tissues [124,125]. Although such probes are currently only available for very few metabolites and are less sensitive than LC–MS (a concentration of ~160 μM can be detected with the ATP sensor, compared to a sensitivity in the picomolar range for LC–MS), these techniques can complement data obtained by LC–MS. Techniques to fractionate [126,127] or isolate subcellular compartments [128] in combination with metabolite analysis via LC–MS may prove useful for the investigation of the subcellular NT and N pools in plant cells. In the future, such fractionation and isolation techniques must gain in resolution. Especially for organelle isolation techniques, quenching strategies must be devised to effectively prevent alterations to the metabolome during organelle preparation.

References

1. Witte, C.-P.; Herde, M. Nucleotide metabolism in plants. *Plant Physiol.* **2020**, *182*, 63–78. [CrossRef]
2. Straube, H.; Niehaus, M.; Zwittian, S.; Witte, C.-P.; Herde, M. Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. *Plant Cell* **2020**. [CrossRef]
3. Hodgson, D.R.W. Physicochemical Aspects of Aqueous and Nonaqueous Approaches to the Preparation of Nucleosides, Nucleotides and Phosphate Ester Mimics. *Adv. Phys. Org.* **2017**, *51*, 187–219. [CrossRef]
4. Liu, B.; Winkler, F.; Herde, M.; Witte, C.-P.; Grosshans, J. A Link between deoxyribonucleotide metabolites and embryonic cell-cycle control. *Curr. Biol.* **2019**, *29*, 1187. [CrossRef]
5. Bielski, R.L. The problem of halting enzyme action when extracting plant tissues. *Anal. Biochem.* **1964**, *9*, 431–442. [CrossRef]
6. Creydt, M.; Fischer, M. Plant metabolomics: Maximizing metabolome coverage by optimizing mobile phase additives for nontargeted mass spectrometry in positive and negative electrospray ionization mode. *Anal. Chem.* **2017**, *89*, 10474–10486. [CrossRef] [PubMed]
7. Rolletschek, H.; Melkus, G.; Grafahrend-Belau, E.; Fuchs, J.; Heinzel, N.; Schreiber, F.; Jakob, P.M.; Borisjuk, L. Combined noninvasive imaging and modeling approaches reveal metabolic compartmentation in the barley endosperm. *Plant Cell* **2011**, *23*, 3041–3054. [CrossRef]
8. De Souza, A.P.; Cocuron, J.-C.; Garcia, A.C.; Alonso, A.P.; Buckeridge, M.S. Changes in whole-plant metabolism during the grain-filling stage in sorghum grown under elevated CO_2 and drought. *Plant Physiol.* **2015**, *169*, 1755–1765. [CrossRef] [PubMed]
9. Nygaard, P. Deoxyribonucleotide pools in plant-tissue cultures. *Physiol. Plant.* **1972**, *26*, 29–33. [CrossRef]

10. Khym, J.X. An analytical system for rapid separation of tissue nucleotides at low pressures on conventional anion exchangers. *Clin. Chem.* **1975**, *21*, 1245–1252. [CrossRef] [PubMed]

11. Dutta, I.; Dutta, P.K.; Smith, D.W.; O'Donovan, G.A. High-performance liquid-chromatography of deoxynucleoside diphosphates and triphosphates in tomato roots. *J. Chromatogr. A* **1991**, *536*, 237–243. [CrossRef]

12. Pabst, M.; Grass, J.; Fischl, R.; Léonard, R.; Jin, C.; Hinterkörner, G.; Borth, N.; Altmann, F. Nucleotide and nucleotide sugar analysis by liquid chromatography-electrospray ionization-mass spectrometry on surface-conditioned porous graphitic carbon. *Anal. Chem.* **2010**, *82*, 9782–9788. [CrossRef]

13. Guerard, F.; Petriacq, P.; Gakiere, B.; Tcherkez, G. Liquid chromatography/time-of-flight mass spectrometry for the analysis of plant samples: A method for simultaneous screening of common cofactors or nucleotides and application to an engineered plant line. *Plant Physiol. Biochem.* **2011**, *49*, 1117–1125. [CrossRef]

14. Kopecna, M.; Blaschke, H.; Kopecny, D.; Vigouroux, A.; Koncitikova, R.; Novak, O.; Kotland, O.; Strnad, M.; Morera, S.; von Schwartzenberg, K. Structure and function of nucleoside hydrolases from physcomitrella patens and maize catalyzing the hydrolysis of purine, pyrimidine, and cytokinin ribosides. *Plant Physiol.* **2013**, *163*, 1568–1583. [CrossRef]

15. Kuskovsky, R.; Buj, R.; Xu, P.; Hofbauer, S.; Doan, M.T.; Jiang, H.; Bostwick, A.; Mesaros, C.; Aird, K.M.; Snyder, N.W. Simultaneous isotope dilution quantification and metabolic tracing of deoxyribonucleotides by liquid chromatography high resolution mass spectrometry. *Anal. Biochem.* **2019**, *568*, 65–72. [CrossRef]

16. Rodríguez-González, P.; García Alonso, J.I. *Isotope Dilution Mass Spectrometry*; Elsevier: Amsterdam, The Netherlands, 2018.

17. Ingle, J. The extraction and estimation of nucleotides and nucleic acids from plant material. *Phytochemistry* **1963**, *2*, 353–370. [CrossRef]

18. Ullrich, J.; Calvin, M. Alcohol-resistant phosphatase activity in chloroplasts. *Biochim. Biophys. Acta* **1962**, *63*, 1–10. [CrossRef]

19. Ullrich, J. Phosphatase action on phosphoglycolic, 3-phosphoglyceric, and phosphoenol pyruvic acids in spinach chloroplast fragments in the presence and absence of high concentrations of methanol. *Biochim. Biophys. Acta* **1963**, *71*, 589–594. [CrossRef]

20. Ikuma, H.; Tetley, R.M. Possible interference by an acid-stable enzyme during extraction of nucleoside diphosphates and triphosphates from higher-plant tissues. *Plant Physiol.* **1976**, *58*, 320–323. [CrossRef] [PubMed]

21. Runeckles, V.C. Formation of alkyl phosphates in wheat leaves. *Nature* **1958**, *181*, 1470–1471. [CrossRef]

22. Baccolini, C.; Witte, C.-P. AMP and GMP catabolism in arabidopsis converge on xanthosine, which is degraded by a nucleoside hydrolase heterocomplex. *Plant Cell* **2019**, *31*, 734–751. [CrossRef] [PubMed]

23. Dietmair, S.; Timmins, N.E.; Gray, P.P.; Nielsen, L.K.; Krömer, J.O. Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol. *Anal. Biochem.* **2010**, *404*, 155–164. [CrossRef]

24. Dahncke, K.; Witte, C.-P. Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in arabidopsis. *Plant Cell* **2013**, *25*, 4101–4109. [CrossRef]

25. Chen, M.; Herde, M.; Witte, C.-P. Of the nine cytidine deaminase-like genes in arabidopsis, eight are pseudogenes and only one is required to maintain pyrimidine homeostasis in vivo. *Plant Physiol.* **2016**, *171*, 799–809. [CrossRef] [PubMed]

26. Chida, J.; Yamane, K.; Takei, T.; Kido, H. An efficient extraction method for quantitation of adenosine triphosphate in mammalian tissues and cells. *Anal. Chim. Acta* **2012**, *727*, 8–12. [CrossRef] [PubMed]

27. Rabinowitz, J.D.; Kimball, E. Acidic acetonitrile for cellular metabolome extraction from Escherichia coli. *Anal. Chem.* **2007**, *79*, 6167–6173. [CrossRef]

28. Au, J.L.; Su, M.H.; Wientjes, M.G. Extraction of intracellular nucleosides and nucleotides with acetonitrile. *Clin. Chem.* **1989**, *35*, 48–51. [CrossRef] [PubMed]

29. Chen, H.; Zhang, B.; Hicks, L.M.; Xiong, L. A nucleotide metabolite controls stress-responsive gene expression and plant development. *PLoS ONE* **2011**, *6*. [CrossRef] [PubMed]

30. Liu, Z.; Rochfort, S. Recent progress in polar metabolite quantification in plants using liquid chromatography-mass spectrometry. *J. Integr. Plant Biol.* **2014**, *56*, 816–825. [CrossRef] [PubMed]

31. Zbornikova, E.; Knejzlik, Z.; Hauryliuk, V.; Krasny, L.; Rejman, D. Analysis of nucleotide pools in bacteria using HPLC-MS in HILIC mode. *Talanta* **2019**, *205*. [CrossRef]

32. Riondet, C.; Morel, S.; Alcaraz, G. Determination of total ribonucleotide pool in plant materials by high-pH anion-exchange high-performance liquid chromatography following extraction with potassium hydroxide. *J. Chromatogr. A* **2005**, *1077*, 120–127. [CrossRef] [PubMed]

33. Brown, P.R. Stability of nucleotide solutions on storage as determined by high-pressure liquid chromatography. *Anal. Biochem.* **1971**, *43*, 305–306. [CrossRef]

34. Barnes, J.; Tian, L.; Loftis, J.; Hiznay, J.; Comhair, S.; Lauer, M.; Dweik, R. Isolation and analysis of sugar nucleotides using solid phase extraction and fluorophore assisted carbohydrate electrophoresis. *MethodsX* **2016**, *3*, 251–260. [CrossRef]

35. Salem, M.A.; Yoshida, T.; de Souza, L.P.; Alseekh, S.; Bajdzienko, K.; Fernie, A.R.; Giavalisco, P. An improved extraction method enables the comprehensive analysis of lipids, proteins, metabolites and phytohormones from a single sample of leaf tissue under water-deficit stress. *Plant J.* **2020**, *103*, 1614–1632. [CrossRef] [PubMed]

36. Soga, T.; Ishikawa, T.; Igarashi, S.; Sugawara, K.; Kakazu, Y.; Tomita, M. Analysis of nucleotides by pressure-assisted capillary electrophoresis-mass spectrometry using silanol mask technique. *J. Chromatogr. A* **2007**, *1159*, 125–133. [CrossRef]

37. Cordell, R.L.; Hill, S.J.; Ortori, C.A.; Barrett, D.A. Quantitative profiling of nucleotides and related phosphate-containing metabolites in cultured mammalian cells by liquid chromatography tandem electrospray mass spectrometry. *J. Chromatogr. B* **2008**, *871*, 115–124. [CrossRef]

38. Nieman, R.H.; Pap, D.L.; Clark, R.A. Rapid purification of plant nucleotide extracts with xad-2,polyvinyl-polypyrrolidone and charcoal. *J. Chromatogr. A* **1978**, *161*, 137–146. [CrossRef]

39. Tanaka, K.; Yoshioka, A.; Tanaka, S.; Wataya, Y. An improved method for the quantitative determination deoxyribonucleoside triphosphates in cell-extracts. *Anal. Biochem.* **1984**, *139*, 35–41. [CrossRef]

40. Uziel, M. Periodate oxidation and amine-catalyzed elimination of terminal nucleoside from adenylate or ribonucleic-acid products of overoxidation. *Biochemistry* **1973**, *12*, 938–941. [CrossRef] [PubMed]

41. Odmark, G.; Kihlman, B.A. Effects of chromosome-breaking purine derivates on nucleic acid synthesis and on levels of adenosine 5′-triphosphate and deoxyadenosine 5′-triphosphate in bean root tips. *Mutat. Res. Fundam. Mol. Mech. Mutagen.* **1965**, *2*, 274–286. [CrossRef]

42. Hennere, G.; Becher, F.; Pruvost, A.; Goujard, C.; Grassi, J.; Benech, H. Liquid chromatography-tandem mass spectrometry assays for intracellular deoxyribonucleotide triphosphate competitors of nucleoside antiretrovirals. *J. Chromatogr. B* **2003**, *789*, 273–281. [CrossRef]

43. Ji, P. *iBonD 2.0-the Most Comprehensive pKa and BDE Database so Far*; Tsinghua University: Beijing, China, 2016. [CrossRef]

44. Chen, X.; Wu, Y.; Huang, L.; Yang, L.; Hong, R.; Yao, H.; Li, S. Magnetic dispersive solid-phase micro-extraction combined with high-performance liquid chromatography for determining nucleotides in anoectochilus roxburghii (Wall.) Lindl. *J. Pharm. Biomed. Anal.* **2019**, *174*, 432–440. [CrossRef] [PubMed]

45. Hennion, M.C. Graphitized carbons for solid-phase extraction. *J. Chromatogr. A* **2000**, *885*, 73–95. [CrossRef]

46. Behmüller, R.; Forstenlehner, I.C.; Tenhaken, R.; Huber, C.G. Quantitative HPLC-MS analysis of nucleotide sugars in Plant Cells following off-line SPE sample preparation. *Anal. Bioanal. Chem.* **2014**, *406*, 3229–3237. [CrossRef] [PubMed]

47. Rautengarten, C.; Heazlewood, J.L.; Ebert, B. Profiling Cell Wall Monosaccharides and Nucleotide-Sugars from Plants. *Curr. Protoc. Plant Biol.* **2019**, *4*, e20092. [CrossRef]

48. Ito, J.; Herter, T.; Baidoo, E.E.K.; Lao, J.; Vega-Sanchez, M.E.; Smith-Moritz, A.M.; Adams, P.D.; Keasling, J.D.; Usadel, B.; Petzold, C.J.; et al. Analysis of plant nucleotide sugars by hydrophilic interaction liquid chromatography and tandem mass spectrometry. *Anal. Biochem.* **2014**, *448*, 14–22. [CrossRef]

49. Guo, M.; Yin, D.; Han, J.; Zhang, L.; Li, X.; He, D.; Du, Y.; Tang, D. Phenylboronic acid modified solid-phase extraction column: Preparation, characterization, and application to the analysis of amino acids in sepia capsule by removing the maltose. *J. Sep. Sci.* **2016**, *39*, 3428–3435. [CrossRef]

50. Xin, P.; Li, B.; Yan, J.; Chu, J. Pursuing extreme sensitivity for determination of endogenous brassinosteroids through direct fishing from plant matrices and eliminating most interferences with boronate affinity magnetic nanoparticles. *Anal. Bioanal. Chem.* **2018**, *410*, 1363–1374. [CrossRef] [PubMed]

51. Chen, L.; Wang, X.; Lu, W.; Wu, X.; Li, J. Molecular imprinting: Perspectives and applications. *Chem. Soc. Rev.* **2016**, *45*, 2137–2211. [CrossRef] [PubMed]

52. Jadda, R.; Madhumanchi, S.; Suedee, R. Novel adsorptive materials by adenosine 5′-triphosphate imprinted-polymer over the surface of polystyrene nanospheres for selective separation of adenosine 5′-triphosphate biomarker from urine. *J. Sep. Sci.* **2019**, *42*, 3662–3678. [CrossRef] [PubMed]

53. Jegourel, D.; Delepee, R.; Breton, F.; Rolland, A.; Vidal, R.; Agrofoglio, L.A. Molecularly imprinted polymer of 5-methyluridine for solid-phase extraction of pyrimidine nucleoside cancer markers in urine. *Bioorg. Med. Chem.* **2008**, *16*, 8932–8939. [CrossRef]

54. Choi, J.-H.; Abe, N.; Tanaka, H.; Fushimi, K.; Nishina, Y.; Morita, A.; Kiriiwa, Y.; Motohashi, R.; Hashizume, D.; Koshino, H.; et al. Plant-growth regulator, imidazole-4-carboxamide, produced by the fairy ring forming fungus Lepista sordida. *J. Agric. Food Chem.* **2010**, *58*, 9956–9959. [CrossRef]

55. Choi, J.-H.; Ohnishi, T.; Yamakawa, Y.; Takeda, S.; Sekiguchi, S.; Maruyama, W.; Yamashita, K.; Suzuki, T.; Morita, A.; Ikka, T.; et al. The Source of "Fairy Rings": 2-Azahypoxanthine and its Metabolite Found in a Novel Purine Metabolic Pathway in Plants. *Angew. Chem.* **2014**, *53*, 1552–1555. [CrossRef] [PubMed]

56. Choi, J.-H.; Wu, J.; Sawada, A.; Takeda, S.; Takemura, H.; Yogosawa, K.; Hirai, H.; Kondo, M.; Sugimoto, K.; Asakawa, T.; et al. N-Glucosides of Fairy Chemicals, 2-Azahypoxanthine and 2-Aza-8-oxohypoxanthine, in Rice. *Org. Lett.* **2018**, *20*, 312–314. [CrossRef] [PubMed]

57. Stevens, W.C.; Hill, D.C. General methods for flash chromatography using disposable columns. *Mol. Divers.* **2009**, *13*, 247–252. [CrossRef] [PubMed]

58. Still, W.C.; Kahn, M.; Mitra, A. Rapid chromatographic technique for preparative separations with modern resolution. *J. Org. Chem.* **1978**, *43*, 2923–2925. [CrossRef]

59. Takemura, H.; Choi, J.-H.; Matsuzaki, N.; Taniguchi, Y.; Wu, J.; Hirai, H.; Motohashi, R.; Asakawa, T.; Ikeuchi, K.; Inai, M.; et al. A Fairy Chemical, Imidazole-4-carboxamide, is Produced on a Novel Purine Metabolic Pathway in Rice. *Sci. Rep.* **2019**, *9*. [CrossRef]

60. Nordstrom, A.; Tarkowski, P.; Tarkowska, D.; Dolezal, K.; Astot, C.; Sandberg, G.; Moritz, T. Derivatization for LC electrospray ionization-MS: A tool for improving reversed-phase separation and ESI responses of bases, ribosides, and intact nucleotides. *Anal. Chem.* **2004**, *76*, 2869–2877. [CrossRef]

61. Jiang, H.-P.; Xiong, J.; Liu, F.-L.; Ma, C.-J.; Tang, X.-L.; Yuan, B.-F.; Feng, Y.-Q. Modified nucleoside triphosphates exist in mammals. *Chem. Sci.* **2018**, *9*, 4160–4167. [CrossRef] [PubMed]

62. Luo, X.-T.; Cai, B.-D.; Jiang, H.-P.; Xiao, H.-M.; Yuan, B.-F.; Feng, Y.-Q. Sensitive analysis of trehalose-6-phosphate and related sugar phosphates in plant tissues by chemical derivatization combined with hydrophilic interaction liquid chromatography-tandem mass spectrometry. *J. Chromatogr. A* **2019**, *1592*, 82–90. [CrossRef]

63. Bjorkman, P.O.; Tillberg, E. Acetylation of cytokinins and modified adenine compounds: A simple and non-destructive derivatization method for gas chromatography—Mass spectrometric analysis. *Phytochem. Anal.* **1996**, *7*, 57–68. [CrossRef]

64. Zhang, H.; Li, Y.; Li, Z.; Lam, C.W.K.; Zhu, P.; Wang, C.; Zhou, H.; Zhang, W. MTBSTFA derivatization-LC-MS/MS approach for the quantitative analysis of endogenous nucleotides in human colorectal carcinoma cells. *J. Pharm. Anal.* **2021**. [CrossRef]

65. Dudley, E.; Bond, L. Mass spectrometry analysis of nucleosides and nucleotides. *Mass Spectrom. Rev.* **2014**, *33*, 302–331. [CrossRef]

66. Salvatore, F.; Paul, R.H.; Colin, F.P.; Marja-Liisa, R. (Eds.) *Liquid Chromatography*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2017; ISBN 978-0-12-805393-5.

67. Sykora, D.; Rezanka, P.; Zaruba, K.; Kral, V. Recent advances in mixed-mode chromatographic stationary phases. *J. Sep. Sci.* **2019**, *42*, 89–129. [CrossRef] [PubMed]

68. Wilson, N.S.; Gilroy, J.; Dolan, J.W.; Snyder, L.R. Column selectivity in reversed-phase liquid chromatography—VI. Columns with embedded or end-capping polar groups. *J. Chromatogr. A* **2004**, *1026*, 91–100. [CrossRef] [PubMed]

69. Zuvela, P.; Skoczylas, M.; Liu, J.J.; Baczek, T.; Kaliszan, R.; Wong, M.W.; Buszewski, B. Column Characterization and Selection Systems in Reversed-Phase High-Performance Liquid Chromatography. *Chem. Rev.* **2019**, *119*, 3674–3729. [CrossRef]

70. Traube, F.R.; Schiffers, S.; Iwan, K.; Kellner, S.; Spada, F.; Mueller, M.; Carell, T. Isotope-dilution mass spectrometry for exact quantification of noncanonical DNA nucleosides. *Nat. Protoc.* **2019**, *14*, 283. [CrossRef] [PubMed]

71. Gustavsson, S.A.; Samskog, J.; Markides, K.E.; Langstrom, B. Studies of signal suppression in liquid chromatography-electrospray ionization mass spectrometry using volatile ion-pairing reagents. *J. Chromatogr. A* **2001**, *937*, 41–47. [CrossRef]

72. Annesley, T.M. Ion suppression in mass spectrometry. *Clin. Chem.* **2003**, *49*, 1041–1044. [CrossRef] [PubMed]

73. Lajin, B.; Goessler, W. Fluoroalkylamines: Novel, highly volatile, fast-equilibrating, and electrospray ionization-mass spectrometry signal-enhancing cationic ion-interaction reagents. *Anal. Chem.* **2020**, *92*, 10121–10128. [CrossRef]

74. Dodbiba, E.; Breitbach, Z.S.; Wanigasekara, E.; Payagala, T.; Zhang, X.; Armstrong, D.W. Detection of nucleotides in positive-mode electrospray ionization mass spectrometry using multiply-charged cationic ion-pairing reagents. *Anal. Bioanal. Chem.* **2010**, *398*, 367–376. [CrossRef] [PubMed]

75. Hsiao, J.J.; van de Bittner, G.C.; Kennedy, A.P.; Wei, T.-C. The use of HILIC zwitterionic phase superficially porous particles for metabolomics analysis. *LC GC N. Am.* **2018**, *36*, 30–35.

76. Kong, Z.; Jia, S.; Chabes, A.L.; Appelblad, P.; Lundmark, R.; Moritz, T.; Chabes, A. Simultaneous determination of ribonucleoside and deoxyribonucleoside triphosphates in biological samples by hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry. *Nucleic Acids Res.* **2018**, *46*. [CrossRef] [PubMed]

77. Huang, Y.; Li, W.; Minakova, A.; Anumol, T.; Keller, A. Quantitative analysis of changes in amino acids levels for cucumber (*Cucumis sativus*) exposed to nano copper. *Nanoimpact* **2018**, *12*, 9–17. [CrossRef]

78. Kate, P.; John, K.L. Determination of selected water-soluble vitamins (thiamine, riboflavin, nicotinamide and pyridoxine) from a food matrix using hydrophilic interaction liquid chromatography coupled with mass spectroscopy. *J. Chromatogr. B* **2021**, 122541. [CrossRef]

79. Warth, B.; Siegwart, G.; Lemmens, M.; Krska, R.; Adam, G.; Schuhmacher, R. Hydrophilic interaction liquid chromatography coupled with tandem mass spectrometry for the quantification of uridine diphosphate-glucose, uridine diphosphate-glucuronic acid, deoxynivalenol and its glucoside: In-house validation and application to wheat. *J. Chromatogr. A* **2015**, *1423*, 183–189. [CrossRef]

80. Hsiao, J.J.; Potter, O.G.; Chu, T.-W.; Yin, H. Improved LC/MS Methods for the analysis of metal-sensitive analytes using medronic acid as a mobile phase additive. *Anal. Chem.* **2018**, *90*, 9457–9464. [CrossRef]

81. Matyska, M.T.; Pesek, J.J.; Duley, J.; Zamzami, M.; Fischer, S.M. Aqueous normal phase retention of nucleotides on silica hydride-based columns: Method development strategies for analytes revelant in clinical analysis. *J. Sep. Sci.* **2010**, *33*, 930–938. [CrossRef]

82. Pesek, J.J.; Matyska, M.T.; Fischer, S.M.; Sana, T.R. Analysis of hydrophilic metabolites by high-performance liquid chromatography-mass spectrometry using a silica hydride-based stationary phase. *J. Chromatogr. A* **2008**, *1204*, 48–55. [CrossRef]

83. Pesek, J.J.; Matyska, M.T.; Loo, J.A.; Fischer, S.M.; Sana, T.R. Analysis of hydrophilic metabolites in physiological fluids by HPLC-MS using a silica hydride-based stationary phase. *J. Sep. Sci.* **2009**, *32*, 2200–2208. [CrossRef]

84. Zhang, W.; Guled, F.; Hankemeier, T.; Ramautar, R. Profiling nucleotides in low numbers of mammalian cells by sheathless CE-MS in positive ion mode: Circumventing corona discharge. *Electrophoresis* **2020**, *41*, 360–369. [CrossRef]

85. Drouin, N.; van Mever, M.; Zhang, W.; Tobolkina, E.; Ferre, S.; Servais, A.-C.; Gou, M.-J.; Nyssen, L.; Fillet, M.; Lageveen-Kammeijer, G.S.M.; et al. Capillary electrophoresis-mass spectrometry at trial by metabo-ring: Effective electrophoretic mobility for reproducible and robust compound annotation. *Anal. Chem.* **2020**, *92*, 14103–14112. [CrossRef] [PubMed]

86. Stafsnes, M.H.; Rost, L.M.; Bruheim, P. Improved phosphometabolome profiling applying isotope dilution strategy and capillary ion chromatography-tandem mass spectrometry. *J. Chromatogr. B* **2018**, *1083*, 278–283. [CrossRef] [PubMed]

87. Schwaiger, M.; Rampler, E.; Hermann, G.; Miklos, W.; Berger, W.; Koellensperger, G. Anion-exchange chromatography coupled to high-resolution mass spectrometry: A powerful tool for merging targeted and non-targeted metabolomics. *Anal. Chem.* **2017**, *89*, 7667–7674. [CrossRef]

88. Walsby-Tickle, J.; Gannon, J.; Hvinden, I.; Bardella, C.; Abboud, M.I.; Nazeer, A.; Hauton, D.; Pires, E.; Cadoux-Hudson, T.; Schofield, C.J.; et al. Anion-exchange chromatography mass spectrometry provides extensive coverage of primary metabolic pathways revealing altered metabolism in IDH1 mutant cells. *Commun. Biol.* **2020**, *3*. [CrossRef]

89. Cech, N.B.; Enke, C.G. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom. Rev.* **2001**, *20*, 362–387. [CrossRef] [PubMed]

90. Xu, X.-B.; Liu, D.-B.; Guo, X.M.; Yu, S.-J.; Yu, P. Improvement of sugar analysis sensitivity using anion-exchange chromatography-electrospray ionization mass spectrometry with sheath liquid interface. *J. Chromatogr. A* **2014**, *1366*, 65–72. [CrossRef]

91. Shi, G.; Wu, J.T.; Li, Y.; Geleziunas, R.; Gallagher, K.; Emm, T.; Olah, T.; Unger, S. Novel direct detection method for quantitative determination of intracellular nucleoside triphosphates using weak anion exchange liquid chromatography/tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1092–1099. [CrossRef]

92. Cech, N.B.; Enke, C.G. Relating electrospray ionization response to nonpolar character of small peptides. *Anal. Chem.* **2000**, *72*, 2717–2723. [CrossRef] [PubMed]

93. Bonfiglio, R.; King, R.C.; Olah, T.V.; Merkle, K. The effects of sample preparation methods on the variability of the electrospray ionization response for model drug compounds. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1175–1185. [CrossRef]

94. Matuszewski, B.K.; Constanzer, M.L.; Chavez-Eng, C.M. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS. *Anal. Chem.* **2003**, *75*, 3019–3030. [CrossRef] [PubMed]

95. Havlikova, L.; Vlckova, H.; Solich, P.; Novakova, L. HILIC UHPLC-MS/MS for fast and sensitive bioanalysis: Accounting for matrix effects in method development. *Bioanalysis* **2013**, *5*, 2345–2357. [CrossRef] [PubMed]

96. Strzelecka, D.; Chmielinski, S.; Bednarek, S.; Jemielity, J.; Kowalska, J. Analysis of mononucleotides by tandem mass spectrometry: Investigation of fragmentation pathways for phosphate- and ribose-modified nucleotide analogues. *Sci. Rep.* **2017**, *7*. [CrossRef] [PubMed]

97. Huber, C.G.; Krajete, A. Sheath liquid effects in capillary high-performance liquid chromatography-electrospray mass spectrometry of oligonucleotides. *J. Chromatogr. A* **2000**, *870*, 413–424. [CrossRef]

98. Apffel, A.; Chakel, J.A.; Fischer, S.; Lichtenwalter, K.; Hancock, W.S. Analysis of oligonucleotides by HPLC-electrospray ionization mass spectrometry. *Anal. Chem.* **1997**, *69*, 1320–1325. [CrossRef]

99. Griffey, R.H.; Greig, M.J.; Gaus, H.J.; Liu, K.; Monteith, D.; Winniman, M.; Cummins, L.L. Characterization of oligonucleotide metabolism in vivo via liquid chromatography electrospray tandem mass spectrometry with a quadrupole ion trap mass spectrometer. *J. Mass Spectrom.* **1997**, *32*, 305–313. [CrossRef]

100. Wu, Z.; Gao, W.; Phelps, M.A.; Wu, D.; Miller, D.D.; Dalton, J.T. Favorable effects of weak acids on negative-ion electrospray ionization mass spectrometry. *Anal. Chem.* **2004**, *76*, 839–847. [CrossRef]

101. Cohen, S.; Megherbi, M.; Jordheim, L.P.; Lefebvre, I.; Perigaud, C.; Dumontet, C.; Guitton, J. Simultaneous analysis of eight nucleoside triphosphates in cell lines by liquid chromatography coupled with tandem mass spectrometry. *J. Chromatogr. B* **2009**, *877*, 3831–3840. [CrossRef]

102. Guo, S.; Duan, J.; Qian, D.; Wang, H.; Tang, Y.; Qian, Y.; Wu, D.; Su, S.; Shang, E. Hydrophilic interaction ultra-high performance liquid chromatography coupled with triple quadrupole mass spectrometry for determination of nucleotides, nucleosides and nucleobases in Ziziphus plants. *J. Chromatogr. A* **2013**, *1301*, 147–155. [CrossRef]

103. Zhou, S.L.; Cook, K.D. Protonation in electrospray mass spectrometry: Wrong-way-round or right-way-round? *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 961–966. [CrossRef]

104. Amad, M.H.; Cech, N.B.; Jackson, G.S.; Enke, C.G. Importance of gas-phase proton affinities in determining the electrospray ionization response for analytes and solvents. *J. Mass Spectrom.* **2000**, *35*, 784–789. [CrossRef]

105. Green-Church, K.B.; Limbach, P.A. Mononucleotide gas-phase proton affinities as determined by the kinetic method. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 24–32. [CrossRef]

106. Birdsall, R.E.; Gilar, M.; Shion, H.; Yu, Y.Q.; Chen, W. Reduction of metal adducts in oligonucleotide mass spectra in ion-pair reversed-phase chromatography/mass spectrometry analysis. *Rapid Commun. Mass Spectrom.* **2016**, *30*, 1667–1679. [CrossRef]

107. Schug, K.; McNair, H.M. Adduct formation in electrospray ionization mass spectrometry II. Benzoic acid derivatives. *J. Chromatogr. A* **2003**, *985*, 531–539. [CrossRef]

108. Harvey, D.J. Collision-induced fragmentation of underivatized N-linked carbohydrates ionized by electrospray. *J. Mass Spectrom.* **2000**, *35*, 1178–1190. [CrossRef]

109. Ernst, M.; Silva, D.B.; Silva, R.R.; Vencio, R.Z.N.; Lopes, N.P. Mass spectrometry in plant metabolomics strategies: From analytical platforms to data acquisition and processing. *Nat. Prod. Rep.* **2014**, *31*, 784–806. [CrossRef]

110. Erngren, I.; Haglof, J.; Engskog, M.K.R.; Nestor, M.; Hedeland, M.; Arvidsson, T.; Pettersson, C. Adduct formation in electrospray ionisation-mass spectrometry with hydrophilic interaction liquid chromatography is strongly affected by the inorganic ion concentration of the samples. *J. Chromatogr. A* **2019**, *1600*, 174–182. [CrossRef] [PubMed]

111. Vessecchi, R.; Crotti, A.E.M.; Guaratini, T.; Colepicolo, P.; Galembeck, S.E.; Lopes, N.P. Radical ion generation processes of organic compounds in electrospray ionization mass spectrometry. *Mini Rev. Org. Chem.* **2007**, *4*, 75–87. [CrossRef]

112. Studzinska, S.; Siecinska, L.; Buszewski, B. On-line electrochemistry/electrospray ionization mass spectrometry (EC-ESI-MS) system for the study of nucleosides and nucleotides oxidation products. *J. Pharm. Biomed. Anal.* **2018**, *158*, 416–424. [CrossRef]

113. Weissberg, A.; Dagan, S. Interpretation of ESI(+)-MS-MS spectra—Towards the identification of "unknowns". *Int. J. Mass Spectrom.* **2011**, *299*, 158–168. [CrossRef]

114. Xu, Y.-F.; Lu, W.; Rabinowitz, J.D. Avoiding misannotation of in-source fragmentation products as cellular metabolites in liquid chromatography-mass spectrometry-based metabolomics. *Anal. Chem.* **2015**, *87*, 2273–2281. [CrossRef] [PubMed]

115. Kondrat, R.W.; McClusky, G.A.; Cooks, R.G. Multiple reaction monitoring in mass spectrometry/mass spectrometry for direct analysis of complex mixtures. *Anal. Chem.* **1978**, *50*, 2017–2021. [CrossRef]

116. Moerlein, S.; Schuster, C.; Paal, M.; Vogeser, M. Collision energy-breakdown curves—An additional tool to characterize MS/MS methods. *Clin. Mass Spectrom.* **2020**, *18*, 48–53. [CrossRef]

117. Glauser, G.; Veyrat, N.; Rochat, B.; Wolfender, J.-L.; Turlings, T.C.J. Ultra-high pressure liquid chromatography-mass spectrometry for plant metabolomics: A systematic comparison of high-resolution quadrupole-time-of-flight and single stage Orbitrap mass spectrometers. *J. Chromatogr. A* **2013**, *1292*, 151–159. [CrossRef] [PubMed]

118. Kaufmann, A. High-resolution mass spectrometry for bioanalytical applications: Is this the new gold standard? *J. Mass Spectrom.* **2020**, *55*. [CrossRef]

119. Chen, M.; Witte, C.-P. A kinase and a glycosylase catabolize pseudouridine in the peroxisome to prevent toxic pseudouridine monophosphate accumulation. *Plant Cell* **2020**, *32*, 722–739. [CrossRef]

120. Chen, M.; Urs, M.J.; Sanchez-Gonzalez, I.; Olayioye, M.A.; Herde, M.; Witte, C.-P. m(6)A RNA degradation products are catabolized by an evolutionarily conserved N-6-Methyl-AMP deaminase in plant and mammalian cells. *Plant Cell* **2018**, *30*, 1511–1522. [CrossRef] [PubMed]

121. Galperin, M.Y.; Moroz, O.V.; Wilson, K.S.; Murzin, A.G. House cleaning, a part of good housekeeping. *Mol. Microbiol.* **2006**, *59*, 5–19. [CrossRef]

122. Senkler, J.; Rugen, N.; Eubel, H.; Hegermann, J.; Braun, H.-P. Absence of complex I implicates rearrangement of the respiratory chain in european mistletoe. *Curr. Biol.* **2018**, *28*, 1606. [CrossRef] [PubMed]

123. Schwaiger, M.; Schoeny, H.; El Abiead, Y.; Hermann, G.; Rampler, E.; Koellensperger, G. Merging metabolomics and lipidomics into one analytical run. *Analyst* **2019**, *144*, 220–229. [CrossRef] [PubMed]

124. Voon, C.P.; Guan, X.; Sun, Y.; Sahu, A.; Chan, M.N.; Gardestrom, P.; Wagner, S.; Fuchs, P.; Nietzel, T.; Versaw, W.K.; et al. ATP compartmentation in plastids and cytosol of Arabidopsis thaliana revealed by fluorescent protein sensing. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E10778–E10787. [CrossRef] [PubMed]

125. Steinbeck, J.; Fuchs, P.; Negroni, Y.L.; Elsaesser, M.; Lichtenauer, S.; Stockdreher, Y.; Feitosa-Araujo, E.; Kroll, J.B.; Niemeier, J.-O.; Humberg, C.; et al. In Vivo NADH/NAD⁺ Biosensing Reveals the Dynamics of Cytosolic Redox Metabolism in Plants. *Plant Cell* **2020**, *32*, 3324–3345. [CrossRef] [PubMed]

126. Arrivault, S.; Guenther, M.; Florian, A.; Encke, B.; Feil, R.; Vosloh, D.; Lunn, J.E.; Sulpice, R.; Fernie, A.R.; Stitt, M. Dissecting the subcellular compartmentation of proteins and metabolites in arabidopsis leaves using non-aqueous fractionation. *Mol. Cell. Proteom.* **2014**, *13*, 2246–2259. [CrossRef] [PubMed]

127. Fuertauer, L.; Kuestner, L.; Weckwerth, W.; Heyer, A.G.; Naegele, T. Resolving subcellular plant metabolism. *Plant J.* **2019**, *100*, 438–455. [CrossRef]

128. Niehaus, M.; Straube, H.; Kuenzler, P.; Rugen, N.; Hegermann, J.; Giavalisco, P.; Eubel, H.; Witte, C.-P.; Herde, M. Rapid affinity purification of tagged plant mitochondria (Mito-AP) for metabolome and proteome analyses. *Plant Physiol.* **2020**, *182*, 1194–1210. [CrossRef]

2.4 Rapid affinity purification of tagged plant mitochondria (Mito-AP) for metabolome and proteome analyses

Markus Niehaus[1], Henryk Straube[1], Patrick Künzler[2], Nils Rugen[2], Jan Hegermann[3], Patrick Giavalisco[4], Holger Eubel[2], Claus-Peter Witte[1] and Marco Herde[1]

[1] Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany

[2] Institute of Plant Genetics, 30419 Hannover, Leibniz Universität Hannover, Germany

[3] Research Core Unit Electron Microscopy, Hannover Medical School (MHH), 30625 Hannover, Germany

[4] Max-Planck-Institute for Biology of Aging, 50931 Köln, Germany

Breakthrough Technologies

Rapid Affinity Purification of Tagged Plant Mitochondria (Mito-AP) for Metabolome and Proteome Analyses[1]

Markus Niehaus,[a] Henryk Straube,[a] Patrick Künzler,[b] Nils Rugen,[b] Jan Hegermann,[c] Patrick Giavalisco,[d] Holger Eubel,[b] Claus-Peter Witte,[a] and Marco Herde[a,2,3]

[a]Leibniz Universität Hannover, Department of Molecular Nutrition and Biochemistry of Plants, 30419 Hannover, Germany
[b]Leibniz Universität Hannover, Institute of Plant Genetics, 30419 Hannover, Germany
[c]Research Core Unit Electron Microscopy, Hannover Medical School (MHH), 30625 Hannover, Germany
[d]Max-Planck-Institute for Biology of Aging, 50931 Köln, Germany

ORCID IDs: 0000-0003-3057-7823 (M.N.); 0000-0001-9286-7784 (H.S.); 0000-0002-9686-5126 (P.K.); 0000-0002-9297-4560 (N.R.); 0000-0002-4636-1827 (P.G.); 0000-0001-7065-178X (H.E.); 0000-0002-3617-7807 (C.-P.W.); 0000-0003-2804-0613 (M.H.).

The isolation of organelles facilitates the focused analysis of subcellular protein and metabolite pools. Here we present a technique for the affinity purification of plant mitochondria (Mito-AP). The stable ectopic expression of a mitochondrial outer membrane protein fused to a GFP:Strep tag in Arabidopsis (*Arabidopsis thaliana*) exclusively decorates mitochondria, enabling their selective affinity purification using magnetic beads coated with Strep-Tactin. With Mito-AP, intact mitochondria from 0.5 g plant material were highly enriched in 30–60 min, considerably faster than with conventional gradient centrifugation. Combining gradient centrifugation and Mito-AP techniques resulted in high purity of >90% mitochondrial proteins in the lysate. Mito-AP supports mitochondrial proteome analysis by shotgun proteomics. The relative abundances of proteins from distinct mitochondrial isolation methods were correlated. A cluster of 619 proteins was consistently enriched by all methods. Among these were several proteins that lack subcellular localization data or that are currently assigned to other compartments. Mito-AP is also compatible with mitochondrial metabolome analysis by triple-quadrupole and orbitrap mass spectrometry. Mito-AP preparations showed a strong enrichment with typical mitochondrial lipids like cardiolipins and demonstrated the presence of several ubiquinones in Arabidopsis mitochondria. Affinity purification of organelles is a powerful tool for reaching higher spatial and temporal resolution for the analysis of metabolomic and proteomic dynamics within subcellular compartments. Mito-AP is small scale, rapid, economic, and potentially applicable to any organelle or to organelle subpopulations.

One challenge in science is to shift the scale in which observations can be made. From tissues via cells to compartments within a cell (e.g. organelles), the scale of scientific observation becomes progressively smaller, and the resolution must increase drastically to allow ever deeper insights into the details of biological processes.

Mitochondria are essential for cellular metabolism, because they are the major ATP-exporting organelles in plants, harbor central biochemical pathways such as the citrate cycle and the respiratory chain, represent a central hub of amino acid metabolism, and support photosynthesis and photorespiration. Objectives in plant mitochondria research are inter alia to create an inventory list of proteins and metabolites residing in these organelles as well as to investigate the dynamics of these molecules, for example upon exposure to different treatments and stresses. To address these objectives, the enrichment of mitochondria relative to other components of the cell is required. The reduction of complexity concomitant with purification also fosters the detection of proteins and metabolites with low abundances, because background and interfering matrix are removed prior to analysis.

To accomplish the isolation of mitochondria from leaf material, centrifugation-based methods have been developed that achieve a high degree of purity (Day et al., 1985; Keech et al., 2005). Nevertheless, a full overview of all mitochondrial proteins is still needed. To generate an inventory of proteins from a particular compartment, proteomic profiling of samples differentially enriched in this compartment was shown to

be a powerful technique (Eubel et al., 2007; Kraner et al., 2017). An even greater challenge is the analysis of subcellular metabolite pools, for example the mitochondrial pools, in particular if the same metabolites are present in several compartments, maybe even in grossly different amounts. Highly purified organelles (mitochondria) are needed to ensure that detected metabolites are truly derived from this organelle and not from other copurified compartments.

To capture the mitochondria as close as possible to their in vivo state, the swiftness of the isolation procedure is even more important than the purity. After removing the mitochondria from their cellular context during cell rupture, posttranslational protein modifications are expected to change quickly within the organelle. Steady-state concentrations of metabolites will also be altered by metabolite conversion, degradation, or leakage from the mitochondria (Roberts et al., 1997; Agius et al., 2001).

Currently, the state-of-the-art technology for large scale mitochondrial isolation is a combination of differential and density gradient centrifugation (DGC) of extracts from cell or callus cultures (Klein et al., 1998) or green leaves (Day et al., 1985; Keech et al., 2005). Although this technique provides comparatively high amounts of mitochondria with reasonably high purity, the procedure is laborious and time-consuming, and it requires a large amount of starting material. A centrifugation-based method on a smaller scale is less time consuming, but yield and purity of the obtained mitochondrial preparation were not reported (Boutry et al., 1984). Following isolation by DGC, mitochondria can be further purified by free-flow electrophoresis using the surface charge on the organelles as an alternative handle for differential isolation (Eubel et al., 2007). Additionally, a fractionation technique using rapid filtration has been employed to swiftly separate plant organelles, including mitochondria, for metabolite analysis, thus minimizing leakage or conversion of metabolites (Lilley et al., 1982; Gardeström and Wigge, 1988). This method requires specialized equipment and the preparation of protoplasts prior to isolation, which might have an undesired impact on the cells. Nonaqueous fractionation is another method for the separation of subcellular metabolite pools and proteins (Gerhardt and Heldt, 1984; Arrivault et al., 2014). This technique has the advantage of rapidly quenching biological activities in the material prior to fractionation, allowing the assessment of metabolites and proteins close to their native state. Whether nonaqueous fractionation can reproducibly separate cytosolic from mitochondrial proteins and metabolites is still under debate (Arrivault et al., 2014; Fürtauer et al., 2019).

Alternative approaches for mitochondria isolation have been developed more recently that do not rely on the general physicochemical characteristics of the organelle, but rather use affinity purification techniques targeting specific proteins present on the mitochondrial surface. An antibody coupled to magnetic beads directed against TOM22, a mitochondrial import receptor

subunit in the outer membrane, allowed the isolation of native mitochondria from mammalian cells (Hornig-Do et al., 2009; Afanasyeva et al., 2018). Recently, an innovative approach employing a recombinant protein for the isolation of mitochondria from mammalian cells with affinity purification was established and used to quantify metabolites and proteins (Chen et al., 2016). The key feature of this technology is the heterologous expression of a chimeric protein with a C-terminal domain integrated into the outer mitochondrial membrane and an N-terminal cytosolic domain consisting of a GFP for microscopic localization and a hemagglutinin-tag (HA) for immunoprecipitation. Employing magnetic beads coupled to anti-HA antibodies, mitochondria decorated with this fusion protein can be affinity purified.

Here, we established an affinity purification scheme for mitochondria from Arabidopsis (*Arabidopsis thaliana*), which we call Mito-AP. Plants pose several unique challenges for Mito-AP, because they possess chloroplasts, which need to be selected against during purification, and because plant extracts are a difficult matrix for any purification method. Mito-AP yields highly enriched mitochondria, requires only a fraction of the starting material typically used for conventional mitochondrial isolation, is scalable, and can be performed considerably faster than the classical centrifugation-based methods. The protocol described here introduces several innovations, which substantially lower the cost of this technique and allow the nondestructive removal of the mitochondria from the magnetic beads.

RESULTS

The Locus At1g55450 Encodes a Suitable Adapter Protein for Mito-AP

The expression of a recombinant chimeric protein on the mitochondrial surface enabling affinity purification is a key feature of Mito-AP. The C-terminal domain of such a protein should be integrated into the outer mitochondrial membrane and the N-terminal domain, including a GFP and an affinity tag, should face the cytosol as a handle for the affinity purification employing magnetic beads (Fig. 1). In contrast to Chen et al. (2016), who used a similar approach for HA-tagged mammalian mitochondria employing a magnetic affinity resin decorated with anti-HA antibodies, we aimed to use the Strep-tag/Strep-Tactin interaction to reduce procedure costs and to potentially allow for mild affinity elution. The Twin-Strep tag (IBA Lifesciences) was chosen, because it supposedly binds more strongly to Strep-Tactin than the single Strep tag.

We did not employ the chimeric protein described by Chen et al. (2016), because plant cells have an elaborate organelle targeting system preventing the misimport of mitochondrial proteins into the chloroplast and vice versa. It therefore cannot be excluded that the mammalian adapter protein may be partially mistargeted
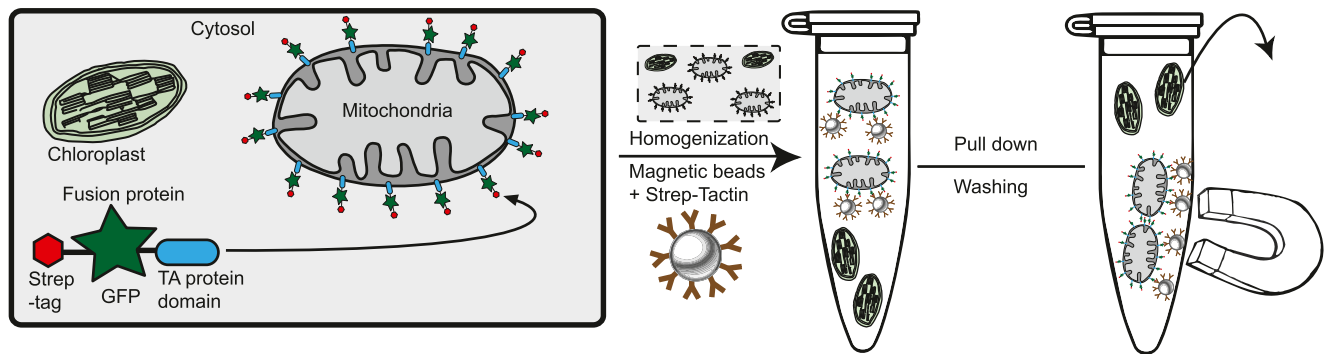
Figure 1. Workflow of Mito-AP. An ectopically expressed protein, which integrates into the outer mitochondrial membrane, fused with an N-terminal Strep-tag and a GFP-tag facing the cytosol can be used as an anchor for affinity purification of mitochondria with Strep-Tactin-coated magnetic beads.

to chloroplasts. In fact, plants ectopically producing mammalian mitochondrial proteins sometimes erroneously direct these proteins into chloroplasts (Maggio et al., 2007). We performed a literature survey to identify plant proteins that are specifically integrated with their C-terminal domain in the outer mitochondrial membrane and feature an N-terminal domain facing the cytosol. Three candidate proteins belonging to the class of tail-anchored proteins were identified: (1) Cb5-D described by Maggio et al. (2007); (2) a putative methyltransferase encoded at locus At1g55450 and described by Marty et al. (2014); and (3) TOM22-V encoded at locus At5g43970.

In the case of Cb5-D, the C-terminal sequence conferring mitochondrial localization has been well defined (Hwang et al., 2004). Therefore, this domain was tested, whereas for the other candidates, full-length proteins were used for the construction of fusion proteins. The expression of the constructs and the subcellular location of the three resulting N-terminal Twin-Strep-tag-GFP fusion variants were assessed in transient assays performed in *Nicotiana benthamiana*. Despite the use of a strong 35S promoter for the respective constructs, Cb5-D and TOM22-V expression were barely detectable by confocal microscopy. By contrast, a clear and strong signal consistent with a mitochondrial location was observed for the protein belonging to the S-adenosyl-L-Met-dependent methyltransferase superfamily encoded at At1g55450. We selected this protein, herein referred to as ADAPTER, for our approach (Supplemental Fig. S1).

A homozygous transgenic Arabidopsis line with intermediate expression (line 11) of the 35S:Twin-Strep-tag:GFP:ADAPTER construct was generated. Confocal microscopy analysis suggested that the mitochondrial membrane was exclusively decorated by this chimeric protein, whereas signals from other organelles, in particular from chloroplasts, were not observed (Fig. 2A). To verify this localization, we created Arabidopsis plants expressing the 35S:Twin-Strep-tag:GFP:ADAPTER construct as well as a construct for mCherry directed to the mitochondria by the *Sc*CoxIV peptide fused to the

N terminus (Maarse et al., 1984). In these plants (line 1), GFP fluorescence from the ADAPTER:GFP fusion surrounded the mCherry signal located in the mitochondrial matrix (Fig. 2B).

Continuous GFP fluorescence from the whole rosette was observed in the Twin-Strep-tag:GFP:ADAPTER line 11, which was selected for all further experiments (Fig. 3A). The highly uniform expression in this line indicates that gene silencing did not occur, which would have resulted in individuals with patchy fluorescence. Silencing was probably prevented by the use of a GFP gene with an intron and by selecting a line with only moderate fluorescence.

Ectopic production of the ADAPTER fusion protein did not change the normal physiological appearance of the plant at any stage during development (Fig. 3, B–E). A recent study demonstrated that a short photoperiod enhances growth phenotypes caused by mutations in a mitochondrial protein (Pétriacq et al., 2017). Therefore, we asked the question whether a potential impact of the ADAPTER fusion protein on mitochondrial function is revealed under short-day conditions. Leaf surface and weight did not differ significantly between the wild-type and the Twin-Strep-tag:GFP:ADAPTER plants (Supplemental Fig. S2, A and B), suggesting that ectopic production of the ADAPTER fusion protein does not interfere with mitochondrial function. Additionally, we did not observe any effect on the root length of seedlings and the germination rate (Supplemental Fig. S2, C and D). Therefore, mitochondria isolated from the transgenic lines are likely representative of mitochondria from the wild type.

Outer Mitochondrial Membranes Can Be Isolated Together with the Mitoplast by Mito-AP

Stable homozygous 35S:Twin-Strep-tag:GFP:ADAPTER plants also possessing a transgene expressing the *Sc*CoxIV:mCherry protein (line 1) were used for affinity purification of mitochondria with commercially available magnetic beads carrying the Strep-Tactin
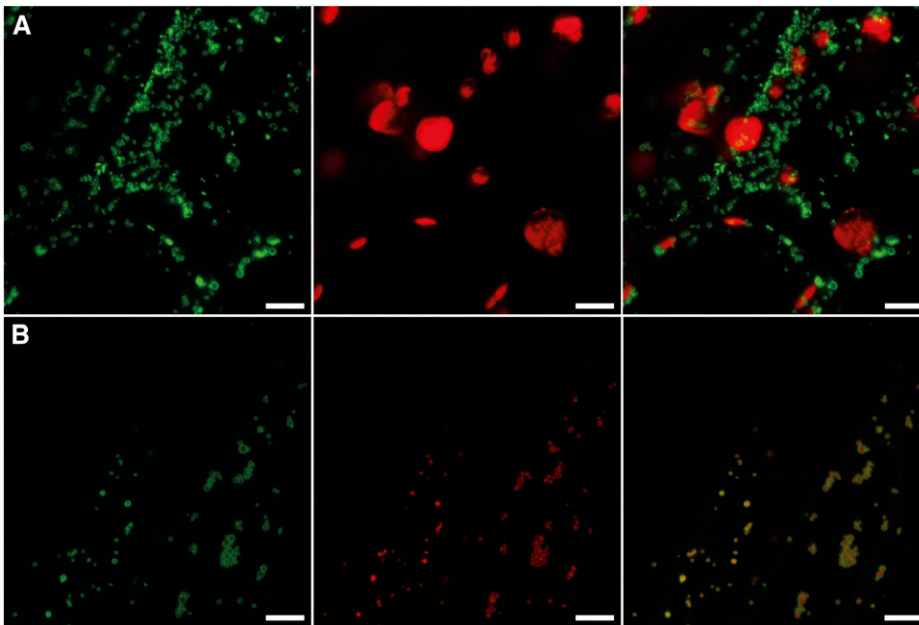
Figure 2. Subcellular localization of the Twin-Strep-tag:GFP:ADAPTER protein in stable transformed Arabidopsis. A, Confocal fluorescence microscopy images of the vascular tissue in leaves expressing the 35S:Twin-Strep-tag:GFP:ADAPTER construct. Shown are GFP (left) and autofluorescence (middle) of chloroplasts, and an overlay of both channels (right). B, Images from leaf vascular tissue coexpressing 35S:Twin-Strep-tag:GFP:ADAPTER (GFP, left) and 35S:*Sc*CoxIV:mCherry (mCherry, middle), and an overlay of both channels (right). Scale bars = 8 μm.

protein (IBA Lifesciences) on their surface. The beads were washed and then analyzed by confocal microscopy. Circular fluorescent structures were observed (Fig. 4) that were not detected on control beads incubated with extracts from wild-type plants. The fluorescence had the spectral emission characteristics of mCherry, suggesting that (1) mitochondria from total cell extracts had been isolated; and (2) these mitochondria maintained their integrity, still confining the mCherry protein to the mitochondrial matrix (Fig. 4). This analysis does not provide a quantitative assessment of mitochondrial integrity, and the presence of outer membranes without a mitoplast or otherwise damaged mitochondria cannot be ruled out (for a more

quantitative assessment of intact mitochondria, see the last paragraph of the "Results" section).

For extracting and washing the beads, an ammonium bicarbonate/sodium chloride buffer previously suggested to be advantageous for downstream liquid chromatography-mass spectrometry (LC-MS) applications with mitochondria isolated from mammalian cells was used (Chen et al., 2016). To ensure that this buffer does not have a negative impact on mitochondrial integrity, we compared it to the mannitol buffer employed for the isolation of plant mitochondria by DGC (Klein et al., 1998). As a control, mitochondria were also exposed to pure water, resulting in hypoosmotic shock and mitochondrial rupture.
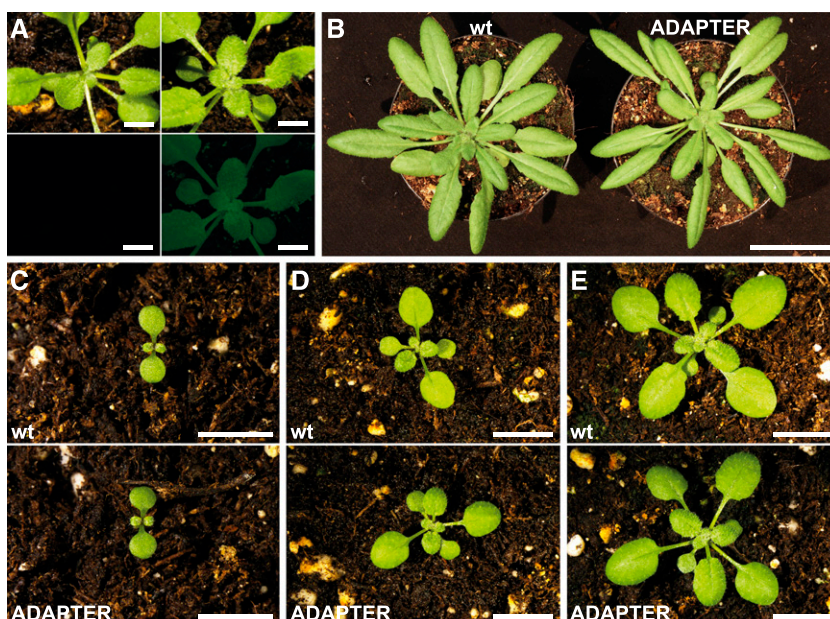

Figure 3. Comparison of growth and GFP fluorescence between the wild type (Col-0) and a stable transformed plant carrying the 35S:Twin-Strep-tag:GFP:ADAPTER construct. A, Photographs (top) and GFP fluorescence images (bottom) of wild-type (left) and transgenic plants (right) 21 d after imbibition (dai). Scale bars = 0.5 cm. B, Growth comparison of wild-type (wt) and transgenic (ADAPTER) plants at 35 dai (scale bar = 3.5 cm) and during development at 9 dai (C), 16 dai (D), and 21 dai (E). Pictures are representative of three plants within each experimental group. Scale bars = 1 cm (C–E).

Figure 4. Colocalization of markers for mitochondrial outer membrane and matrix after affinity purification. Representative confocal fluorescence microscopy images of magnetic bead clusters with bound mitochondria originating from plants coexpressing 35S:Twin-Strep-tag:GFP:ADAPTER (GFP, left) and 35S:*Sc*CoxIV:mCherry (mCherry, middle), and an overlay of both channels (right). Scale bars = 10 μm.
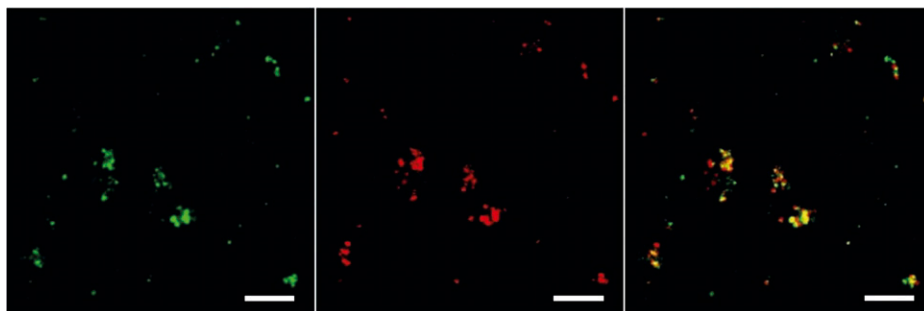
Mitochondria were isolated using DGC, pelleted by centrifugation, and resuspended in either mannitol (original buffer from the DGC protocol), potassium chloride/bicarbonate, sodium chloride/bicarbonate, or water, and then incubated for 20 min on ice. The preparations were centrifuged, and both the pellet and the supernatant were analyzed by SDS gel electrophoresis. A high amount of protein in the supernatant was only observed from mitochondrial preparations that had been incubated in water (Supplemental Fig. S3), suggesting that mitochondrial integrity was compromised under these conditions. By contrast, all other mitochondrial treatments resulted in only minor amounts of protein in the supernatants. Since there was no difference between mitochondrial preparations exposed to mannitol or sodium chloride/bicarbonate buffer, we concluded that the sodium chloride/bicarbonate buffer is suitable for isolation of plant mitochondria.

The Surface Area of the Isolation Matrix Rather Than Its Loading with Strep-Tactin Limits the Yield of Mitochondria

Initial experiments were conducted with commercially available magnetic beads of 1-μm diameter decorated with Strep-Tactin (IBA Lifesciences), which were developed for the isolation of entire cells. Although the 1-μm beads are suitable for mitochondrial isolation in principle (Supplemental Fig. S4), it might be advantageous to use smaller beads, which have been suggested to increase the yield of mitochondria (Chen et al., 2016). However, Strep-Tactin-coated nanoparticles are not commercially available. Therefore, we generated Strep-Tactin-coated iron oxide particles in house by chemically linking Strep-Tactin (IBA Lifesciences) to iron oxide beads (Chemicell). When testing the separation of magnetic beads of different sizes from the liquid phase with an external magnet, we observed that beads with a 100-nm diameter were more difficult to separate than larger beads with a 200-nm diameter.

Magnetic beads are available either coated with starch, whose hydroxyl groups can be activated with cyanogen bromide, or coated with a sugar polymer possessing carboxyl groups, which can be activated with carbodiimide. Both activated groups react with the amino groups of Strep-Tactin, allowing coupling of the protein to the matrix. We performed both coupling regimes and compared the performance of the resulting matrices. With 200-nm starch-coated beads, a higher purity of mitochondria was obtained than with the other matrices. Purity was judged by the relative enrichment of proteins predicted to be located in mitochondria (SUBA4 database; Hooper et al., 2017). Proteins were identified and quantified using shotgun proteomics, where protein abundance was determined by a label-free quantification algorithm (Schwanhäusser et al., 2011) providing intensity-based absolute quantification (iBAQ) values (Supplemental Fig. S5). Compared to the commercial 1-μm beads, the 200-nm starch-coated beads performed only slightly better in terms of enrichment of mitochondrial proteins (Supplemental Fig. S5). However, with equal bead volumes used in the purification, a 3-fold higher protein yield in the samples derived from the 200-nm material was obtained, thus confirming the initial hypothesis that higher yields can be obtained with smaller beads (Supplemental Fig. S6).

We also quantified the amount of Strep-Tactin coupled to the surface of equal volumes of the commercial 1-μm and the in-house-made 200-nm matrices and observed ~20-fold more Strep-Tactin on the 1-μm beads (Supplemental Fig. S7). It appears that for isolation of mitochondria, the 5-fold greater surface area of the 200-nm beads is more important for the yield than is the absolute number of Strep-Tactin moieties on the surface.

Mitochondria Isolated by Affinity Purification Are Strongly Enriched

Next, we compared mitochondria preparations derived from Mito-AP to preparations made using DGC, currently accepted as the gold standard for mitochondria isolation from plants. Both methods were performed with 5-week-old 35S:Twin-Strep-tag:GFP:ADAPTER plants grown under short-day conditions, and both preparations were repeated on three different days. We used 220-fold more plant material for DGC than for Mito-AP (110 g versus 0.5 g; Supplemental Fig. S8, A and B). A similar or even slightly better relative protein yield was obtained with Mito-AP

(Supplemental Fig. S8C), but far less total protein was purified due to the lower amount of starting material used.

For the Mito-AP protocol, mitochondria bound to the magnetic beads were further processed in two different ways. In method 1, the beads were boiled in SDS loading buffer for elution of the proteins. In method 2, an affinity elution step with biotin was employed to potentially increase purity, because unspecific proteins bound to the bead surface will not be eluted and can be removed together with the affinity matrix by centrifugation. However, we observed that elution with biotin was not fully quantitative (Supplemental Fig. S9). Therefore, we supported the biotin elution by adding proteinase K to the elution buffer, subjecting all proteins outside of the mitochondria to proteolytic degradation, including the Twin-Strep-tag:GFP:ADAPTER fusion protein connecting the mitochondria to the beads as well as any remaining unspecific protein contaminants. We hypothesized that this step might improve elution and increase the enrichment of mitochondrial proteins in the preparation, but that proteins residing in the outer mitochondrial membrane might be lost.

As expected, most proteins in the crude extracts originated from plastids. The abundance of mitochondrial proteins was strongly increased by DGC as well as by Mito-AP (Fig. 5; Supplemental Table S1). Consistent with our expectation, mitochondrial proteins were not as highly enriched by method 1 as by method 2, which led to >70% of iBAQs originating from mitochondrial proteins. The ADAPTER protein, known to reside on the mitochondrial surface, was only detected in the sample derived from method 1 but not method 2,

confirming the efficiency of proteinase K digestion. The addition of proteinase K improved the yield 1.7-fold (Supplemental Fig. S9). One can envisage that such a yield increase could also be achieved by upscaling the method without the proteinase K treatment, resulting in improved yield and purity. Most interestingly, mitochondrial proteins were enriched to a higher degree by Mito-AP (method 2) than by DGC (Fig. 5, A and B).

However, we had obtained a higher level of purity with DGC in the past (Klodmann et al., 2011; Senkler et al., 2017; Rugen et al., 2019). Also, the variation in mitochondrial protein enrichment across the replicates was high at ~8.5%, whereas the variation for Mito-AP replicates with both methods was considerably lower (3.3% for method 1 and 2.5% for method 2; Supplemental Fig. S10). Reasons for this are currently unknown. Most likely, the expression of the fusion protein has some impact on the sedimentation behavior of the mitochondria during centrifugation. Since the extraction of the mitochondrial band from the gradient is a manual process and subject to a certain degree of variation in terms of volume, position, and disturbance of the gradient, mitochondrial purity may suffer when the mitochondria of the transgenic line are extracted in the same way as previously performed for the wild type. Thus, minor adjustments to the DGC procedure are expected to yield organelle purities in the transgenic line that are comparable to those reported for wild-type mitochondria.

In a tandem purification experiment, we assessed whether purity might be further increased when mitochondria isolated via DGC are used as input for Mito-AP enrichment. With this approach, purities of >90%
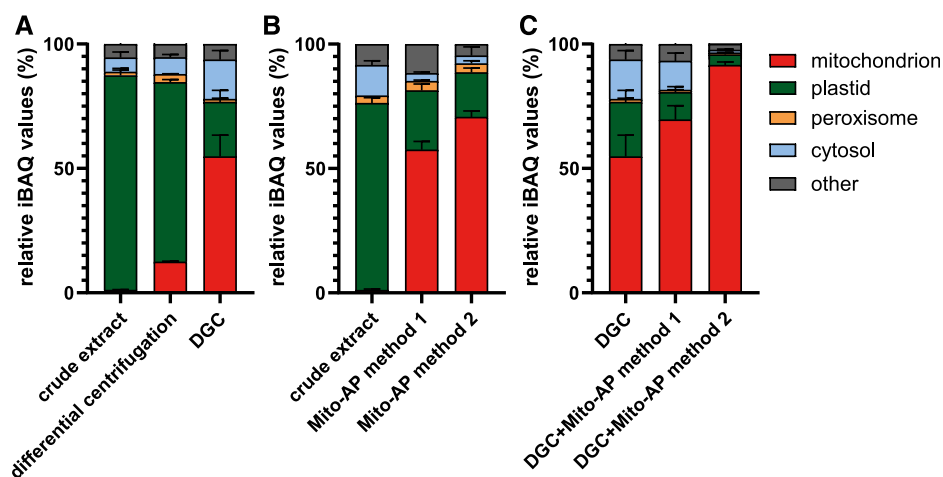

Figure 5. Enrichment of mitochondria by different purification methods. Plants were grown under short-day conditions for 4 weeks and subjected to either differential centrifugation, DGC, Mito-AP, or a combination of DGC and Mito-AP (DGC + Mito-AP). iBAQ values for every identified protein were assigned to their respective categories (mitochondrion, plastid, peroxisome, cytosol, or other), referred to total iBAQs of the corresponding sample, and plotted as percentages. The categories were defined by the SUBAcon annotation provided by SUBA4. Categorized iBAQ values are shown from crude extract, samples after only differential centrifugation, and samples after DGC (A); from crude extracts and samples after Mito-AP (method 1, boil-off; method 2, elution with biotin and proteinase K; B); and from samples after DGC and after tandem purification with DGC followed by the two different Mito-AP procedures (C). DGC data in A and C are the same. Error bars represent the sd ($n = 3$ biological replicates).

were achieved, suggesting that each method removes a different spectrum of contaminants (Fig. 5C).

Mito-AP can be performed considerably faster than DGC, requiring only about 30 min for method 1 and 60 min for method 2. The direct comparison of both methods shows that Mito-AP results in better purity than DGC and appears to be more reproducible. Mito-AP requires only comparatively small amounts of plant material and is quite economic when a Strep-Tactin affinity matrix is used. One Mito-AP employing 6.25 mg of Strep-Tactin-coated 200-nm beads and using 500 mg of Arabidopsis leaf material yields ~10 μg of mitochondrial protein and costs about 10 euros in consumables. The Strep-Tactin matrix is ~5-fold less expensive than comparable commercial beads coated with anti-HA antibody, which were employed by Chen et al. (2016) for the affinity purification of mammalian mitochondria.

Prediction of Mitochondrial Protein Localization by Correlation Analysis

Although many studies have contributed to a comprehensive inventory of the mitochondrial proteome, it is an ongoing challenge to compile a complete set of mitochondrial proteins. We hypothesize that this challenge can be addressed by monitoring and correlating the relative abundances of proteins in samples from distinct experimental approaches for enriching mitochondria. This study offered an opportunity to test this idea, because several different strategies were employed for generating samples enriched in mitochondria. In such samples, the relative abundances of proteins known to be located in mitochondria will strongly correlate, because their abundance directly reflects the ratio of mitochondria to other cellular components. By contrast, the relative abundance of a contaminant protein will depend on whether a particular enrichment strategy is suitable for its removal, and therefore, the relative abundance of contaminants may correlate to each other but not to proteins truly localized in mitochondria. Thus, correlation can be used as a tool to predict mitochondrial localization. Proteins that cluster with known mitochondrial proteins in such an analysis are likely to reside in mitochondria as well.

Because the combination of DGC and Mito-AP resulted in a higher mitochondrial purity (Fig. 5C) than any of these techniques alone, each method appears to remove different contaminants. Also, the two different versions of Mito-AP have distinct protein enrichment profiles. To increase the coverage of both mitochondrial and contaminating proteins, we employed a longer (4-h) liquid chromatography separation prior to shotgun MS analysis for all three biological replicates of (1) DGC alone, (2) Mito-AP method 1, (3) Mito-AP method 2, (4) DGC combined with Mito-AP method 1, and (5) DGC combined with Mito-AP method 2.
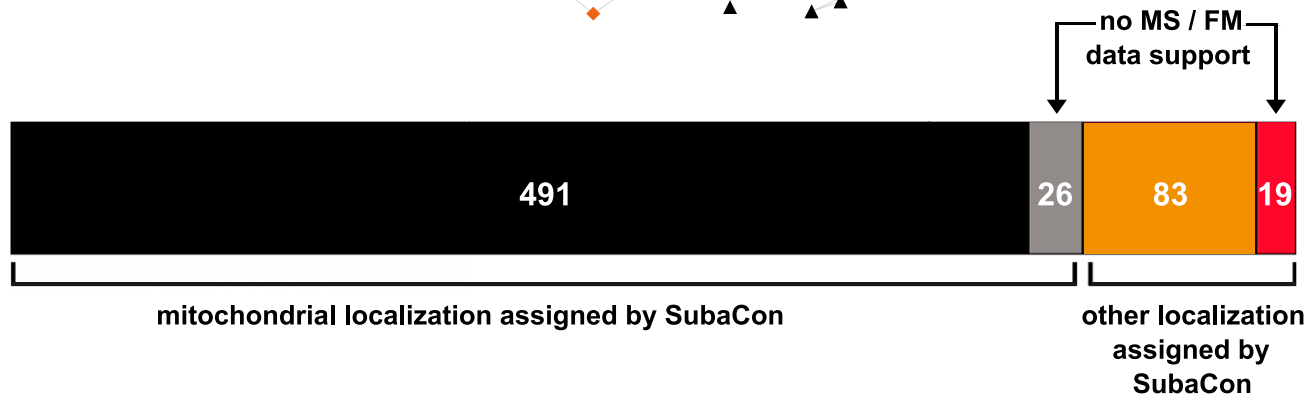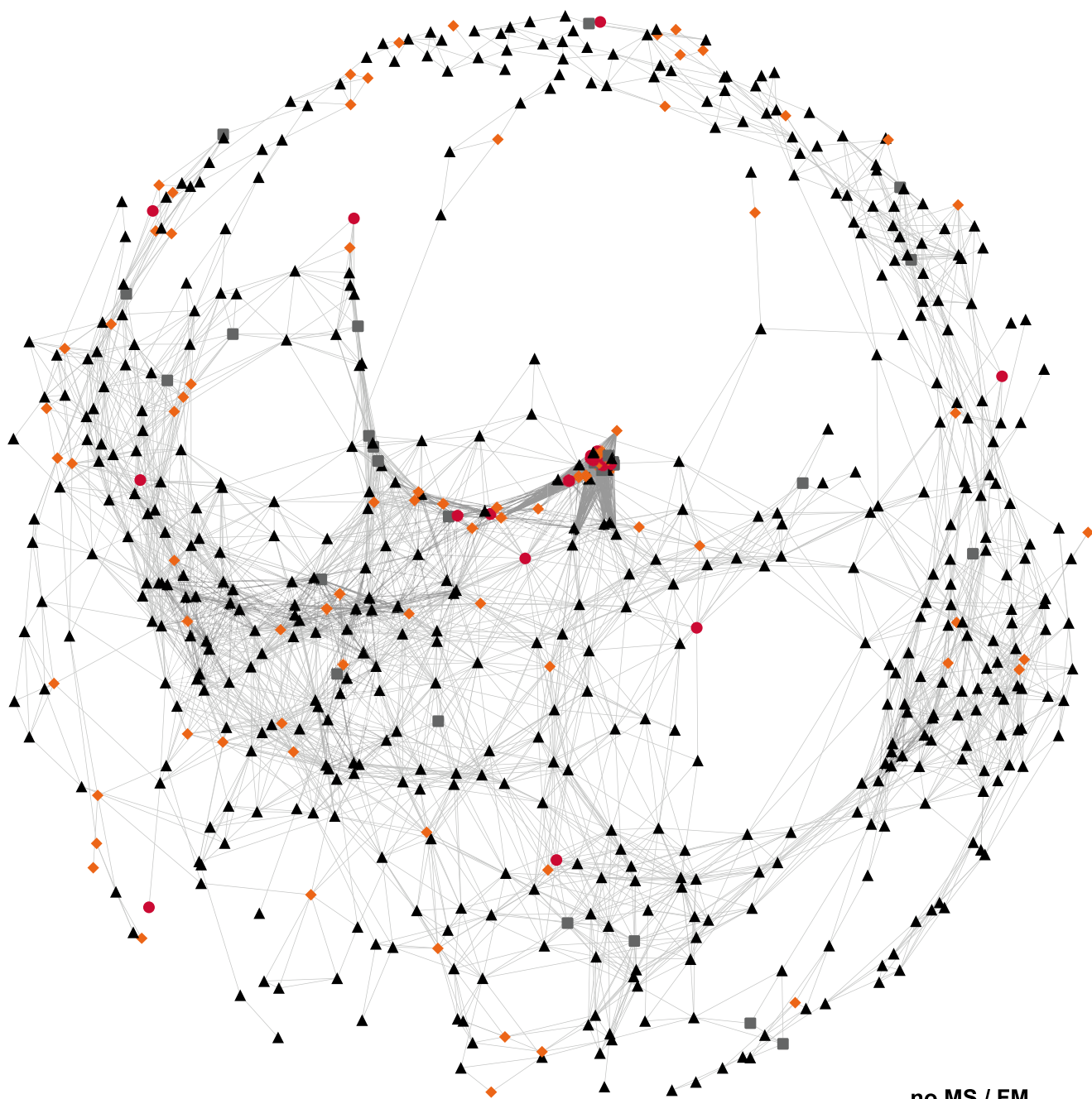
For the individual replicates of each method, the iBAQ value of each protein was normalized to the total number of iBAQs detected in the respective sample. For each of the five methods, a relative mean iBAQ value for each protein was calculated from the three replicates. With this data, a correlation matrix using Pearson's correlation coefficient was built. All correlations above a cutoff of 0.99 were displayed as a weighted network placing proteins with stronger correlation closer together.

The correlation analysis revealed one cluster containing 619 proteins that is highly enriched in proteins predicted by SUBAcon (Hooper et al., 2017) to be localized in mitochondria (517 of 619 [82%]; Fig. 6; Supplemental Table S2). Of these 517 proteins assigned to mitochondria, 491 (shown as black triangles in the cluster) have available supporting MS or fluorescence microscopy data. For the remaining 26 proteins, experimental evidence of their mitochondrial localization is provided here (Fig. 6, gray squares). For the other 102 proteins in the above-mentioned cluster of 619 proteins (18%), mitochondrial localization is not predicted according to SUBAcon. For 19 of these, no experimental data are available (Fig. 6, red circles). Therefore, this study provides evidence for mitochondrial location of these proteins despite a different SUBAcon prediction. For 83 of the 102 proteins not assigned to mitochondria by SUBAcon, either MS or fluorescent microscopy data are available (Fig. 6, orange diamonds), and 62 of these proteins have at least once been found in mitochondria by either technique. The remaining 21 proteins are not predicted by SUBAcon to be in mitochondria, nor is there current experimental evidence indicating their presence in mitochondria. Further analysis of these 21 proteins shows that only five were demonstrated by fluorescence microscopy to reside in another compartment. Although we cannot completely resolve this discrepancy, it is possible that in some instances protein variants arising from alternative transcripts might be differentially located. This might also be one possible explanation for the remaining 16 proteins, for which localization information is based on MS analysis of isolated cellular compartments. It is worth noting that for half of these proteins, the experimental evidence is limited to a single MS study.

In summary, in the mitochondrial cluster presented here, we have identified previously uncharacterized candidate proteins that exhibit mitochondrial localization as a starting point for further analysis and experimental assessment.

Mitochondrial Metabolites Can Be Detected and Quantified in Affinity-Purified Mitochondria

Our initial motivation to develop a purification protocol for mitochondria was to quantify nucleotide subpools in these organelles. Therefore, we tested whether nucleotides can be detected when mitochondria are enriched using method 1 and then directly

101

no MS / FM
data support

| 491 | 26 | 83 | 19 |

mitochondrial localization assigned by SubaCon

other localization
assigned by
SubaCon

extracted on the beads. Unfortunately, metabolites could hardly be detected in these extracts. The addition of labeled nucleotide standards suggested that metabolites interact with the magnetic beads, preventing their detection in the mass spectrometer. We reasoned that the removal of the magnetic beads before metabolite extraction might alleviate the problem. To this end, method 2 was developed, in which the mitochondria are eluted and the affinity matrix is removed prior to extraction. Extracts from method 2 were analyzed for NAD, NADH, AMP, ADP, ATP, NADP, and nicotinamide mononucleotide (NMN). For all these metabolites, robust signals were recorded that were several orders of magnitude stronger than signals from the negative control, where beads had been incubated with extracts from wild-type plants and treated according to method 2 (Fig. 7).

The results demonstrate that with the protocol presented here, it is possible to obtain sufficient amounts of metabolites, aiding in the characterization of enriched mitochondria, as suggested before (Ikuma, 1970). However, from the NADH/NAD and ATP/ADP ratios, it is clear that the mitochondria lost activity during Mito-AP (Roberts et al., 1997, Agius et al., 2001). This strongly suggests that metabolite concentrations in vivo cannot be determined with Mito-AP, except in the case of stable metabolites, such as lipids, or for entire metabolite groups, such as the adenylates (Fig. 7).

Metabolites were also detected in the negative controls of the recently described Mito-AP from mammalian cells, but unfortunately the actual magnitude of this background noise was not shown (Chen et al., 2016). It is possible that this background arose from metabolites directly bound to the matrix, because Chen et al. (2016) extracted the metabolites from the mitochondria in the presence of the magnetic beads. We minimized such effects by removing the beads prior to extraction, but it cannot be fully excluded that a minor amount of matrix remained in our preparations, resulting in some background.

To assess whether there was any metabolite contamination of our mitochondrial preparations from intact plastids, we used ADP-Glc as a marker metabolite. It is generated by ADP-Glc pyrophosphorylase for starch biosynthesis, which normally occurs in the
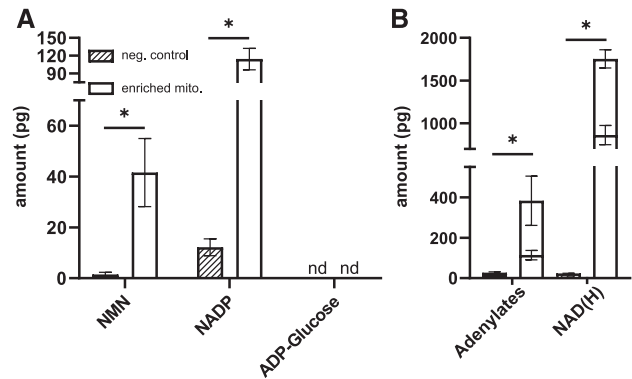

Figure 7. Polar metabolites in mitochondria enriched by Mito-AP. Seedlings carrying the 35S:Twin-Strep-tag:GFP:ADAPTER construct (enriched mito.) or Col-0 (neg. control) were grown for 10 d in liquid media and processed with Mito-AP (method 2). Polar metabolites were identified and quantified with a triple quadrupole MS in these preparations. A, NMN, NADP, and ADP Glc. The ADP-Glc content was below the detection limit of 0.65 fg (see Supplemental Fig. S10). B, Adenylates and reduced and oxidized NADH, each shown as a pool in one column, to emphasize that pools of these metabolites can be measured but that the concentrations of individual compounds within those pools do not represent the concentrations in vivo. For the adenylates the column is separated into AMP (lower), ADP (middle), and ATP (upper), whereas for the nicotinamides, NAD is shown in the lower part of the column and NADH in the upper part. The values for ATP were too low to be visualized. Error bars represent the SD ($n = 4$ biological replicates, derived from independent Mito-APs of distinct samples grown in parallel). Significant differences (at $P < 0.05$) were determined with Student's t test and marked with an asterisk. Calculation of the false discovery rate (FDR) was done using the two-stage linear step-up procedure described in Benjamini et al. (2006). The statistical analysis was done with a FDR of 1%. nd, Not detectable. Standard curves for different metabolite concentrations in water were used for absolute quantification.

plastids. Under special circumstances, cytosolic starch biosynthesis might occur (Villand and Kleczkowski, 1994); however, the process has never been observed in mitochondria. Dilutions of ADP-Glc down to 0.67 fg on column were detected with a robust signal-to-noise ratio by our analytical platform (Supplemental Fig. S11), but ADP-Glc was not detected in any of the samples obtained with Mito-AP (Fig. 7), indicating that intact plastids were not present.

Figure 6. (Continued.)
Correlation analysis of protein abundances in samples from DGC, the two Mito-AP methods, and tandem DGC-Mito-AP purifications. The normalized iBAQ values for all identified proteins from DGC and Mito-AP methods 1 and 2 as well as from tandem DGC + Mito-AP (methods 1 and 2) were determined. An average of the three biological replicates from each method was used to calculate Pearson's correlation coefficients for all possible combinations of two proteins. The correlations are displayed as an interaction graph using a cutoff value of 0.99. Proteins are nodes and the length of an edge is negatively correlated with the strength of the correlation. Black triangles and gray squares, proteins classified by SUBAcon as being located in mitochondria; gray squares, locations predicted only in silico, without experimental support (by MS or fluorescence microscopy [FM]); orange diamonds and red circles, proteins classified by SUBAcon as being located in other cellular compartments and not mitochondria; red circles, locations predicted only in silico, without experimental support; orange diamonds, experimental data are available. The same color scheme was used for the schematic below the interaction graph, representing the proportions of the respective protein groups.

Besides adenylates, the lipid composition of enriched mitochondria was also analyzed, and the abundances of individual lipids were compared between the Mito-AP and the negative control derived from magnetic beads incubated with the wild-type extract. Several cardiolipins, known to occur in mitochondria (Schlame et al., 1993; Zhou et al., 2016), were exclusively detected in extracts of the Mito-AP and absent in the control. Two cardiolipins were also detected in the control, but were over three orders of magnitude more abundant in the Mito-AP extracts (Fig. 8; Supplemental Table S3). Ubiquinone (Q), a membrane-soluble electron carrier in the electron transport chain of mitochondria, was strongly enriched in samples containing mitochondria. Qs are classified according to the number of isoprenoid side-chain units, which usually rank from 6 (for yeast [Q6]) to 10 (for humans [Q10]). Interestingly, Q7, Q8, Q9, and Q10 were detected (Fig. 8), although it has been reported that Q9 is the sole form of Q in Arabidopsis (Hirooka et al., 2003; Yoshida et al., 2010; Liu and Lu, 2016). The roles of these additional Qs in plant mitochondria are currently unknown, but it is tempting to speculate that they may function in nonrespiratory pathways, similar to the involvement of human Q10 in signaling (Schmelzer et al., 2007). The strong enrichment of mitochondria by Mito-AP in combination with the sensitive detection by MS probably explains why several Qs could be detected here, thus highlighting that resolution is indeed gained by the focused analysis of isolated mitochondria. Previous studies on cardiolipins (Zhou et al., 2016) and Qs (Yoshida et al., 2010)

had the advantage of employing a rapid quenching of the sample, abolishing any possibility for metabolite conversion. The cardiolipin species we observed could previously only be detected using a solid-phase extraction protocol (Zhou et al., 2016), suggesting that Mito-AP maybe useful for the initial detection of metabolites in samples of enriched mitochondria without need for further sample preparation. However, such findings are pending confirmation from samples that are more efficiently quenched than is possible with Mito-AP.

Chlorophyll, associated with thylakoid membranes of the chloroplast, was also slightly enriched, by ~2-fold, in Mito-AP samples compared to controls (Supplemental Table S3). When mitochondria were bound to the beads, we observed that the beads became adhesive to each other, probably caused by mitochondria interconnecting several beads. This network of beads may occasionally trap thylakoid fragments which cannot be removed by washing. In fact, we occasionally observed some red auto-fluorescence in our mitochondria preparations, but never corresponding to the size of an intact chloroplast (Supplemental Fig. S12).

Interestingly, galactolipids, which are associated with the chloroplast, were also enriched in the Mito-AP samples. For the majority of these lipid species, only a slight enrichment of 10- to 20-fold was observed, but in one case a 200-fold enrichment was obtained (Supplemental Table S3). As for the chlorophylls, the physical entrapment of plastidic fragments during Mito-AP might explain this observation. It is also conceivable that some of these galactolipids actually reside in the mitochondria, because transfer of galactolipids from plastids to mitochondria has been observed, particularly in conditions of phosphate starvation (Jouhet et al., 2004). In summary, this proof-of-concept study demonstrates the feasibility of metabolite quantification in mitochondria upon isolation using the Mito-AP protocol, which paves the way for further studies addressing problems of metabolite conversion during the isolation procedure.
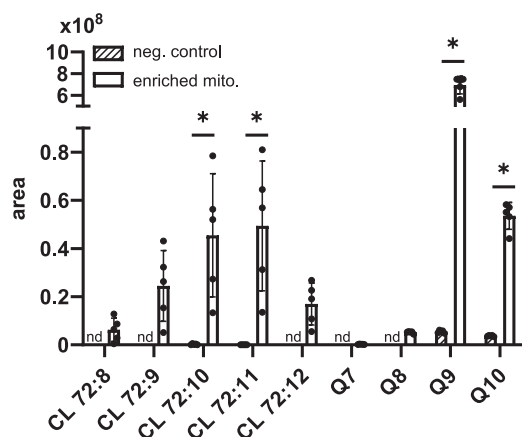

Figure 8. Enrichment of the mitochondrial cardiolipins (CL) and Qs in mitochondria enriched by Mito-AP. Seedlings carrying the 35S:Twin-Strep-tag:GFP:ADAPTER construct (enriched mito.) or Col-0 (neg. control) were grown for 10 d in liquid media for Mito-AP (method 2). Samples were extracted and lipids were identified and quantified with an orbitrap MS. Error bars represent the SD ($n = 5$ biological replicates, derived from independent Mito-APs of distinct samples grown in parallel). Significant differences (at $P < 0.05$) were determined with Student's t test and marked with an asterisk. Calculation of the FDR was done using the two-stage linear step-up procedure described in Benjamini et al. (2006). The statistical analysis was done with a FDR of 1%. nd, Not detectable.

Mitochondria Isolated by Mito-AP Have a Membrane Potential, Respire, and Show Respiratory Control

Previous studies on mitochondria isolation techniques reported that the isolated organelles remained physiologically active (Keech et al., 2005). We also asked the question whether mitochondria isolated by Mito-AP maintain their membrane potential and respiratory control. However, this could not be addressed with the sodium chloride extraction buffer, because it was chosen for its compatibility with downstream LC-MS measurements and its reported potential to suppress metabolic activity (Kong et al., 2018). Different buffers containing either mannitol or Suc as osmotic agents were used instead.

First, we tested whether mitochondria in intact roots of the 35S:Twin-Strep-tag:GFP:ADAPTER plants have

104

membrane potential, which would indicate that the mitochondrial tag does not grossly disturb mitochondrial function or integrity. The fluorescent dye tetramethyl rhodamine methyl ester (TMRM) is only sequestered in mitochondria in the presence of a membrane potential (Brand and Nicholls, 2011; Schwarzländer et al., 2012). TMRM was found in mitochondria of the transgenic line as well as the wild type (Supplemental Fig. S13, A and B). The signal was quenched by adding carbonyl cyanide 3-chlorophenylhydrazone (CCCP), an uncoupling agent that abolishes the membrane potential. The same technique was applied to mitochondria isolated with the Mito-AP protocol using the modified extraction buffer containing 10 mM succinate. In these samples, the green fluorescence originating from the Twin-Strep-tag:GFP:ADAPTER protein coincided with the fluorescence of the TMRM dye in 73% of the cases. This rate dropped to 32% upon addition of the uncoupler CCCP (Supplemental Fig. S14). Moreover, the signal from these mitochondria was considerably weaker. These results suggest that >70% of the mitochondria enriched by Mito-AP are sufficiently intact to maintain a membrane potential.

Next, we assessed the respiration of mitochondria isolated by Mito-AP in a Seahorse Analyzer (Agilent; Rogers et al., 2011) and recorded the oxygen consumption rate (OCR) before and after addition of ADP, as well as in response to the inhibitors oligomycin and antimycin (Fig. 9). A basal respiration rate of 29 ± 3 nmol O_2 min^{-1} μg^{-1} protein was observed. Addition of ADP led to a 1.3-fold increase in OCR, which reverted to the basal level after addition of the ATP synthase inhibitor oligomycin. Furthermore, addition of antimycin, an inhibitor of complex III, reduced the OCR to
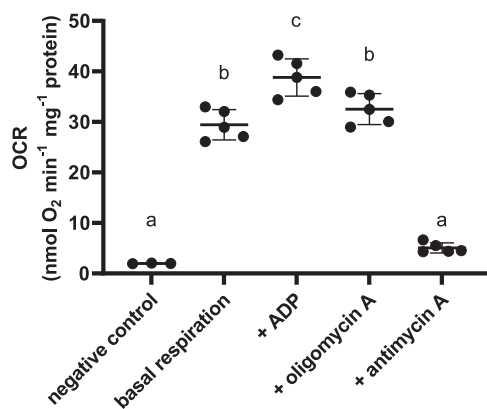
background levels. These data demonstrate that mitochondria with respiratory activity can be isolated by Mito-AP, and that they maintain some degree of respiratory control, as indicated by the before-to-after OCR ratio upon addition of ADP or oligomycin (respiratory control ratio; Fig. 9).

These results were confirmed by the morphological assessment of Mito-AP-isolated organelles as deduced by electron microscopy. Mitochondria display outer and inner membranes as well as an ultrastructure consistent with intact organelles isolated in previous studies (Supplemental Fig. S15; de Virville et al., 1998; Logan and Leaver, 2000; Eubel et al., 2007). In some instances, a pronounced gap between the inner and outer membrane can be observed, suggesting either some shrinkage of the mitoplasts or an inflation of the outer membrane. Reasons for this are currently unknown. We also observed grana stacks and other unidentified compartments consistent with plastidic fragments observed in confocal microscopy and the presence of nonmitochondrial proteins in the preparations (Fig. 5; Supplemental Fig. S12).

DISCUSSION

Here we demonstrate that plant mitochondria can be enriched from small amounts of plant material in sufficient amounts to support proteome and metabolome analyses using affinity purification. The application and further development of this technique offers the potential to increase the resolution of proteome and metabolome experiments. Because comparatively little plant material is needed for Mito-AP, the biological context from which the mitochondria are obtained can be well defined. It can be envisioned that mitochondrial isolation may be restricted to only a certain tissue or only tissues in defined developmental states, for example leaves of different ages. Mito-AP also offers the prospect to exclusively decorate mitochondria of certain cell types or following environmental stimuli. For this, the Twin-Strep:ADAPTER protein would need to be expressed under the control of a cell-type-specific promoter or a promoter induced by a particular stimulus. Previous studies have shown the power of promoter-reporter fusions, for example for the elucidation of the root cell type-specific transcriptomes (Moussaieff et al., 2013) and for the isolation of cell-specific mitochondria from *Caenorhabditis elegans* as well as mouse (*Mus musculus*; Ahier et al., 2018; Bayraktar et al., 2019). Yet another possibility is the transfer of the affinity purification principle to other compartments of plant cells, for example to peroxisomes or chloroplasts. Along this line, an immunoprecipitation protocol has been developed for the enrichment of plant secretory vesicles to analyze their proteomes (Heard et al., 2015).

The affinity purification protocol presented here has several features that will facilitate its adaptation and use by other laboratories. First, the construct was

Figure 9. Activity of mitochondria isolated by Mito-AP from Arabidopsis. OCRs were measured for mitochondria bound to magnetic beads before (basal respiration) and after addition of ADP and in response to oligomycin A and antimycin A. In total, five replicates were analyzed for mitochondria bound to beads and three replicates for the negative control. The negative control represents the basal respiration of the empty beads. Error bars represent the SD. Statistical analysis was done using a one-way ANOVA with Tukey's honestly significant difference (HSD) mean-separation test. Lowercase letters indicate significant differences ($P < 0.05$).

105

generated using the MoClo system (Patron et al., 2015), which is highly modular, allowing an independent exchange of different elements including the tags, the promoter, or the ADAPTER itself. Second, in contrast to DGC methods, no specialized equipment such as an ultracentrifuge is required to perform Mito-AP. Third, usage of the Strep-tag affinity system instead of a system based on antibodies (Chen et al., 2016) is cost effective. Mito-AP can be performed with a commercially available matrix already coated with Strep-Tactin. Alternatively, it is straightforward to attach Strep-Tactin to prepared iron oxide particles, further lowering the costs and increasing the yield (Supplemental Fig. S6).

Recently, a simple protocol was published for the economic production of iron oxide nanoparticles with carboxyl groups on the surface, called Bio on Magnetic Beads (Oberacker et al., 2019). We are currently investigating the suitability of these particles for Strep-Tactin loading. The availability of large quantities of affordable affinity matrix would permit an increase in the amount of beads used per preparation. This might improve the yield or accelerate the sample processing, because at high yields the magnetic separations do not need to be fully quantitative and can therefore be performed more swiftly. With shorter processing times, mitochondria could be isolated even closer to their native state. This is still a problem, because the ADP/ATP ratio (Fig. 6) showed that the mitochondria isolated by the current Mito-AP protocol are metabolically inactive (Stitt et al., 1982). An additional solution might be to prepare the mitochondria in the presence of buffers containing a broad selection of mitochondrial metabolite precursors suitable to maintain mitochondrial function during isolation. The successful reactivation of mitochondria with substrates for mitochondrial respiration was already demonstrated in this study (Fig. 9). In comparison to traditional isolation techniques, Mito-AP is performed in small volumes, which considerably reduces the costs of such an approach.

Conceptually, affinity elution from the matrix should increase the purity, because unspecifically bound contaminants will not be released and are removed together with the beads. Indeed, this was observed, because Mito-AP involving an elution step with biotin and proteinase K (method 2) resulted in higher purity than Mito-AP with method 1 (Fig. 5). However, it would be desirable to omit the proteinase K treatment, because it is time consuming and damages the mitochondrial surface proteins. Unfortunately, elution only with biotin is not efficient enough, probably because there are too many interactions between the Twin-Strep-decorated mitochondria and the Strep-Tactin on the matrix. The coating density of the matrix and the decoration density of the mitochondria should be further fine-tuned and optimized to develop a protocol that allows complete affinity elution. This might also reduce the number of beads connected to each other via mitochondria and thus limit the potential physical entrapment of contaminants in this network.

Additionally, different bead surfaces should be tested with the aim to suppress the unspecific binding of metabolites, which would further reduce the background currently observed in the controls (Figs. 7 and 8). It is clear that metabolites can be interconverted in a very short period of time (Dietz, 2017) and that therefore the duration of the Mito-AP protocol is still too long to avoid such processes. However, the impact of mutants on the mitochondrial metabolome still can be studied by focusing on relative changes in metabolites rather than absolute amounts. Also, groups of metabolites that can be interconverted into each other can be combined and considered as pools. Alternatively, improved quenching techniques prior to the Mito-AP protocol may address this challenge. Furthermore, Mito-AP will allow the identification of low-abundance metabolites that cannot be detected in a crude extract (Fig. 8) because of ion suppression effects.

Plant mitochondria have been analyzed numerous times by proteomic approaches. However, definition of the complete set of proteins residing in this organelle and the discernment of these proteins from background and false positives is an ongoing struggle. Previous studies indicated an inconsistency between computational prediction for mitochondrial localization and experimental evidence from proteomic studies. Several nuclear-encoded proteins with a clear mitochondrial target peptide were not detected in mitochondria by proteomics (Heazlewood et al., 2005, 2007). To increase sensitivity and to eliminate more false positives, tandem purification strategies for mitochondria have been employed combining DGC with free-flow electrophoresis (Eubel et al., 2007). However, both methods are very time consuming and require specialized equipment and expertise. Here, we used a combination of DGC and Mito-AP and performed a correlation analysis for protein abundance integrating the data from all enrichment strategies. Irrespective of the technique used, true mitochondrial proteins should be enriched and form a closely correlated cluster. This was observed, and in addition to known mitochondrial proteins, previously uncharacterized candidate proteins were found in this cluster (Fig. 6). More detailed analyses are required to validate these mitochondrial protein candidates, which might provide clues about yet unknown processes taking place in plant mitochondria.

An important consideration is whether mitochondria isolated by Mito-AP are of the same quality as that described using centrifugation-based methods. Since the quantities of mitochondria isolated by Mito-AP are comparatively low, standard methods for quality control of mitochondria, for example an oxygen electrode assay to assess respiration, cannot be applied easily. Here, we demonstrate oxygen consumption of plant mitochondria using a Seahorse Analyzer, which is uniquely suited for minute sample amounts. Recently, the oxygen consumption of mammalian mitochondria bound to beads was observed using the same technique; however, in contrast to our study, the impact of

106

ADP on respiration was not reported (Ahier et al., 2018), suggesting that the methodology for assessing isolated mitochondria with this technology is still in its infancy. In our hands, the method required optimization of the amount of mitochondria and the concentration of ADP used. Currently, the minimum assay temperature in the Seahorse analyzer cannot fall below 29°C, which differs from previously described procedures for assessing the activity of plant mitochondria. The respiration rates and the respiratory control we report here (Fig. 9) are lower than that described previously for Arabidopsis (Kerbler et al., 2019). However, it is possible that these differences can in part be explained by the current imperfections of the assessment technology, which needs further optimization.

Electron microscopy pictures suggested the presence of predominantly intact mitochondria with an outer and inner membrane in the Mito-AP preparation (Supplemental Fig. S15), but the presence of strongly fragmented mitochondria cannot be excluded. In line with this result, we demonstrated that 73% of the isolated mitochondria showed a membrane potential (Supplemental Fig. S14). Although no abnormal phenotypes were observed in the Twin-Strep-tag:GFP:ADAPTER plants, another explanation for a lower respiration rate/respiratory control is that expression of the Twin-Strep-tag:GFP:ADAPTER protein has an impact on mitochondrial functionality. This potential effect might be attenuated by further reducing the abundance of the protein with a weaker promoter or by removing the catalytic domain of the ADAPTER protein.

In summary, this study shows that Mito-AP allows rapid enrichment of plant mitochondria from small amounts of plant material. This work should, however, be regarded as a proof-of-concept pilot study that requires further experimentation to determine whether tagging the mitochondria interferes with their function, especially under stress conditions. Using this protocol, the quantification of mitochondrial metabolite subpools in the absence of any significant contamination was achieved, and previously uncharacterized candidate proteins for the complete mitochondrial proteome were identified. We envisage that affinity purification of organelles will increase the sensitivity of metabolome and proteome approaches. The potential of this method to target particular organelle subpopulations, in combination with the short processing time required for affinity isolation, will increase the spatial and temporal resolution for the investigation of plant organelles in the future.

MATERIALS AND METHODS

Plant Material and Cultivation

For proteomic analysis, Arabidopsis (*Arabidopsis thaliana*) plants were cultivated in soil (Steckmedium, Klasmann-Deilmann) in a climate chamber under short-day conditions (8 h light/16 h darkness, 22°C, 85 μmol s^{-1} m^{-2} light, and 65% humidity), as described in Senkler et al. (2017). Plants were harvested after

6 weeks. For metabolome analysis, plants were cultivated at 22°C in sterile 100-mL flasks (10 mg seeds per flask) in liquid culture (0.5 × Murashige and Skoog, 0.125% [w/v] MES, pH 5.8) in a shaker with an artificial light source emitting 45 μmol s^{-1} m^{-2} light. For the first 3 d, the shaker (New Brunswick Innova 42, Eppendorf) was set to 30 rpm and subsequently changed to 50 rpm. Plants were harvested after 10 d. For phenotypical characterization, plants were cultivated in 8-cm pots filled with soil (Steckmedium, Klasmann-Deilmann) in a climate chamber under long-day conditions (16 h light/8 h darkness, 22°C day, 20°C night, 100 μmol s^{-1} m^{-2} light, and 70% humidity). For leaf surface area and rosette fresh weight analysis, plants were cultivated in trays with soil (Steckmedium, Klasmann-Deilmann) under short-day conditions as described above. For root length and germination assays, plants were grown on modified Murashige and Skoog medium (3 mm CaCl$_2$, 1.5 mm MgSO$_4$, 1.25 mm KH$_2$PO$_4$, 18.7 mm KNO$_3$, 0.1 mm FeSO$_4$, 0.1 mm Na$_2$EDTA, 0.13 mm MnSO$_4$, 0.1 mm BO$_3$, 0.03 mm ZnSO$_4$, 1 μm Na$_2$MoO$_4$, 0.1 μm CuSO$_4$, 0.1 μm NiCl$_2$, 0.5 g L^{-1} MES, and 8 g L^{-1} phytoagar, pH adjusted to 5.7 with KOH) under long-day conditions as described above.

Subcellular Localization and Confocal Microscopy

For subcellular localization, the constructs were coexpressed in *Nicotiana benthamiana* leaves for 4 d and analyzed by confocal microscopy as described by Dahncke and Witte (2013). For confocal microscopy, the Leica True Confocal Scanner SP8 microscope equipped with an HC PL APO CS2 40× 1.10 water immersion objective (Leica Microsystems) was used. Acquired images were processed using Leica Application Suite Advanced Fluorescence software.

Magnetic Bead Preparation

The activation of the polysaccharide resins was described previously (Kohn and Wilchek, 1984) and coupling of Strep-Tactin to the magnetic beads was performed according to a protocol provided by Chemicell, with minor modifications. In detail, 400 μL uncoated magnetic beads (25 mg mL^{-1}) with a GlcA polymer matrix (fluidMAG ARA, 4115-5, Chemicell) were mixed with 150 μL freshly prepared activation buffer containing 10 mg 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide and mixed in a shaker for 10 min at room temperature (RT). Between the different steps, the beads were always pelleted using a magnetic separator. After activation, the beads were washed twice by suspending them in 1 mL water. For coating with Strep-Tactin, the bead pellet was resuspended in 250 μL water, 20 μL Strep-Tactin solution was added (5 mg mL^{-1} in phosphate-buffered saline [PBS]; 2-1204-005, IBA Lifesciences), and the slurry was incubated for 2 h under constant shaking. The beads were washed three times with 1 mL water and stored in 400 μL 0.05% (w/v) sodium azide solution.

500 μL beads with a starch polymer matrix (25 mg mL^{-1}; Chemicell, fluidMAG –D, 4101-5) were washed once with 1 mL freshly prepared activation buffer containing 0.2 m sodium hydrogen carbonate (pH range 8.4 to 8.7), pelleted, and resuspended in 500 μL activation buffer. After addition of 100 μL 5 m cyanogen bromide in acetonitrile, the beads were mixed and incubated for 10 min on ice, then washed twice with 1 mL PBS buffer and resuspended in 500 μL PBS. To ensure a homogenous coating, the beads were placed for 2 min in an ultrasonication bath. For coating, 500 μL magnetic beads were incubated with 20 μL Strep-Tactin solution (5 mg mL^{-1} in PBS; 2-1204-005, IBA Lifesciences) for 2 h under constant shaking. The beads were washed three times with 1 mL water and stored in 500 μL 0.05% (w/v) sodium azide solution.

Cloning

Cloning was performed using the MoClo system described by Engler et al. (2014). For this, several intermediate vectors in addition to already published vectors had to be created. The primers used are listed in Supplemental Table S4.

Turbo-GFP was amplified from pICSL50016 using P626 and P820 and cloned into pAGM1287, resulting in the vector pAGM1287_TurboGFP+Intron_without_Stop (V151). The TOM22-V fragment sequence was synthesized by Integrated DNA Technologies (IDT) and cloned into pAGM1301, resulting in the vector pAGM1301_level_1_TOM22-V (V152). The Cb5-ᴅ fragment sequence was synthesized by IDT and cloned into pAGM1301, resulting in the vector pAGM1301_level_0_Cb5-ᴅ (V153). The codon-optimized DNA sequence coding for the Twin-Strep-tag (28 amino acids: WSHPQFEK-GGGSGGGSGG-SA-WSHPQFEK) was synthesized by IDT and cloned into pAGM1276, resulting in the vector pAGM1276_Twin-Strep (V154). Amplification of the ADAPTER (encoded at locus At1g55450) using complementary DNA from

107

Arabidopsis was performed by using the primers P1023 + P1024 and P1025 + P1026, integrating a mutation to remove a *Bbs*I recognition site without changing the amino acid sequence. PCR products were digested, ligated, and cloned into pAGM1301, resulting in the vector pAGM1301_level_0_ADAPTER (V159).

The level_1 vector containing a phosphinothricin resistance for plant selection was generated by combining pICH87633, pICH43844, and pICH41421 with the recipient vector pICH47742 in a digestion/restriction reaction with *Bsa*I resulting in pICH47742_Basta_pos2_fwd (V166). For cloning the level_1 *Sc*CoxIV:mCherry vector, pICH45089, pAGM1482, pICSL80007, and pICH41432 were digested/ligated together with the recipient vector pICH47742, resulting in pICH47742 _pos2_fwd_ *Sc*CoxIV:mCherry (V168). For cloning the level_1 vector containing a 35S:Twin-Strep-tag:GFP:TOM22-V, pICH41373, pAGT707, V154, V151, V152, and pICH41421 were combined with the recipient vector pICH47732 in a digestion/restriction reaction with *Bsa*I resulting in pICH47732_pos1_fwd_35S:Twin-Strep-tag:GFP:TOM22-V (H405). For cloning the level_1 vector containing 35S:Twin-Strep-tag:GFP:Cb5-ᴅ, pICH41373, pAGT707, V154, V151, V153, and pICH41421 and the recipient vector pICH47732 were combined in a digestion/restriction reaction with *Bsa*I resulting in pICH47732_pos1_fwd_35S:Twin-Strep-tag:GFP:Cb5-ᴅ (H407). For cloning the level_1 vector containing the 35S:Twin-Strep-tag:GFP:ADAPTER, pICH41373, pAGT707, V154, V151, V159, and pICH41421 were combined with the recipient vector pICH47732 in a digestion/restriction reaction with *Bsa*I resulting in pICH47732_pos1_fwd_35S:Twin-Strep-tag:GFP:ADAPTER (H433).

For the level_2 vector containing 35S:Twin-Strep-tag:GFP:ADAPTER and a Basta resistance gene, H433, V166, and pICH41744 were combined with the recipient vector pAGM4723 in a digestion/restriction reaction with *Bbs*I resulting in the vector pAGM4723_35S:Twin-Strep-tag:GFP:ADAPTER_Basta (H500). The level_2 vector pAGM4723_35S:Twin-Strep-tag:GFP:ADAPTER_*Sc*CoxIV:mCherry (H501) was cloned by combining H433, pICH41744, V168, and pAGM4723 in a digestion/restriction reaction with *Bpi*I. All materials (plants and vectors) are made available upon request.

Leaf Area Quantification

To quantify the leaf area, RGB (red/green/blue) pictures were taken from the same distance every day and loaded into ImageJ. To create a binary image, the pictures colors were split into separate channels (Image → color → split channels). The green channel image was then further processed to reduce background signal by minimum filtering with a radius of 50 pixels followed by maximum filtering with a radius of 50 pixels and a contrast enhancement with a saturation of 0.01%. Subsequently a binary picture was created by setting an auto-threshold. The binary picture was modified (euclidian distance map binary options → close [20 iterations]) and the region of interest was projected onto the scaled original picture to retrieve the leaf area of the respective sample.

Density Gradient Isolation of Mitochondria

A modified version of the protocol of Keech et al. (2005) was used. All steps were carried out at 4°C. Arabidopsis rosette leaves were ground for 10 min in disruption buffer using a mortar and pestle (0.3 ᴍ Suc, 60 mᴍ TES, pH 8 [KOH], 25 mᴍ sodium pyrophosphate, 10 mᴍ KH_2PO_4, 2 mᴍ EDTA, 1 mᴍ Gly, 1% [w/v] PVP 40, 1% [w/v] bovine serum albumin [BSA], 50 mᴍ sodium ascorbate, and 20 mᴍ Cys) in the presence of sand. Two milliliters of disruption buffer was used for each gram of plant material. The homogenate was filtered through 4 layers of Miracloth and the dry cake was ground in disruption buffer for another 5 min before being filtered as outlined above. The pooled filtrate was centrifuged at 300*g* for 5 min and the resulting supernatant was centrifuged at 2,500*g* for 5 min once or twice, depending on the size of the pellet. A mitochondria-enriched pellet was produced by centrifuging the supernatant for 15 min at 15,100*g*. The pellet was subsequently resuspended in 1 mL washing buffer (0.3 ᴍ Suc, 10 mᴍ TES, pH 7.5 [KOH], and 10 mᴍ KH_2PO_4) and subjected to two strokes in a Dounce homogenizer. Mitochondria were further enriched on continuous Percoll gradients (50% [v/v] Percoll, 0.3 ᴍ Suc, 10 mᴍ TES [pH 7.5 with KOH], 1 mᴍ EDTA, 10 mᴍ KH_2PO_4, and 1 mᴍ Gly). Percoll gradients were established by centrifugation of the gradients at 69,400*g* for 40 min prior to loading. One milliliter of homogenized material was layered upon each of the 12 gradients, followed by centrifugation at 17,400*g* for 20 min. Mitochondria, concentrated near the bottom of the tube as a cloudy layer, were quantitatively removed using a glass pipette, resuspended in washing buffer, and centrifuged

at 17,200*g* for 15 min. This was repeated two to four times until the pellets were solid. Finally, pellets of all gradients were pooled in a last washing step.

Isolation of Plant Mitochondria with Mito-AP

For Mito-AP, 500 mg plant material was ground using a mortar and pestle with some sand and 1 mL of ice cold, freshly prepared extraction buffer (32 μL mL^{-1} BioLock [IBA Lifesciences], 100 mᴍ ammonium bicarbonate, and 200 mᴍ NaCl, pH 8) on ice for 3 min. The extract was transferred to a 2-mL reaction tube and centrifuged at 4°C for 5 min at 1,000*g*. Aliquots of magnetic beads (25 mg mL^{-1}; 500 μL for metabolome analysis and 250 μL for proteome analysis) were prepared by washing beads twice with 1 mL washing buffer (100 mᴍ ammonium bicarbonate and 200 mᴍ NaCl, pH 8) followed by resuspension in 500 μL washing buffer. The supernatant of the extract (800 μL) was added to the beads and incubated for 5 min at 4°C with continuous inversion. The beads were separated from the remaining liquid with a magnetic separator for 2 min. The supernatant was discarded and 1 mL washing buffer was added. Beads were resuspended by inverting and gentle shaking of the reaction tube. The washing was repeated three times. For method 1, the beads were directly incubated with 20 μL 2× SDS loading buffer and heated for 5 min at 95°C. For method 2, the beads were further treated by resuspension in 1 mL elution buffer (50 mᴍ biotin, 100 mᴍ ammonium bicarbonate, and 200 mᴍ NaCl, pH 8). The mixture was incubated for 5 min on ice and inverted every minute. Subsequently, 950 μL digest buffer (200 μg mL^{-1} proteinase K, 4 mᴍ $CaCl_2$, 100 mᴍ ammonium bicarbonate and 200 mᴍ NaCl, pH 8) was added and the slurry incubated at RT for 10 min. Proteinase K was inhibited by the addition of 50 μL 100 mᴍ phenylmethylsulfonyl fluoride (in pure isopropanol). The beads were separated for 2 min with a magnetic separator and the liquid was transferred with a cut tip, to prevent damaging the mitochondria, to a new 2-mL reaction tube followed by centrifugation at 4°C for 10 min at 14,500*g*. The supernatant was carefully removed and the pellet was stored at −80°C for further analysis.

Proteomic Analysis

Proteins from all crude extracts were first quantified with Bradford reagent to ensure a similar loading of SDS gels. Samples were resuspended in 20 μL 2× SDS loading buffer and heated for 5 min at 95°C. The proteins from all samples were additionally quantified by comparison with a BSA standard on the preparative SDS gel (Luo et al., 2006) employing a Li-cor Odyssey FC with Image Studio software. Sample preparation and proteomic analysis were performed according to Rugen et al. (2019). The volume used for resuspension of the peptides was adjusted for each sample to account for differences in the initial protein amounts. In fact, the total number of iBAQs in all samples with enriched mitochondria varied only by a factor of 3 (Supplemental Table S1). We found ~10-fold fewer iBAQs in crude extracts for differential and density gradient centrifugation (DGC), potentially indicating interference of the extraction buffer with in-gel quantification.

Correlation Analysis

Peptides from DGC, Mito-AP, and DGC combined with Mito-AP (method 1 and 2, respectively) were reanalyzed using a 4-h gradient for LC-MS/MS and otherwise treated as described above. All iBAQs were normalized to the total iBAQs in the respective sample and a correlation matrix, based on the Pearson correlation coefficient, was constructed in R (version 3.6.0; https://www.R-project.org). Correlations with a coefficient >0.99 were visualized by Cytoscape version 3.7.1 (Shannon, 2003), including an expression correlation plugin) with an edge-weighted spring-embedded layout, based on the correlation coefficient. Edges with a higher correlation coefficient are shorter.

Metabolome Analysis

The pellet of mitochondria derived from Mito-AP method 2 was solubilized with 0.5 mL 80:20 MeOH:5 mᴍ ammonium acetate (pH 9.5, adjusted with NH_3), vortexed, and sonicated for 2 min at RT. Samples were centrifuged for 10 min at 40,000*g*, and supernatants were dried down in a vacuum centrifuge concentrator (RVC 2-25 CD plus, Thermo Fisher Scientific) until dry. The resulting metabolites were resuspended in 11 μL 5 mᴍ ammonium acetate (pH 9.5, adjusted with NH_3). The chromatography of 10 μL solution was performed on a 1290 Infinity HPLC (Agilent Technologies) using a Hypercarb column (50 mm, 4.6-mm diameter; Thermo Fisher Scientific) with mobile phase A consisting of

5 mM ammonium acetate (pH 9.5, adjusted with NH₃) and mobile phase B consisting of pure acetonitrile. With a flow rate of 0.6 mL min⁻¹, the gradient was 0–7 min, 4% to 20% B; 7–11 min 10 s, 20% to 30% B; 11 min 10 s–15 min 10 s, 50% B; 15 min 10 s–15 min 20 s, 50% to 100% B; 15 min 20 s–17 min 42 s, 100% B; 17 min 42 s–18 min, 100% to 4% B; and 4% B until the end of the gradient (23 min). With this gradient, metabolites elute in the following order: NMN (4.6 min), ATP (6.1 min), ADP (6.33 min), AMP (6.35 min), NADP (9.1 min), and NAD/NADH (11.6 min). The masses for the metabolites were quantified on a 6470 triple quadrupole mass spectrometer (Agilent Technologies). The LC-Agilent Jet Stream-electrospray ionization-tandem MS measurements were conducted under multiple reaction monitoring in positive ion mode. Agilent Jet Stream-electrospray ionization source conditions were set as follows: gas temperature, 250°C; gas flow, 12 L min⁻¹; nebulizer gas, 20 psi; sheath gas temperature, 395°C; sheath gas flow, 12 L min⁻¹; capillary voltage, 3000 V; and nozzle voltage, 500 V. Absolute amounts were calculated by comparison with the respective external standard curves of known concentrations. Identification of compounds was confirmed by retention times identical with the external standards. For quantification, the multiple reaction monitoring with the highest signal was used as quantifier. Other transitions, if available, were used as additional qualifiers. Transitions, fragmentors, and collision energies for all compounds can be found in Supplemental Table S5. For quantification, the Agilent Mass Hunter Quantitative Analysis Software was used.

Lipidomic Analysis

The pellet of mitochondria from Mito-AP method 2 was extracted according to a described method for the extraction and analysis of lipids, metabolites, proteins, and starch (Salem et al., 2016). In brief, frozen beads were resuspended in 1 mL extraction buffer (methyl tert-butyl ether:MeOH, 3:1 [v/v]) and the samples were incubated on a shaker for 45 min before a 15-min sonication in an ice-cooled sonication bath was applied. Phase separation was achieved by adding 0.65 mL of water:MeOH (3:1). Lipids were analyzed from the organic (methyl tert-butyl ether) phase by ultra-performance LC-MS. For this purpose, an Acquity iClass (Waters) ultra-performance LC was connected to a Q-Exactive HF (Thermo Scientific) high-resolution MS. Samples were measured in positive ionization mode as described previously (Giavalisco et al., 2011; Hummel et al., 2011; Salem and Giavalisco, 2018;). Data analysis was performed using targeted peak extraction and integration using the Trace Finder software (Version 4.1, Thermo Scientific).

Visualization of the Mitochondrial Membrane Potential

Twelve-day-old seedlings originating from liquid culture were washed twice with water and incubated in 500 μL one-half strength Murashige and Skoog medium containing 40 nM TMRM (dissolved in dimethyl sulfoxide) for at least 15 min. For analyzing the decoupled state of the mitochondria, 40 μM CCCP (dissolved in dimethyl sulfoxide) was added additionally to the solution to a final concentration of 40 μM and incubated for 30 min.

Isolated mitochondria bound to magnetic beads were resuspended in 100 μL mitochondrial assay solution buffer (70 mM Suc, 219 mM mannitol, 5 mM HEPES, 1 mM EGTA, and 0.5% [w/v] BSA, pH 7.2) containing 10 mM succinate and treated as described.

The analysis of the samples was done by using confocal microscopy as described above. The TMRM dye was excited at 552 nm and the emission between 600 and 615 nm was visualized.

Respiratory Activity Measurement

For the respiratory activity, 500 mg plant material was ground up using a mortar and pestle with some sand and 1 mL ice cold, freshly prepared extraction buffer (32 μL mL⁻¹ BioLock [IBA Lifesciences]; 5 mM dithiothreitol, 300 mM Suc, 5 mM KH₂PO₄, 10 mM TES, 10 mM NaCl, 2 mM MgSO₄, and 0.1% [w/v] BSA, pH 7.2) on ice for 2 min. The extract was transferred to a 2 mL reaction tube and centrifuged at 4°C for 5 min at 1,000g. Five hundred microliters of magnetic beads (25 mg mL⁻¹) was prepared by washing twice with 1 mL washing buffer (300 mM Suc, 5 mM KH₂PO₄, 10 mM TES, 10 mM NaCl, 2 mM MgSO₄, and 0.1% [w/v] BSA, pH 7.2), followed by resuspension in 500 μL washing buffer. The subsequent steps were performed as described in "Isolation of Plant Mitochondria with Mito-AP."

For the measurement of the respiration through succinate dehydrogenase (complex II), the beads were resuspended in respiration buffer 2 (300 mM Suc,

5 mM KH₂PO₄, 10 mM TES, 10 mM NaCl, 2 mM MgSO, and 0.1% [w/v] BSA, 10 mM succinate, 500 μM ATP, and 500 μM n-propyl gallate, pH 7.2) at RT.

The OCR measurements were performed using the Seahorse XFe 96 Analyzer (Agilent Technologies). The XFe96 Sensor Cartridge was prepared according to the manufacturer's instructions. The ports were filled as follows: port A, 20 μL 60 mM ADP, pH adjusted to 7.2 with KOH; port B, 22 μL 50 μg mL⁻¹ oligomycin A; port C, 24 μL 80 mM antimycin A; port D, 26 μL buffer. All chemicals were prepared in the corresponding respiration buffer.

Resuspended beads were loaded into the XF96 cell culture microplate and the plate was subsequently centrifuged for 5 min at 2,200g. The wells were filled with the corresponding respiration buffer to a final volume of 200 μL. The four corners of the plate were not used for sample measurements.

The respiration assay was performed at 29°C (internal heater was turned off). After equilibration for 12 min, the baseline was measured by mixing for 30 s, waiting 10 s, and measuring for 2 min. The same procedure was done after the injection of the single ports. After the measurements, the protein amount was quantified for every sample.

Transmission Electron Microscopy

Samples, prepared as described above for the respiratory measurements, were fixed in 0.15 M HEPES, pH 7.35, containing 1.5% (w/v) formaldehyde and 1.5% (w/v) glutaraldehyde for 30 min at RT and overnight at 4°C. Samples were immobilized in 2% agarose (w/v) and then incubated in aqueous solutions of 1% (w/v) OsO4 (2 h at RT) and 1% (w/v) uranyl acetate (overnight at 4°C). After dehydration in acetone, samples were embedded in Epon, and 60-nm sections were stained with uranyl acetate and lead citrate (Reynolds, 1963). Images were recorded using a Morgagni transmission electron microscope (FEI) with a side-mounted Veleta charge-coupled device camera.

Accession Numbers

Sequences for the ADAPTER protein and TOM22-V are provided by Araport with the accession numbers At1g55450 and At5g43970, respectively. The Cb5-D sequence is available under the GenBank accession AY578730.

Supplemental Material

The following supplemental materials are available.

Supplemental Figure S1. Transient expression of different tail-anchored proteins fused to GFP in *N. benthamiana*.

Supplemental Figure S2. Quantitative comparison of phenotypical traits between the wild type (Col-0) and a stable 35S:Twin-Strep-tag:GFP:ADAPTER construct.

Supplemental Figure S3. Integrity of isolated mitochondria in different buffers.

Supplemental Figure S4. Plant mitochondria bound to a commercial matrix of 1-μm diameter.

Supplemental Figure S5. Different bead sizes and coating materials influence the enrichment of mitochondria.

Supplemental Figure S6. Quantification of the relative protein amounts from mitochondria enriched with different matrices.

Supplemental Figure S7. Quantification of Strep-Tactin bound to different magnetic matrices.

Supplemental Figure S8. Plant material needed for DGC and Mito-AP and relative mitochondrial protein yield obtained by both methods.

Supplemental Figure S9. Quantification of the total protein yield obtained from elution with and without proteinase K.

Supplemental Figure S10. Variation of mitochondrial purity from biological repeats processed by DGC or Mito-AP.

Supplemental Figure S11. ADP-Glc standard curve.

Supplemental Figure S12. Contamination with plastidic fragments.

Supplemental Figure S13. Visualization of the mitochondrial membrane potential in vivo.

Supplemental Figure S14. Visualization of the mitochondrial membrane potential in vitro.

Supplemental Figure S15. Electron micrographs of isolated mitochondria.

Supplemental Table S1. Proteins identified by proteomics for the determination of mitochondrial purity.

Supplemental Table S2. Proteins identified in a 4-h LC gradient for correlation analysis.

Supplemental Table S3. Lipids identified by LC-MS in Mito-AP (method 2) from 35S:Twin-Strep-tag:GFP:ADAPTER plants or Col-0 plants.

Supplemental Table S4. Primer sequences.

Supplemental Table S5. MRM parameters used for detection of polar metabolites with a triple quadrupole MS.

ACKNOWLEDGMENTS

LITERATURE CITED

Agius SC, Rasmusson AG, Møller IM (2001) NAD(P) turnover in plant mitochondria. Aust J Plant Physiol **28:** 461–470

Ahier A, Dai C-Y, Tweedie A, Bezawork-Geleta A, Kirmes I, Zuryn S (2018) Affinity purification of cell-specific mitochondria from whole animals resolves patterns of genetic mosaicism. Nat Cell Biol **20:** 352–360

Afanasyeva MA, Ustiugova AS, Golyshev SA, Kopylov AT, Bogolyubova AV, Demin DE, Belousov PV, Schwartz AM (2018) Isolation of large amounts of highly pure mitochondria for "omics" studies. Biochemistry (Mosc) **83:** 76–85

Arrivault S, Guenther M, Florian A, Encke B, Feil R, Vosloh D, Lunn JE, Sulpice R, Fernie AR, Stitt M, et al (2014) Dissecting the subcellular compartmentation of proteins and metabolites in Arabidopsis leaves using non-aqueous fractionation. Mol Cell Proteomics **13:** 2246–2259

Bayraktar EC, Baudrier L, Özerdem C, Lewis CA, Chan SH, Kunchok T, Abu-Remaileh M, Cangelosi AL, Sabatini DM, Birsoy K, et al (2019) MITO-Tag mice enable rapid isolation and multimodal profiling of mitochondria from specific cell types in vivo. Proc Natl Acad Sci USA **116:** 303–312

Benjamini Y, Krieger AM, Yekutieli D (2006) Adaptive linear step-up procedures that control the false discovery rate. Biometrika **93:** 491–507

Boutry M, Faber AM, Charbonnier M, Briquet M (1984) Microanalysis of plant mitochondrial protein synthesis products: Detection of variant polypeptides associated with cytoplasmic male sterility. Plant Mol Biol **3:** 445–452

Brand MD, Nicholls DG (2011) Assessing mitochondrial dysfunction in cells. Biochem J **435:** 297–312

Chen WW, Freinkman E, Wang T, Birsoy K, Sabatini DM (2016) Absolute quantification of matrix metabolites reveals the dynamics of mitochondrial metabolism. Cell **166:** 1324–1337

Dahncke K, Witte C-P (2013) Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in Arabidopsis. Plant Cell **25:** 4101–4109

Day DA, Neuburger M, Douce R (1985) Biochemical-characterization of chlorophyll-free mitochondria from pea leaves. Aust J Plant Physiol **12:** 219–228

de Virville JD, Alin MF, Aaron Y, Remy R, Guillot-Salomon T, Cantrel C (1998) Changes in functional properties of mitochondria during growth

cycle of Arabidopsis thaliana cell suspension cultures. Plant Physiol Biochem **36:** 347–356

Dietz K-J (2017) Subcellular metabolomics: The choice of method depends on the aim of the study. J Exp Bot **68:** 5695–5698

Engler C, Youles M, Gruetzner R, Ehnert T-M, Werner S, Jones JDG, Patron NJ, Marillonnet S (2014) A golden gate modular cloning toolbox for plants. ACS Synth Biol **3:** 839–843

Eubel H, Lee CP, Kuo J, Meyer EH, Taylor NL, Millar AH (2007) TECHNICAL ADVANCE: Free-flow electrophoresis for purification of plant mitochondria by surface charge. Plant J **52:** 583–594

Fürtauer L, Küstner L, Weckwerth W, Heyer AG, Nägele T (2019) Resolving subcellular plant metabolism. Plant J **100:** 438–455

Gardeström P, Wigge B (1988) Influence of photorespiration on ATP/ADP ratios in the chloroplasts, mitochondria, and cytosol, studied by rapid fractionation of barley (Hordeum vulgare) protoplasts. Plant Physiol **88:** 69–76

Gerhardt R, Heldt HW (1984) Measurement of subcellular metabolite levels in leaves by fractionation of freeze-stopped material in nonaqueous media. Plant Physiol **75:** 542–547

Giavalisco P, Li Y, Matthes A, Eckhardt A, Hubberten H-M, Hesse H, Segu S, Hummel J, Köhl K, Willmitzer L (2011) Elemental formula annotation of polar and lipophilic metabolites using ^{13}C, ^{15}N and ^{34}S isotope labelling, in combination with high-resolution mass spectrometry. Plant J **68:** 364–376

Heard W, Sklenář J, Tomé DFA, Robatzek S, Jones AM (2015) Identification of regulatory and cargo proteins of endosomal and secretory pathways in Arabidopsis thaliana by proteomic dissection. Mol Cell Proteomics **14:** 1796–1813

Heazlewood JL, Tonti-Filippini J, Verboom RE, Millar AH (2005) Combining experimental and predicted datasets for determination of the subcellular location of proteins in Arabidopsis. Plant Physiol **139:** 598–609

Heazlewood JL, Verboom RE, Tonti-Filippini J, Small I, Millar AH (2007) SUBA: The Arabidopsis Subcellular Database. Nucleic Acids Res **35:** D213–D218

Hirooka K, Bamba T, Fukusaki E, Kobayashi A (2003) Cloning and kinetic characterization of Arabidopsis thaliana solanesyl diphosphate synthase. Biochem J **370:** 679–686

Hooper CM, Castleden IR, Tanz SK, Aryamanesh N, Millar AH (2017) SUBA4: The interactive data analysis centre for Arabidopsis subcellular protein locations. Nucleic Acids Res **45**(D1): D1064–D1074

Hornig-Do H-T, Günther G, Bust M, Lehnartz P, Bosio A, Wiesner RJ (2009) Isolation of functional pure mitochondria by superparamagnetic microbeads. Anal Biochem **389:** 1–5

Hummel J, Segu S, Li Y, Irgang S, Jueppner J, Giavalisco P (2011) Ultra performance liquid chromatography and high resolution mass spectrometry for the analysis of plant lipids. Front Plant Sci **2:** 54

Hwang YT, Pelitire SM, Henderson MPA, Andrews DW, Dyer JM, Mullen RT (2004) Novel targeting signals mediate the sorting of different isoforms of the tail-anchored membrane protein cytochrome b5 to either endoplasmic reticulum or mitochondria. Plant Cell **16:** 3002–3019

Ikuma H (1970) Necessary conditions for isolation of tightly coupled higher plant mitochondria. Plant Phys **45:** 773–781

Jouhet J, Marechal E, Baldan B, Bligny R, Joyard J, Block MA (2004) Phosphate deprivation induces transfer of DGDG galactolipid from chloroplast to mitochondria. J Cell Biol **167:** 863–874

Keech O, Dizengremel P, Gardestrom P (2005) Preparation of leaf mitochondria from Arabidopsis thaliana. Physiol Plant **124:** 403–409

Kerbler SM, Taylor NL, Millar AH (2019) Cold sensitivity of mitochondrial ATP synthase restricts oxidative phosphorylation in Arabidopsis thaliana. New Phytol **221:** 1776–1788

Klein M, Binder S, Brennicke A (1998) Purification of mitochondria from Arabidopsis. In JM Martinez-Zapater, and J Salinas, eds, Arabidopsis Protocols. Humana Press, Totowa, NJ, pp 49–53

Klodmann J, Senkler M, Rode C, Braun HP (2011) Defining the protein complex proteome of plant mitochondria. Plant Physiol **157:** 587–598

Kohn J, Wilchek M (1984) The use of cyanogen bromide and other novel cyanylating agents for the activation of polysaccharide resins. Appl Biochem Biotechnol **9:** 285–305

Kong Z, Jia S, Chabes AL, Appelblad P, Lundmark R, Moritz T, Chabes A (2018) Simultaneous determination of ribonucleoside and deoxyribonucleoside triphosphates in biological samples by hydrophilic

interaction liquid chromatography coupled with tandem mass spectrometry. Nucleic Acids Res 46: e66

Kraner ME, Müller C, Sonnewald U (2017) Comparative proteomic profiling of the choline transporter-like1 (CHER1) mutant provides insights into plasmodesmata composition of fully developed *Arabidopsis thaliana* leaves. Plant J 92: 696–709

Lilley RM, Stitt M, Mader G, Heldt HW (1982) Rapid fractionation of wheat leaf protoplasts using membrane filtration. Plant Physiol 70: 965–970

Liu M, Lu S (2016) Plastoquinone and ubiquinone in plants: Biosynthesis, physiological function and metabolic engineering. Front Plant Sci 7: 1898

Logan DC, Leaver CJ (2000) Mitochondria-targeted GFP highlights the heterogeneity of mitochondrial shape, size and movement within living plant cells. J Exp Bot 51: 865–871

Luo S, Wehr NB, Levine RL (2006) Quantitation of protein on gels and blots by infrared fluorescence of Coomassie blue and Fast Green. Anal Biochem 350: 233–238

Maarse AC, Van Loon AP, Riezman H, Gregor I, Schatz G, Grivell LA (1984) Subunit IV of yeast cytochrome c oxidase: Cloning and nucleotide sequencing of the gene and partial amino acid sequencing of the mature protein. EMBO J 3: 2831–2837

Maggio C, Barbante A, Ferro F, Frigerio L, Pedrazzini E (2007) Intracellular sorting of the tail-anchored protein cytochrome *b*5 in plants: A comparative study using different isoforms from rabbit and *Arabidopsis*. J Exp Bot 58: 1365–1379

Marty NJ, Teresinski HJ, Hwang YT, Clendening EA, Gidda SK, Sliwinska E, Zhang D, Miernyk JA, Brito GC, Andrews DW, et al (2014) New insights into the targeting of a subset of tail-anchored proteins to the outer mitochondrial membrane. Front Plant Sci 5: 426

Moussaieff A, Rogachev I, Brodsky L, Malitsky S, Toal TW, Belcher H, Yativ M, Brady SM, Benfey PN, Aharoni A (2013) High-resolution metabolic mapping of cell types in plant roots. Proc Natl Acad Sci USA 110: E1232–E1241

Oberacker P, Stepper P, Bond DM, Höhn S, Focken J, Meyer V, Schelle L, Sugrue VJ, Jeunen G-J, Moser T, et al (2019) Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation. PLoS Biol 17: e3000107

Patron NJ, Orzaez D, Marillonnet S, Warzecha H, Matthewman C, Youles M, Raitskin O, Leveau A, Farré G, Rogers C, et al (2015) Standards for plant synthetic biology: A common syntax for exchange of DNA parts. New Phytol 208: 13–19

Pétriacq P, de Bont L, Genestout L, Hao J, Laureau C, Florez-Sarasa I, Rzigui T, Queval G, Gilard F, Mauve C, et al (2017) Photoperiod affects the phenotype of mitochondrial complex I mutants. Plant Physiol 173: 434–455

Reynolds ES (1963) The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. J Cell Biol 17: 208–212

Roberts J, Aubert S, Gout E, Bligny R, Douce R (1997) Cooperation and competition between adenylate kinase, nucleoside diphosphokinase, electron transport, and ATP synthase in plant mitochondria studied by ^{31}P-nuclear magnetic resonance. Plant Physiol 113: 191–199

Rogers GW, Brand MD, Petrosyan S, Ashok D, Elorza AA, Ferrick DA, Murphy AN (2011) High throughput microplate respiratory measurements using minimal quantities of isolated mitochondria. PLoS One 6: e21746

Rugen N, Straube H, Franken LE, Braun H-P, Eubel H (2019) Complexome profiling reveals association of PPR proteins with ribosomes in the mitochondria of plants. Mol Cell Proteomics 18: 1345–1362

Salem MA, Giavalisco P (2018) Semi-targeted lipidomics of plant acyl lipids using UPLC-HR-MS in combination with a data-independent acquisition mode. Methods Mol Biol 1778: 137–155

Salem MA, Jüppner J, Bajdzienko K, Giavalisco P (2016) Protocol: A fast, comprehensive and reproducible one-step extraction method for the rapid preparation of polar and semi-polar metabolites, lipids, proteins, starch and cell wall polymers from a single sample. Plant Methods 12: 45

Schlame M, Brody S, Hostetler KY (1993) Mitochondrial cardiolipin in diverse eukaryotes. Comparison of biosynthetic reactions and molecular acyl species. Eur J Biochem 212: 727–735

Schmelzer C, Lindner I, Vock C, Fujii K, Döring F (2007) Functional connections and pathways of coenzyme Q10-inducible genes: An insilico study. IUBMB Life 59: 628–633

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. Nature 473: 337–342

Schwarzländer M, Logan DC, Johnston IG, Jones NS, Meyer AJ, Fricker MD, Sweetlove LJ (2012) Pulsing of membrane potential in individual mitochondria: A stress-induced mechanism to regulate respiratory bioenergetics in *Arabidopsis*. Plant Cell 24: 1188–1201

Senkler J, Senkler M, Eubel H, Hildebrandt T, Lengwenus C, Schertl P, Schwarzländer M, Wagner S, Wittig I, Braun H-P (2017) The mitochondrial complexome of *Arabidopsis thaliana*. Plant J 89: 1079–1092

Shannon P (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498–2504

Stitt M, Lilley RM, Heldt HW (1982) Adenine nucleotide levels in the cytosol, chloroplasts, and mitochondria of wheat leaf protoplasts. Plant Physiol 70: 971–977

Villand P, Kleczkowski LA (1994) Is there an alternative pathway for starch biosynthesis in cereal seeds. Z Natforsch C J Biosci 49: 215–219

Yoshida K, Shibata M, Terashima I, Noguchi K (2010) Simultaneous determination of in vivo plastoquinone and ubiquinone redox states by HPLC-based analysis. Plant Cell Physiol 51: 836–841

Zhou Y, Peisker H, Dörmann P (2016) Molecular species composition of plant cardiolipin determined by liquid chromatography mass spectrometry. J Lipid Res 57: 1308–1321
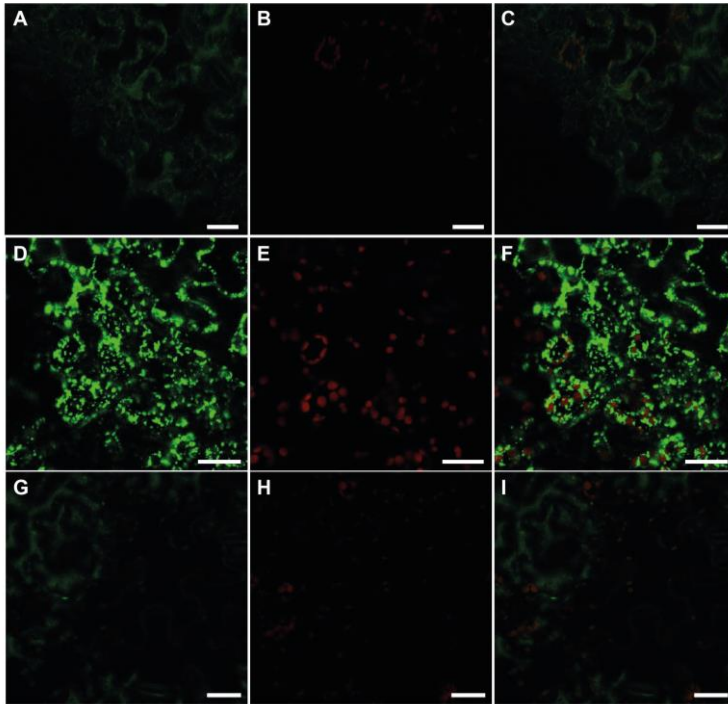
111

Supplemental Figure S1. Transient expression of different tail-anchored proteins fused to GFP in *Nicotiana benthamiana.* Confocal fluorescence microscopy images of *N. benthamiana* transiently expressing 35S:Twin-Strep-tag:GFP:TOM22-V (A to C), 35S:Twin-Strep-tag:GFP:ADAPTER (D to F), and 35S:Twin-Strep-tag:GFP:Cb5-D (G to I). GFP fluorescence (A, D, G), chlorophyll autofluorescence (B, E, H) and overlay of both channels (C, F, I). Scale bars, 25 μm.
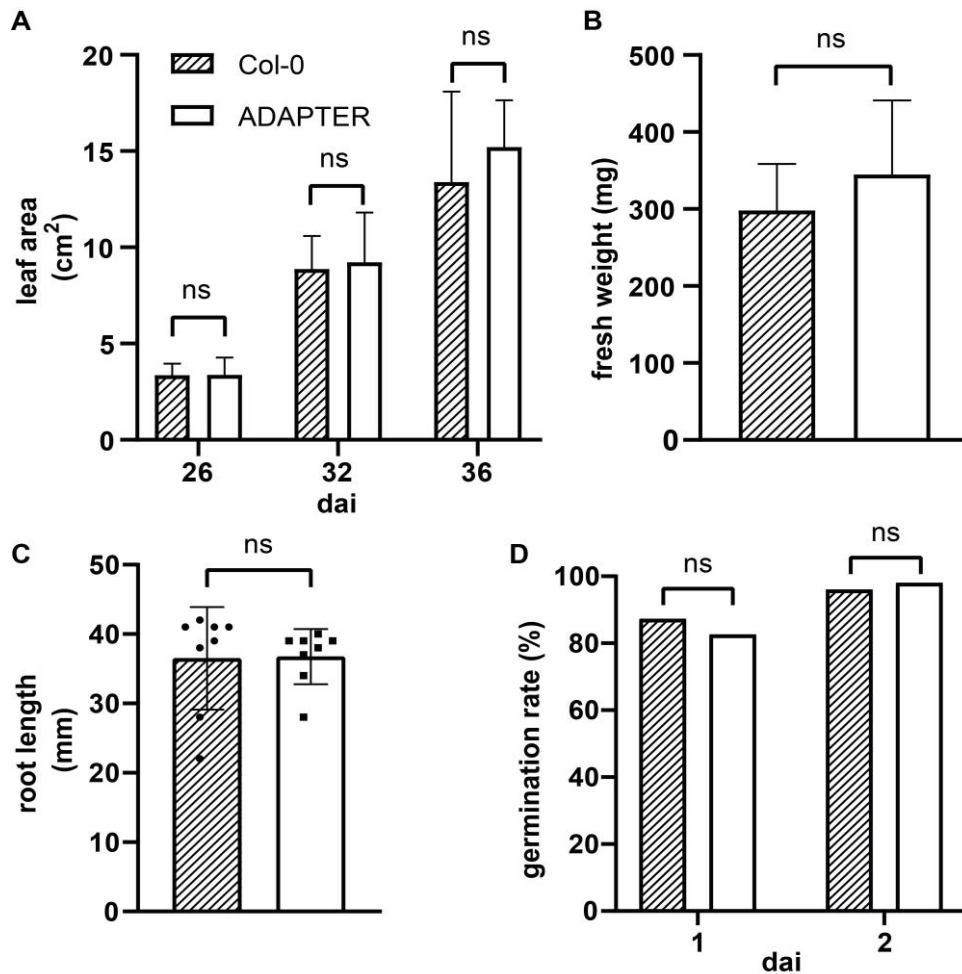
Supplemental Figure S2. Quantitative comparison of phenotypical traits between wild type (Col-0) and a stable transformed plant carrying the 35S:Twin-Strep-tag:GFP:ADAPTER. (A) Leaf area of plants at 26, 32 and 36 days after imbibition (dai) grown under short day conditions, n = 20. (B) Rosette fresh weight of plants 36 dai grown under short day conditions. (C) Root length of plants 12 dai grown on modified ½ MS media. Eight plants of each line were grown in parallel on one plate. The experiment was repeated three times with similar results. Error bars are SD. Significant differences (at $p < 0.05$) were determined with Student´s t-test and marked with a star. Calculation of the False Discovery Rate (FDR) was done using the two-stage linear step-up procedure described in Benjamini et al., 2006. The statistical analysis was done with a FDR of 1 %. (D) Germination rate of seeds on modified ½ MS media. In total 150 seeds were evaluated. For statistical evaluation the chi-square test X^2 was employed with (1, n = 150) = 1.281, p = 0.257705 for 1 dai and (1, n = 150) = 1.0309, p = 0.309941 for 2 dai. 'ns' denotes for 'not significant'.
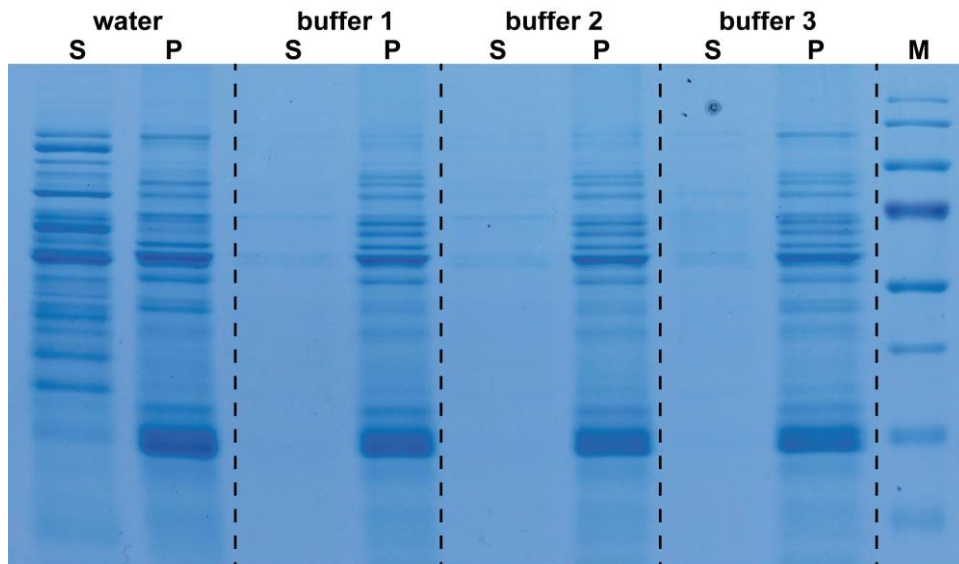
Supplemental Figure S3. Integrity of isolated mitochondria in different buffers. Plant mitochondria isolated from *Arabidopsis thaliana* with differential and density gradient centrifugation (DGC) were carefully resuspended in water, or buffer 1 (100 mM NaCl, 100 mM ammonium bicarbonate, pH 8), or buffer 2 (100 mM KCl, 100 mM ammonium bicarbonate, pH 8) or buffer 3 (400 mM mannitol, 1 mM EGTA, 10 mM Tricine, pH 7.2). These samples were centrifuged at 14500g for 10 minutes at 4°C and the supernatants were discarded. This procedure was rapidly repeated to ensure complete removal of the original buffer used for DGC. The resulting pellets were resuspended in the respective buffers and incubated on ice for 20 min and centrifuged again. SDS loading buffer was added to the supernatants (S) and pellets (P) and the samples were analyzed by SDS gel electrophoresis. For the marker (M) the PageRuler Prestained Protein Ladder (Thermo Fisher Scientific) was used. The gel was stained with colloidal Coomassie. For all proteome and metabolome analyses buffer 1 was employed, but the NaCl concentration was increased to 200 mM.
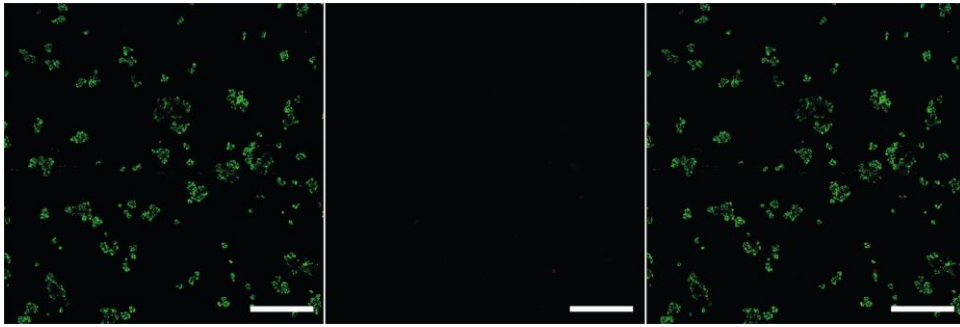
Supplemental Figure S4. Plant mitochondria bound to a commercial matrix of 1 μm diameter. Representative confocal fluorescence microscopy images of magnetic beads (1-μm diameter, IBA Lifesciences) with attached mitochondria extracted from *A. thaliana* carrying the 35S:Twin-Streptag: GFP:ADAPTER construct. GFP fluorescence (left panel), chloroplast autofluorescence (middle panel) and overlay of both channels (right panel). Scale bars, 25 μm.
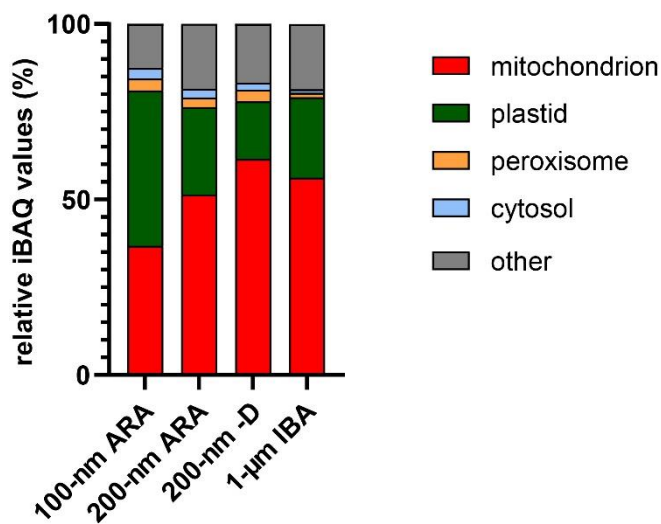

Supplemental Figure S5. Different bead sizes and coating materials influence the enrichment of mitochondria. 10-day-old seedlings of *A. thaliana* plants grown in liquid culture and expressing 35S:Twin-Strep-tag:GFP:ADAPTER were subjected to Mito-AP (method 1) with different bead types. Beads of 100-nm or 200-nm diameter coated in-house with Strep-Tactin using carboxyl (ARA; glucoronic acid as matrix polymer) or hydroxyl (-D; starch as matrix polymer) functional groups on the surface were used, or beads of 1-μm diameter already provided with a Strep-Tactin coating (IBA Lifesciences) were employed. iBAQ values for every identified protein were assigned to the respective categories (mitochondrion, plastid, peroxisome, cytosol, other), referred to total iBAQs of the corresponding sample, and plotted as a percentage. The categories were defined by the SUBAcon annotation provided by SUBA4.
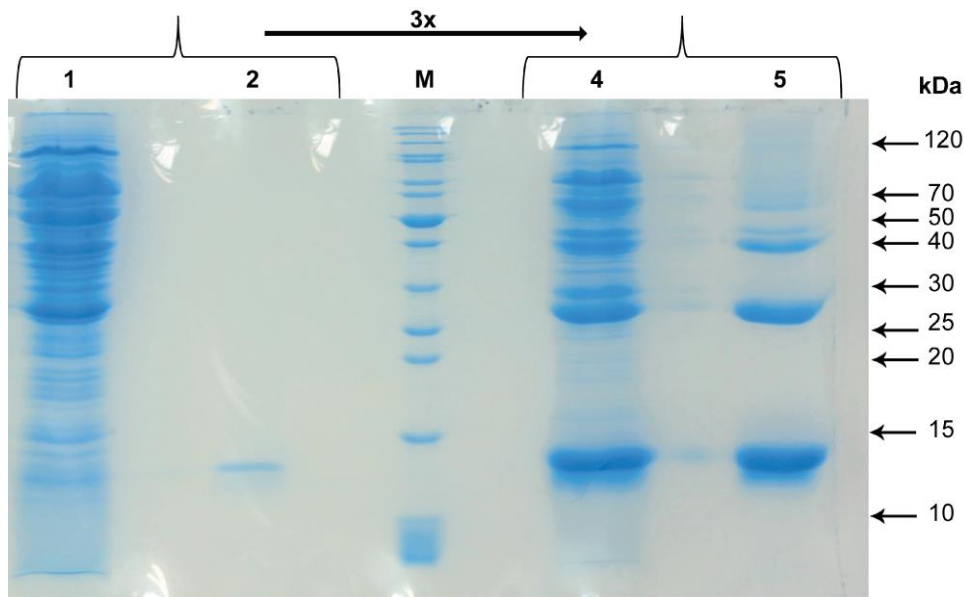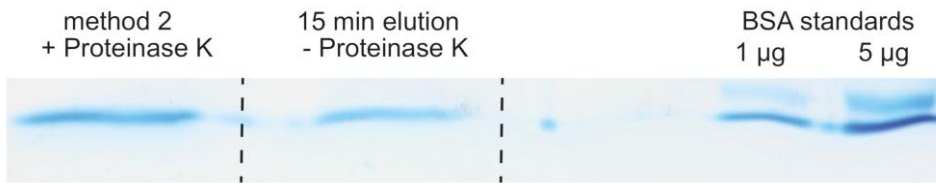
Supplemental Figure S6. Quantification of the relative protein amounts from mitochondria enriched with different matrices. 10-day-old seedlings of *A. thaliana* plants grown in liquid culture and expressing 35S:Twin-Strep-tag:GFP:ADAPTER were subjected to Mito-AP (method 1) with either 200-nm in-house-coated nanoparticles (lane 1) or 1-µm commercially available beads (lane 4). For all preparations, the same volume of beads was used. To assess which proteins originate from the bead coating and not from mitochondria, empty 200-nm and 1-µm beads were boiled in SDS buffer. The resulting proteins are shown in lanes 2 and 5, respectively. All bands were quantified with a Li-Cor Odyssee FC and signals from empty beads were subtracted from the corresponding loaded beads and the ratio between both signals was calculated. Three times more proteins were purified with the 200-nm beads compared to the 1-µm matrix. For the marker (M) the PageRuler Unstained Protein Ladder (Thermo Fisher Scientific) was used.
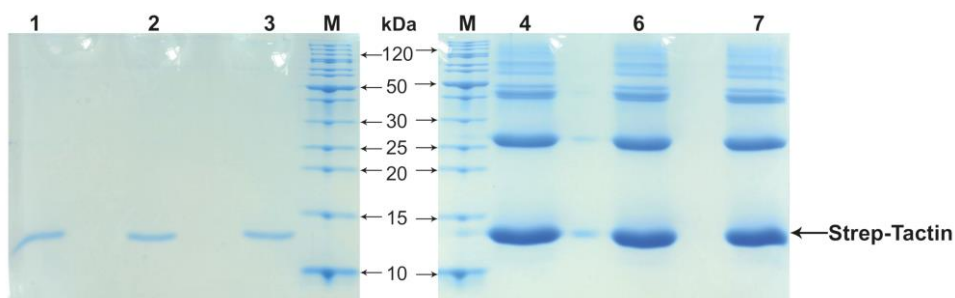
Supplemental Figure S7. Quantification of Strep-Tactin bound to different magnetic matrices. Strep-Tactin bound to in-house-coated 200-nm beads (left panel, lanes 1 to 3) and commercial 1-µm beads (right panel, lanes 4 to 6) was quantified. The same volumes of beads were treated with SDS loading buffer, incubated at 95°C, and samples were analyzed by SDS gel electrophoresis. The marked bands (Strep-Tactin monomer has a size of about 14 kDa) were quantified by analyzing BSA standards of defined concentrations on the same gel (not shown). On average 0.709 ± 0.027 µg Strep-Tactin was bound on the 200-nm beads (250 µl, 2.3% (v/v) slurry), whereas the 1-µm beads contained almost 20 times more Strep-Tactin (16.09 ± 0.32 µg in 115 µl 5% (v/v) slurry). For the marker (M) the PageRuler Unstained Protein Ladder (Thermo Fisher Scientific) was used.
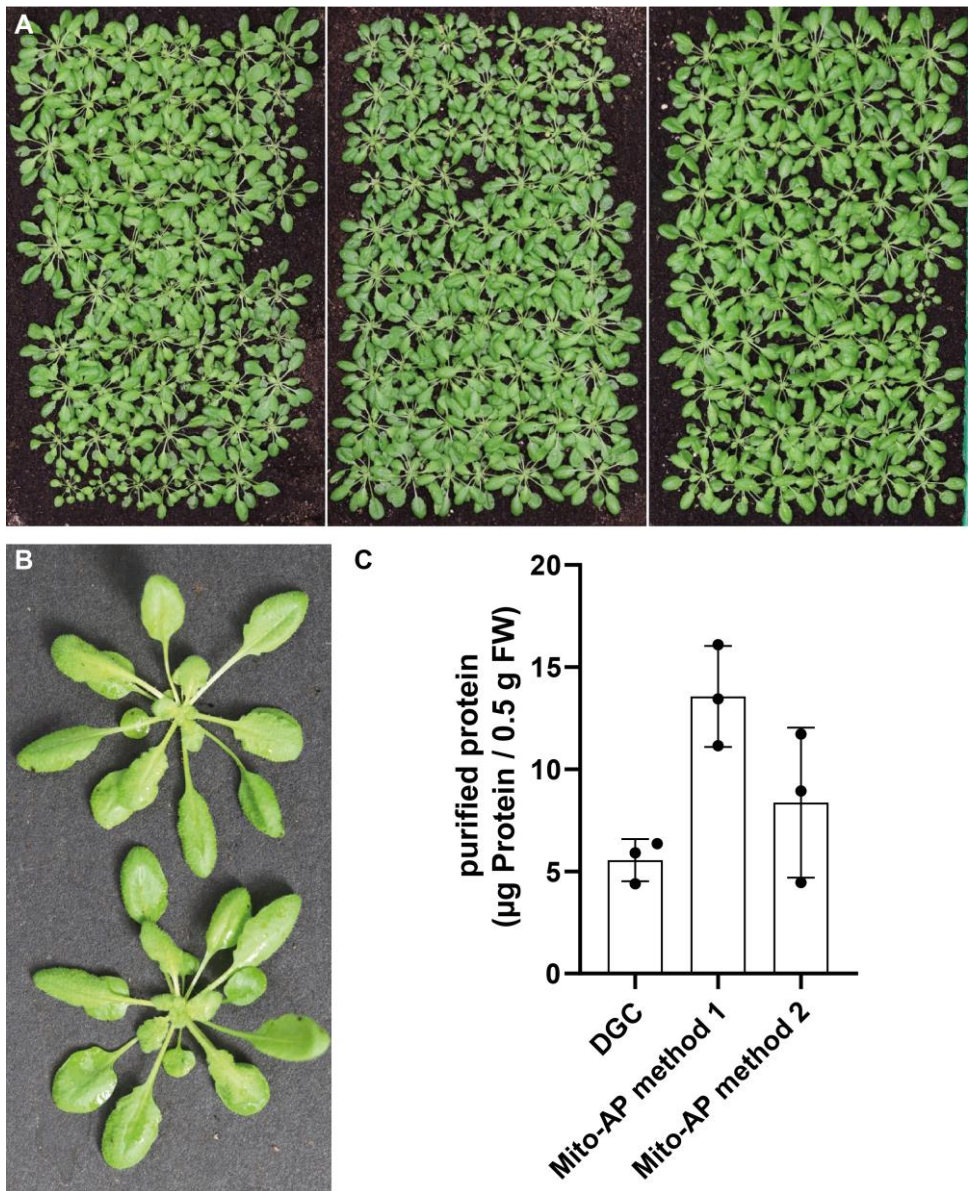
Supplemental Figure S8. Plant material needed for DGC and Mito-AP and relative mitochondrial protein yield obtained by both methods. *A. thaliana* plants were grown under short day conditions in a climate chamber for 5 weeks. (A) For the differential and density gradient centrifugation (DGC) around 110 g of plant material was used. (B) For Mito-AP (method 1 or method 2) 500 mg of plant material was used. (C) The total yield of protein referred to plant fresh weight input (FW) for all methods. Error bars are SD (n = 3 biological replicates)

method 2   15 min elution   BSA standards
+ Proteinase K   - Proteinase K   1 µg   5 µg

Supplemental Figure S9. Quantification of the total protein yield obtained from elution with and without proteinase K. 35S:Twin-Strep-tag:GFP:ADAPTER *A. thaliana* plants were subjected to either Mito-AP with proteinase K (method 2) or to the identical protocol but with omission of proteinase K and an extension of the time for elution to 15 min to match the incubation time of proteinase K (15 min elution). The resulting samples were analyzed by SDS gel electrophoresis. The observed band (stained with colloidal Coomassie) represents the total protein in the sample. These were quantified with a Li-Cor Odyssee FC and the total protein amounts were calculated by comparison with a BSA standard run on the same gel. 3.39 µg protein were present in lane 1 (method 2) and 1.97 µg protein were detected in lane 2 (15 min elution).
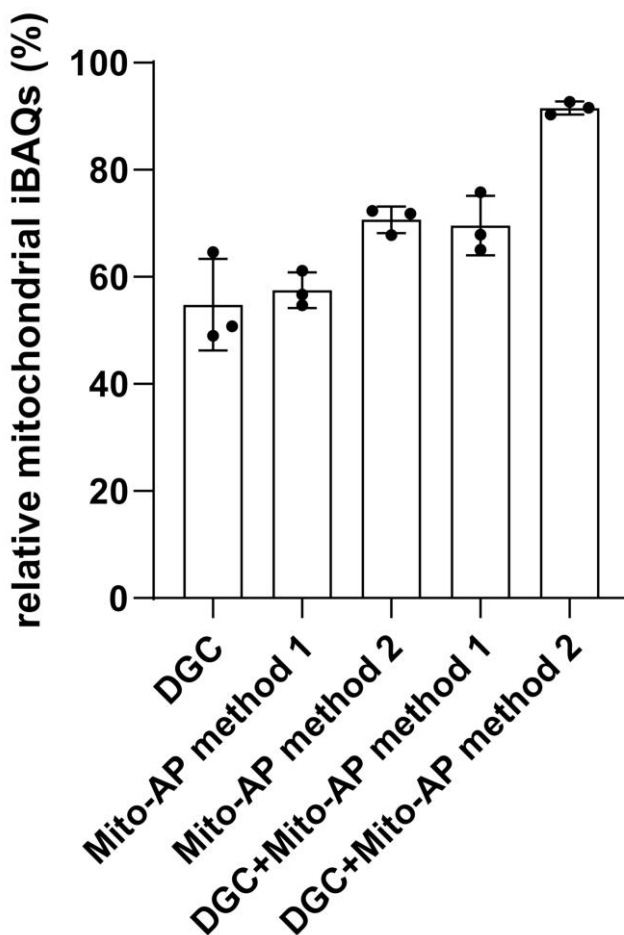
Supplemental Figure S10. Variation of mitochondrial purity from biological repeats processed by DGC or Mito-AP. The data for the mitochondrial category as described in Figure 5 is shown. Each of

119

the three replicates is a data point. Biological repeats were performed on consecutive days. Error bars are SD.
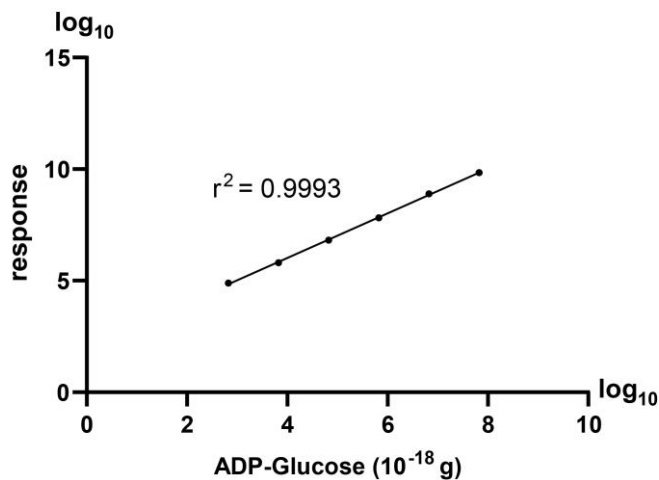

Supplemental Figure S11. ADP-glucose standard curve. ADP-glucose dissolved in different amounts in water were measured via LC-MS. The metabolite input in attogram was logarithmized. Linear regression was performed on non-logarithmized values with a response curve showing a linearity with $r^2 = 0.99$. The lowest detected amount of 0.67 fg on column was measured with a signal to noise ratio greater than 6.

Supplemental Figure S12. Contamination with plastidic fragments. Representative confocal microscopy overview pictures of 200-nm beads coated with Strep-Tactin bound to mitochondria derived from leaves of *A. thaliana* plants carrying the 35S:Twin-Strep-tag:GFP:ADAPTER construct. The mitochondria were purified using Mito-AP method 1. Bound mitochondria possess GFP fluorescence (GFP, left panel). Signals consistent with chloroplast autofluorescence (middle panel) and the overlay of both channels (right panel) are shown. Scale bars, 100 µm. Insert, higher magnification of several mitochondria bound to beads. Scale bar, 25 µm.
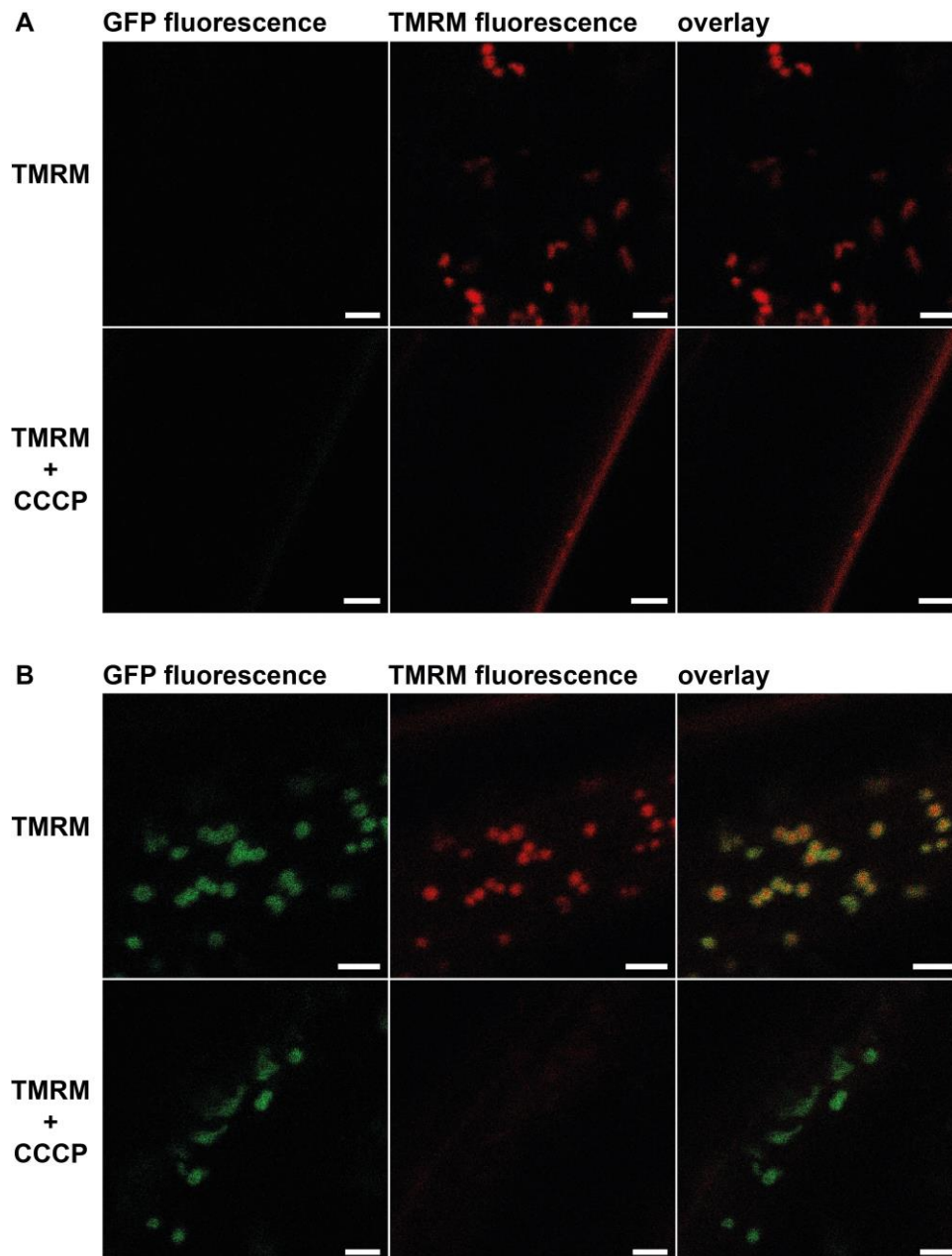
Supplemental Figure S13. Visualization of the mitochondrial membrane potential *in vivo*. Confocal fluorescence microscopy images of the root tissue of 12-day-old seedlings from a shaking culture. Seedlings were treated either with tetramethyl rhodamine methyl ester (TMRM) (upper panels) or a combination of TMRM and carbonyl cyanide 3-chlorophenylhydrazone (CCCP) (lower panels). The cell wall shows unspecific signal in the TMRM channel. (A) Wild type Col-0 and (B) plants expressing the 35S:Twin-Strep-tag:GFP:ADAPTER construct. In case of the 35S:Twin-Strep-tag:GFP:ADAPTER plants, a ring shaped structure surrounding the TMRM fluorescence signal was observed, indicating that this signal originates from mitochondria. Scale bars, 3 µm.
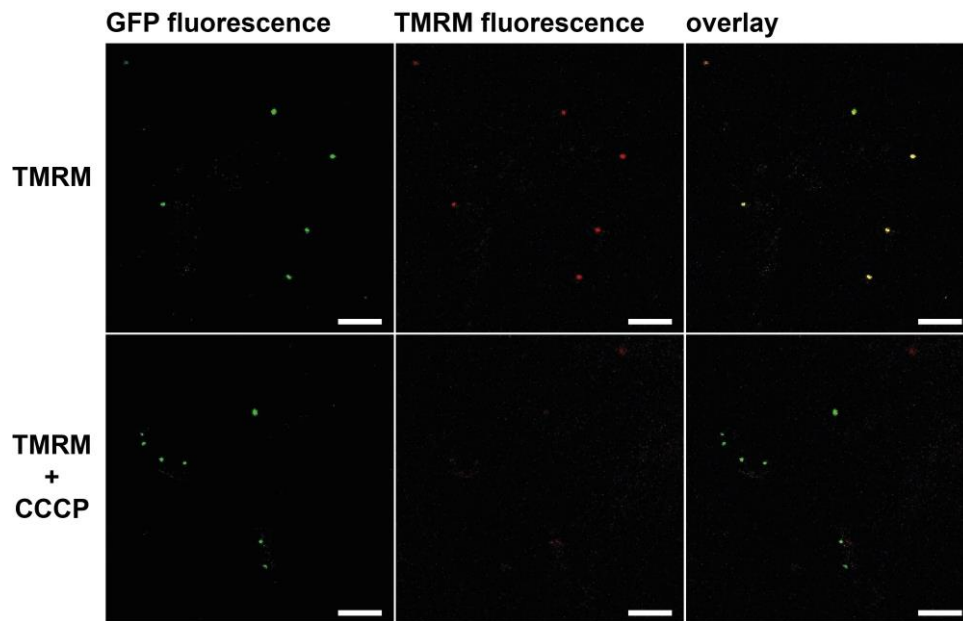
Supplemental Figure S14. Visualization of the mitochondrial membrane potential *in vitro.* Confocal fluorescence microscopy images of mitochondria bound to magnetic beads isolated from plants expressing the 35S:Twin-Strep-tag:GFP:ADAPTER construct. Isolated mitochondria were resuspended in MAS buffer containing 10 mM succinate and treated either with TMRM (upper panels) or a combination of TMRM and CCCP (lower panels). In total 219 GFP-fluorescent particles were counted for the TMRM treatment of which 160 showed a co-localization with the TMRM-fluorescence (73%). For the TMRM + CCCP treatment 183 GFP-fluorescent particles were counted of which 58 showed a co-localization with a weaker TMRM-fluorescence (31%). Scale bars, 10 μm.
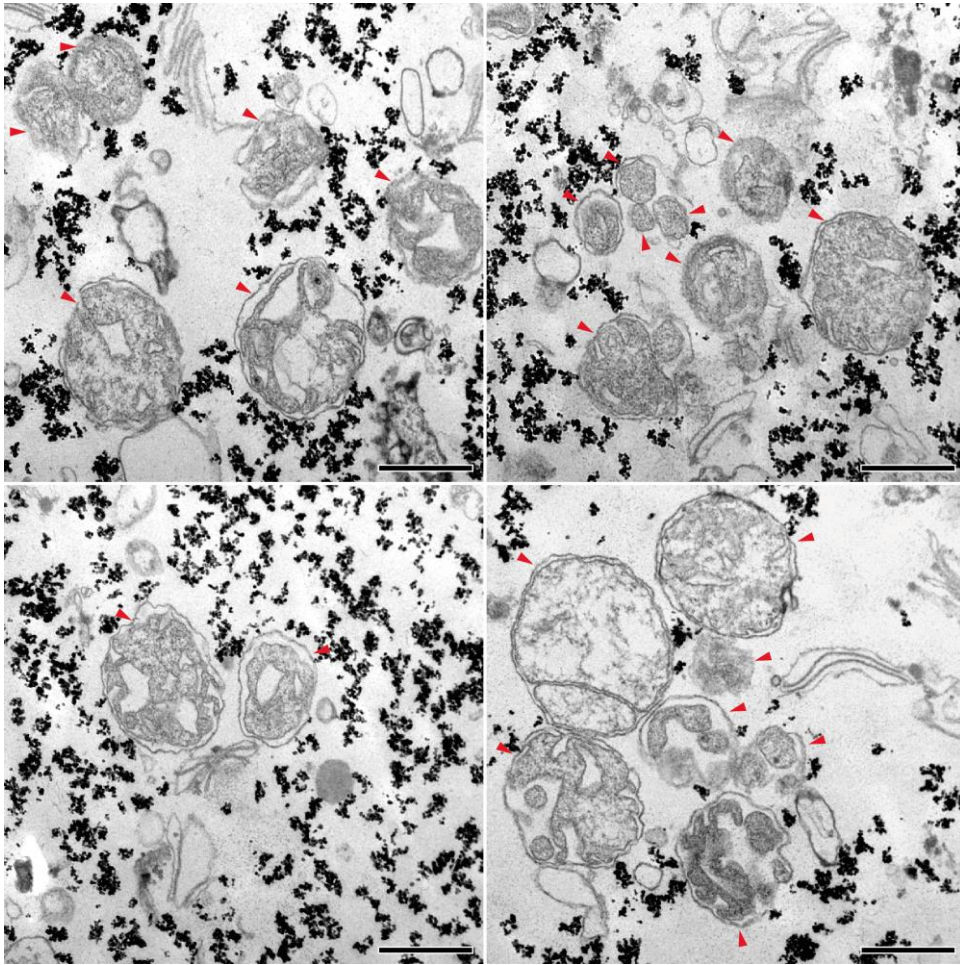
Supplemental Figure 15. Electron micrographs of isolated mitochondria. Representative transmission electron microscopy (TEM) images of mitochondria bound to magnetic beads isolated from plants expressing the 35S:Twin-Strep-tag:GFP:ADAPTER construct. Small dense particles are the iron oxide core of the magnetic beads. Mitochondria are indicated by red arrowheads. Scale bars, 400 nm.

Supplemental Table S4. Primer sequences

name	sequence
P1023	TTTGGTCTCAACATTTCGGCTGCCTTGTCGGATAAGTTAG
P1024	TTTGGTCTCAATCTTCGGACATTGTTGTGGGAG
P1025	TTTGGTCTCAAGATCAAATGGTTGCTCTAGTCGGCGGA
P1026	TTTGGTCTCAACAAAAGCCTAACTGTTCTTCCTTGCAGAG
P626	TTGAAGACAAAATGGAGTCTGATGAGTCT
P820	TTGAAGACAACGAACCTTCCTCACCAGCATC

Supplemental Table S5. MRM parameters used for the detection of polar metabolites with a triple quadrupole mass spectrometer. The precursor ion, the product ions (used for quantification, quant.; used for qualification, qual.), the fragmentor, and the collision energy (CV, in volt) is provided for every measured metabolite.

	Precursor Ion [M+H]$^+$ (m/z)	Product Ion (m/z)	Fragmentor	Collision Energy (V)
ADP	428	136 (quant.)	132	25
ADP-Glucose	590,1	136 (quant.)	105	49
		428 (qual.)	105	13
AMP	348,1	136 (quant.)	132	25
ATP	508	136 (quant.)	120	45
		410 (qual.)	120	15
NMN	335,07	123 (quant.)	101	13
		97 (qual.)	101	33
NAD$^+$	664,1	136 (quant)	144	48
		542,1 (qual.)	144	17
NADH	666,13	136 (quant.)	155	58
NADP	745,1	136 (quant.)	159	64

2.5 The nucleotide metabolome of germinating Arabidopsis thaliana reveals a central role of thymidine for chloroplast development

Markus Niehaus[1], Henryk Straube[1], André Specht[1], Chiara Baccolini[2], Claus-Peter Witte[1] and Marco Herde[1]

[1] Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany
[2] Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam, Germany

The nucleotide metabolome of germinating *Arabidopsis thaliana* seeds reveals a central role for thymidine phosphorylation in chloroplast development

Research Article

Markus Niehaus [iD] ,[1] **Henryk Straube** [iD] ,[1] **André Specht** [iD] ,[1] **Chiara Baccolini** [iD] ,[1,†]
Claus-Peter Witte [iD] [1] **and Marco Herde** [iD] [1,*]

1   Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany

*Author for correspondence: mherde@pflern.uni-hannover.de
†Present address: Max Planck Institute of Molecular Plant Physiology, Potsdam, 14476, Germany.
M.N. and M.H. designed the study. M.N., H.S., C.B., and A.S. acquired the experimental data. M.N. and H.S. analyzed the data and M.N., H.S., C.P.W., and M.H. interpreted the data. M.N., H.S., C.P.W., and M.H. wrote the manuscript. All authors read and revised the manuscript and agreed on the final version.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell) is Marco Herde (mherde@pflern.uni-hannover.de).

Abstract

Thymidylates are generated by several partially overlapping metabolic pathways in different subcellular locations. This interconnectedness complicates an understanding of how thymidylates are formed in vivo. Analyzing a comprehensive collection of mutants and double mutants on the phenotypic and metabolic level, we report the effect of de novo thymidylate synthesis, salvage of thymidine, and conversion of cytidylates to thymidylates on thymidylate homeostasis during seed germination and seedling establishment in Arabidopsis (*Arabidopsis thaliana*). During germination, the salvage of thymidine in organelles contributes predominantly to the thymidylate pools and a mutant lacking organellar (mitochondrial and plastidic) thymidine kinase has severely altered deoxyribonucleotide levels, less chloroplast DNA, and chlorotic cotyledons. This phenotype is aggravated when mitochondrial thymidylate de novo synthesis is additionally compromised. We also discovered an organellar deoxyuridine-triphosphate pyrophosphatase and show that its main function is not thymidylate synthesis but probably the removal of noncanonical nucleotide triphosphates. Interestingly, cytosolic thymidylate synthesis can only compensate defective organellar thymidine salvage in seedlings but not during germination. This study provides a comprehensive insight into the nucleotide metabolome of germinating seeds and demonstrates the unique role of enzymes that seem redundant at first glance.

Introduction

The molecular processes of synthesis, degradation, and salvage of deoxyribonucleotides (dNTs; relevant abbreviations are also listed in Supplemental Table S1), the building blocks

of DNA, are central to the maintenance of genetic information. In two ways dNTs are crucial for efficient and faithful DNA replication: (1) the absolute concentration of dNT

IN A NUTSHELL

Background: All living organisms, including plants, possess DNA and rely on the active synthesis of its building blocks, called deoxyribonucleotides (dNTs; A, G, C, and T). DNA is made not only in the nucleus but also in chloroplasts and mitochondria; thus, dNTs must be either made in or transported to these organelles. Several alternative pathways transport and synthesize one of these nucleotides (T) for the nucleus and the organelles. These processes are important during seed germination and seedling establishment since a dry seed does not contain any dNTs. Thus, they have to be made prior to DNA replication.

Question: What is the concentration of dNTs during germination and seedling establishment? What is the importance of every alternative pathway that results in the formation of the "T" nucleotide? What are the consequences of lacking "T" for DNA synthesis in the nucleus and the organelles during germination and seedling establishment?

Findings: The concentration of nucleotides in germinating Arabidopsis seeds is constantly rising. The ratio of the different nucleotides seems to be first tailored to support DNA synthesis within chloroplasts. Analyzing a collection of mutants which are impaired in processes for the formation of "T", we found that the phosphorylation of thymidine within organelles is the most important process. Plants that are unable to phosphorylate thymidine have a reduced amount of DNA in chloroplasts and seedlings are pale. Interestingly, this effect is weaker in established seedlings and we assume that transport processes that change during seedling establishment are responsible for this phenomenon.

Next steps: Although we found thymidine to be a crucial metabolite for "T" formation we do not know which process generates it. We think that the transport of precursors for "T" over the plastidic membrane changes during seedling establishment, and our aim now is to identify transporters and characterize the process further.

triphosphates (dNTPs) and the concentration ratios of dNTPs to each other affect the mutation rate and the rate of DNA synthesis (Kohalmi et al., 1991; Buckland et al., 2014; Le Ret et al., 2018) and (2) the abundance of noncanonical dNTPs that arise from enzymatic or nonenzymatic processes can cause mutations. Thus, sanitizing dNT pools prevents DNA damage (Nagy et al., 2014). In some cases, the effects of purity and size of dNTP pools on the DNA are tightly interconnected. In mammals, the usage of the noncanonical deoxyuridine triphosphate (dUTP) by DNA polymerases depends on the size of the deoxythymidine triphosphate (dTTP) pool; thus, enzymes involved in dTTP synthesis influence its pool size and thereby modulate the effect of noncanonical dNTPs (Anderson et al., 2011; Schmidt et al., 2019; Martínez-Arribas et al., 2020). Some enzymes are hypothesized to contribute to both dNTP purity and pool size. For example, the enzyme dephosphorylating dUTP (dUTP PYROPHOSPHATASE 1 [DUT1]) has a role in sanitizing the dNTP pool (Dubois et al., 2011) but also a function in dTTP synthesis, since the product of this reaction (deoxyuridine monophosphate, [dUMP]) is also a precursor of dTTP (Guillet et al., 2006; Vértessy and Tóth, 2009).

For reproduction, many plants form dormant seeds, which upon stimulus germinate to reestablish a plant. Seed vigor is of high economic importance, as it determines successful germination and crop yield (Rajjou et al., 2012; Waterworth et al., 2015). In the dormant seed, nucleic acids are damaged and need to be repaired during germination (Waterworth et al., 2015). Much research has focused on how RNA and DNA are repaired during germination, but little is known

about the metabolism of nucleic acid precursors, besides its importance for proper seedling development (Ashihara et al., 2020). Furthermore, dNTPs are mostly discussed as fueling DNA repair, but little attention has been paid to the aspect that imbalanced or too low dNTP concentrations cause faulty DNA replication and mutations (Buckland et al., 2014; Pedroza-García et al., 2019). Dormant seeds of Arabidopsis (*Arabidopsis thaliana*) and wheat (*Triticum aestivum*) have been shown not to contain dNTPs (Schimpff et al., 1978; Castroviejo et al., 1979; Straube et al., 2021a). Therefore, dNTPs must be synthesized during germination, consistent with a strong rise of dNTP concentrations after imbibition. These dNTP concentrations quickly reach a plateau, coinciding with the onset of DNA synthesis (Castroviejo et al., 1979).

DNA synthesis not only occurs in the nucleus, but also in mitochondria and chloroplasts. The nuclear, mitochondrial, and plastidic replication machineries need access to dNTPs, either requiring the presence of enzymes synthesizing dNTPs at each of these cellular locations or transport of dNTPs. Furthermore, for the on-site synthesis of dNTPs, transport processes are required, namely of precursors like deoxyribonucleosides (dNs) or dNT monophosphates or diphosphates (dNMPs, dNDPs). Organellar (i.e. chloroplastic and mitochondrial) dNTPs are currently thought to be supplied by a combination of transport and organellar synthesis where different transport processes, as well as distinct biosynthetic routes, are functionally redundant. This model is supported by a lack of phenotypic alterations in mutant plants lacking only a single metabolic route (Gorelova et al., 2017;

Pedroza-García et al., 2019). Because most of these single knockouts have no altered phenotype, in this study, we assessed the influence of the different metabolic processes that contribute to dNT formation and especially the dTTP pool in plants on a molecular level, analyzing the dNT metabolome as a more sensitive readout.

It is not only necessary to ensure the dNTP supply, but the dNTP pool sizes in the cytosol and the organelles must be tailored to the different demands of each replication machinery. In mammalian cells, dNTPs are more abundant at the onset of DNA synthesis in the S-phase, suggesting a co-ordination of dNTP supply and demand (Franzolin et al., 2013; Stillman, 2013). In germinating seeds, the replication of the different genomes is asynchronous. Strong replication of the plastid genome precedes the mitochondrial and nuclear replication for at least 24 h (Barroco et al., 2005; Paszkiewicz et al., 2017). Therefore, the demand for dNTPs also differs between the nucleus, mitochondria, and chloroplasts during germination.

All dNTs can be synthesized from the respective ribonucleotide diphosphates (NDPs) by the action of a RIBONUCLEOTIDE REDUCTASE (RNR) localized in the cytosol (Lincker et al., 2004; Wang and Liu, 2006). The dNDPs require phosphorylation by NUCLEOTIDE DIPHOSPHATE KINASES (NDKs) to form dNTPs. However, for the synthesis of thymidylates, the RNR generates deoxyuridine diphosphate (dUDP) (Zrenner et al., 2006; Witte and Herde, 2020), which is dephosphorylated to dUMP. Then dUMP is methylated to deoxythymidine monophosphate (dTMP) before being phosphorylated to dTTP. The transfer of the methyl moiety to dUMP is catalyzed by DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE (DHFR-TS). In Arabidopsis, three isoforms of DHFR-TS exist; DHFR-TS1 is located in the cytosol, DHFR-TS2 is present exclusively in mitochondria while the third isoform probably has no catalytic function (Gorelova et al., 2017; Corral et al., 2018). Mutants lacking single DHFR-TS isoforms are phenotypically normal, suggesting that they are functionally redundant and that thymidylates (dTMP, deoxythymidine diphosphate [dTDP], or dTTP) can be transported across the mitochondrial membrane. Combining *dhfr-ts1* and *dhfr-ts2* loss-of-function mutants results in albino seeds occurring at a frequency supporting Mendelian inheritance, which emphasizes the crucial role for both isoforms in embryo development (Gorelova et al., 2017). Interestingly, the molecular basis for the conversion of the RNR product dUDP to the DHFR-TS substrate dUMP is unknown. Two alternative scenarios for this conversion can be envisioned. Either dUDP is dephosphorylated by a so far unknown dUDP phosphatase or dUDP is first phosphorylated to dUTP by a nucleotide diphosphate kinase (Zrenner et al., 2006; Witte and Herde, 2020) and subsequently converted to dUMP by the pyrophosphatase DUT1. A direct dephosphorylation of dUDP by DUT1 in plants seems unlikely, because DUT1 belongs to a protein family that is generally unable to dephosphorylate dNDPs (Vértessy and Tóth, 2009). In contrast to other

metabolic routes, the formation of dTMP by RNR and DHFR-TS does not require dNs or other dNTs; thus, we refer to this pathway here as de novo T synthesis.

A fine-tuning of deoxycytidilate and thymidylate pools was observed in rice (*Oryza sativa*), which has an endogenous DEOXYCYTIDINE MONOPHOSPHATE (dCMP) DEAMINASE (DCD) that converts dCMP to dUMP, the substrate for dTMP synthesis by DHFR-TS (Xu et al., 2014). The protein is exclusively localized in mitochondria and a mutation abolishing enzyme function causes a strong increase in deoxycytidine triphosphate (dCTP) and a moderate reduction of dTTP levels concomitant with chlorosis and white leaf areas in young and in fully grown plants (Xu et al., 2014; Niu et al., 2017). A homolog of DCD was identified in Arabidopsis but no altered phenotype was observed in the respective loss-of-function mutant and the subcellular localization is unknown (Niu et al., 2017).

The formation of dTMP by several processes is followed by phosphorylation of dTMP to dTDP catalyzed by dTMP KINASE (TMPK, previously referred to as ZEU1; Ronceret et al., 2008). The *TMPK* gene generates transcripts that contain a long and a short open reading frame (ORF) encoding two proteins localizing in mitochondria and the cytosol, respectively (Ronceret et al., 2008). A possible dual localization of TMPK not only in mitochondria but also in chloroplasts has not been investigated yet since localization experiments were performed in onion (*Allium cepa*) peels, an experimental system in which plastidic localization can be overlooked (Ronceret et al., 2008; Osaki and Kodama, 2017).

Most dNs can be salvaged by DEOXYNUCLEOSIDE KINASE phosphorylating deoxyguanosine (dG), deoxyadenosine (dA), and deoxycytidine (dC) while for thymidine (dT) a separate enzyme, THYMIDINE KINASE (TK1), is required, which can also phosphorylate deoxyuridine (dU; Clausen et al., 2012). One TK1 isoform (TK1a) is located in the cytosol (Xu et al., 2015) while the second isoform (TK1b) is present in mitochondria and chloroplasts (Xu et al., 2015; Pedroza-García et al., 2019). In contrast to *tk1a*, *tk1b* mutant seedlings have chlorotic cotyledons but later develop normally (Clausen et al., 2012; Pedroza-García et al., 2019). Interestingly, a *tk1a tk1b* double mutant shows an albino phenotype and development is arrested at the four-leaf stage (Clausen et al., 2012; Pedroza-García et al., 2019), suggesting that de novo T synthesis alone is not sufficient to sustain DNA synthesis but that the salvage of dT is required to support the thymidylate pools. It has been reported that this double mutant can be rescued with carbohydrates in the media, which was attributed to a stimulation of de novo T synthesis (Pedroza-García et al., 2019). In addition, we recently demonstrated that dT can also be degraded in Arabidopsis by NUCLEOSIDE HYDROLASE 1 (NSH1; Straube et al., 2021a), influencing thymidine and thymidylate pools.

The expression profiles of genes potentially involved in thymidylate formation (*TK1a*, *TK1b*, *DHFR-TS1*, *DHFR-TS2*, *DCD*, *DUT1*, and *TMPK*) were investigated during germination and seedling establishment (Narsai et al., 2011). While the
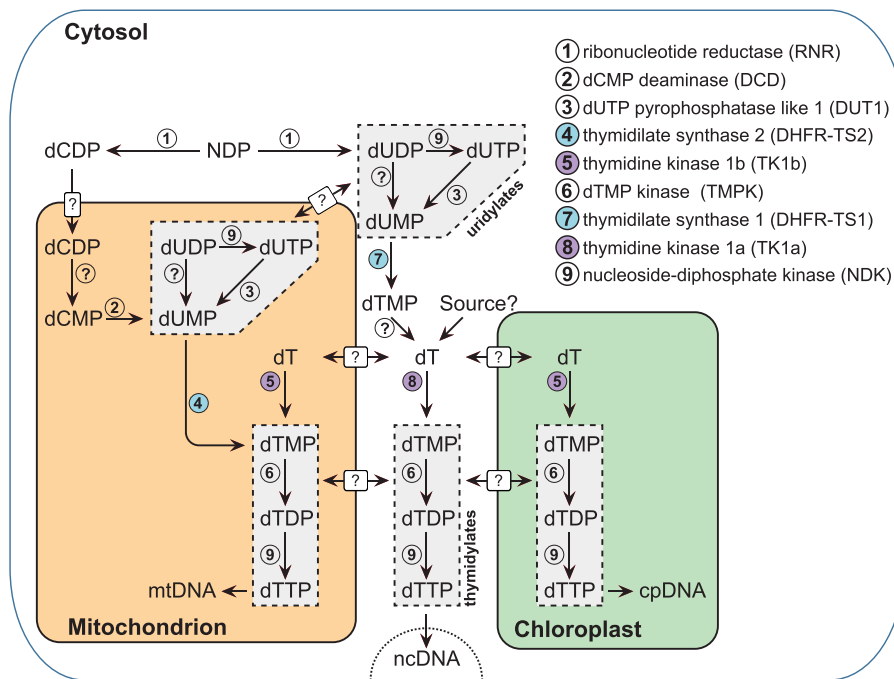
Figure 1 Proposed model of thymidylate metabolism in *A. thaliana* during germination and seedlings establishment. Enzymes putatively involved in the formation of thymidylate pools and their subcellular localization are shown. Question marks indicate currently unknown enzymes or transporters allowing an exchange of nucleosides and nucleotides between the cytosol, the chloroplast, and the mitochondrion. Enzymes of the de novo thymidylate synthesis pathway are colored in blue while enzymes of the salvage pathway are colored in purple. The localization of DUT1 in plastids is not shown. Deoxycytidine diphosphate (dCDP); thymidine (dT); deoxythymidine monophosphate (dTMP).

transcript of *TK1b* is abundant and present already in dry seeds, transcripts of all other genes are increased rapidly upon transfer to long-day growth conditions (Supplemental Figure S1). This observation suggests that the enzymes hypothesized to contribute to thymidylate formation (Figure 1) play a role during germination and seedling establishment.

Both de novo T synthesis of dNTPs via RNR (reductive pathway) and the salvage of dNs seem to be important during germination, with salvage activity preceding the activity of RNR in time. Several studies concerning the DNA synthesis during germination used the application of labeled dT, or sometimes dA and dC, and their incorporation into DNA as a readout for DNA synthesis (Castroviejo et al., 1979; Dellaquila et al., 1980; Strugala and Buchowicz, 1984; Marciniak et al., 1987; Stasolla et al., 2002). The reductive pathway was assessed with labeled cytidine and uridine whose incorporation into DNA requires the action of RNR (Schimpff et al., 1978; Stasolla et al., 2002). Soon after imbibition, salvage activity assessed by incorporation of dNs into DNA was already active, whereas activity of the reductive pathway was only observed at later stages of germination and depended on seed age (Schimpff et al., 1978). Fresh wheat (*T. aestivum*) seeds contained an unknown compound inhibiting the RNR, but dN incorporation (salvage) was not affected by this inhibitor (Schimpff et al., 1978; Baumann et al., 1984). Further evidence that the salvage pathway is active early on is derived from experiments utilizing hydroxyurea, an inhibitor of the RNR, showing a steady incorporation of labeled dT in the first hours of germination (Thornton et al., 1993). These

experiments all rely on the exogenous feeding of nucleosides and demonstrate that germinating seeds have the capacity to incorporate dNs by salvage. However, it is not known whether salvage actually takes place under conditions where pool sizes are not manipulated. The contribution of salvage to the dNT pools in vivo remains unclear. Furthermore, little is known about the amounts and origins of dNs in dormant or germinating seeds (Straube et al., 2021a). Hitherto it was speculated that dNs are derived from DNA degradation or dNTP dephosphorylation, but no evidence has been presented so far (Bryant, 1980).

In this study, we monitored the nucleotide and nucleoside metabolome in germinating seeds and demonstrate that dNTs are synthesized very early in germination, long before the onset of DNA synthesis. Furthermore, we created and analyzed a collection of mutants abolished in enzymatic activities of all known processes for thymidylate formation. We conclude that thymidine salvage in organelles contributes predominantly in germinating seeds to the thymidylate pool and has a strong impact on other dNT pools as well as on chloroplast DNA (cpDNA) synthesis during early seedling establishment.

Results

An alternative ORF encodes a DUT1 version that localizes in mitochondria and chloroplasts

To investigate the impact of the different metabolic routes on the formation of the thymidylate pool, a collection of
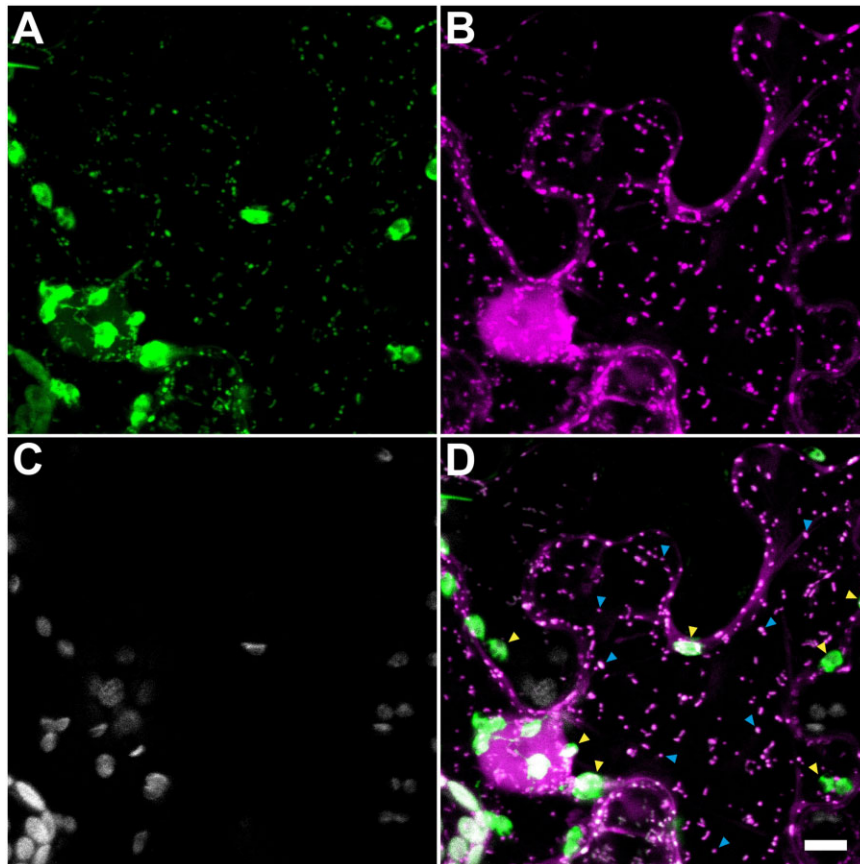
Figure 2 Subcellular localization of DUT1 translated from a full-length transcript. Confocal microscopy images of a *N. benthamiana* leaf transiently expressing DUT1 as C-terminal-tagged eYFP fusion protein (Pro$_{35S}$:*DUT1:eYFP*) from a longer transcript than previously annotated containing 5'-prolonged ORF encoding 41 additional amino acids at the N-terminal end. A mitochondrial marker protein (*Sc*COXIV:mCherry) was co-expressed. Images were taken 3 days after agrobacteria infiltration. A, eYFP:DUT1, B, mitochondrial marker Sc.COXIV:mCherry, C, chloroplast autofluorescence, and D, overlay of A, B, and C. Exemplary, selected mitochondria, and chloroplasts showing dual localization are marked with blue and yellow arrowheads, respectively. Scale bars = 10 μm.

mutants lacking these processes was created. Mutant plants lacking steps of salvage, C-to-T conversion, and de novo T synthesis (Figure 1) located in cytosol and organelles, were available from previous studies (Clausen et al., 2012; Gorelova et al., 2017; Niu et al., 2017). However, DUT1, which may be involved in the formation of dTMP via the generation of the DHFR-TS substrate dUMP (Figure 1), is less characterized and a complete loss-of-function mutant of *DUT1* is likely not viable (Dubois et al., 2011). We observed the presence of several RNA-seq reads mapping upstream of the 5'-UTR (untranslated region) predicted by The Arabidopsis Information Resource, version 10, suggesting the presence of an alternative transcription start site. This longer transcript contains an extended ORF that encodes the original protein plus 41 additional amino acids at the N-terminus that are predicted to constitute a transit peptide. The coding sequence (CDS) of enhanced yellow fluorescent protein (eYFP) was fused to the 3'-end of the extended ORF and the construct was transiently expressed under the control of a constitutive promoter (Pro$_{35S}$) in *Nicotiana benthamiana* leaves. A signal consistent with cytosolic and nuclear as well as mitochondrial and plastidic localization of

the fusion protein was detected with confocal microscopy (Figure 2A). The mitochondrial and plastidic localization was corroborated by an overlay with the signal of a mitochondrial marker (yeast COXIV transit peptide fused to mCherry fluorescent protein; Nelson et al., 2007) and the chloroplast autofluorescence, respectively (Figure 2, B–D).

Since the mitochondrial localized DUT1 might alone provide the substrate for DHFR-TS2, whose loss-of-function does not lead to phenotypic alterations (Gorelova et al., 2017), we speculated that a mutant abolished in organellar DUT1 function is viable. We used a CRISPR (clustered regularly interspaced short palindromic repeats) approach to introduce a lesion between the first and second start codon of the respective ORFs (Supplemental Figure S2, A and B) that results in exclusive disruption of the ORF encoding the organellar localized DUT1. A homozygous *dut1$_{org}$* mutant lacking the T-DNA used for its creation was isolated and is phenotypically inconspicuous (Supplemental Figure S2C). DNA from the *dut1$_{org}$* mutant was used as a template to amplify the full-length CDS including the genetic lesion. A construct encoding this mutated variant of DUT1 fused to eYFP was generated. In contrast to the wild-type fusion

protein, the mutant variant was exclusively found in the cytosol/nucleus (Supplemental Figure S2D), demonstrating that our approach to disrupt only the ORF encoding the organellar localized DUT1 was successful.

It has been reported that the subcellular localization of DHFR-TS may vary depending on the tissue (Gorelova et al., 2017). Therefore, we determined the localization of all organellar proteins involved in the formation of the thymidylate pool to demonstrate their subcellular location in an establishing seedling. The subcellular localization of DCD, DHFR-TS2, and TK1b in roots and cotyledons of germinating seeds (2 days after transfer to long-day growth conditions) was assessed using transgenic plants constitutively expressing the respective eYFP fusion proteins (Supplemental Figures S3 and S4). In line with previous studies, DHFR-TS2 was exclusively observed in mitochondria while TK1b was localized in mitochondria and chloroplasts (Supplemental Figures S3, A–D and S4, B and C; Gorelova et al., 2017; Corral et al., 2018; Pedroza-García et al., 2019). Consistent with results for the DCD in rice, we also observed a mitochondrial localization of the homolog in Arabidopsis (Supplemental Figures S3, E and F and S4A). We found that TMPK is present in chloroplasts in addition to the already described localization in mitochondria (Ronceret et al., 2008) by expressing the TMPK:eYFP fusion protein in *N. benthamiana* (Supplemental Figures S3G and S4E).

Only the disruption of organellar salvage and de novo T synthesis alters the phenotype of seedlings

While a complete blockage of either de novo T synthesis or salvage of dT in the cytosol and the organelles is lethal (*tk1a tk1b* or *dhfr-ts1 dhfr-ts2*), loss-of-function mutants of any single gene are phenotypically normal with the exception of the *tk1b* mutant whose cotyledons are chlorotic (Clausen et al., 2012; Xu et al., 2015; Gorelova et al., 2017; Pedroza-García et al., 2019). We performed a side-by-side comparison of seedlings grown on soil from all single mutants (*tk1a* cytosolic salvage; *tk1b* salvage in organelles, *dhfr-ts1* cytosolic de novo T synthesis, *dhfr-ts2* mitochondrial de novo T synthesis; *dut1_org* putative function in de novo T synthesis in mitochondria; *dcd* C-to-T conversion in mitochondria). We photographed individual seedlings (Figure 3A) and quantified the green and yellow pixels (Figure 3B) in the respective micrographs from different time points to gain quantitative insight into the chlorosis previously observed for the *tk1b* mutant. Consistent with the previous reports, only *tk1b* mutants displayed chlorosis, further emphasizing the importance of dT salvage in mitochondria and chloroplasts for the fitness of the seedling. We also determined relative growth rates from 3 to 4 and 4 to 5 days after transfer to long-day growth conditions of plants from all genotypes and did not observe any significant differences between mutants and wild-type (Supplemental Figure S5).

We reasoned that the phenotype of the *tk1b* mutant is probably caused by increased thymidine and decreased thymidylate pools and that therefore visualizing the impact of

other processes such as the de novo T synthesis might be facilitated in the *tk1b* mutant background. Plants of this background were thus crossed with *dhfr-ts1*, *dhfr-ts2*, *dcd*, and *dut1_org* mutants, respectively. Interestingly, the *tk1b dhfr-ts2* double mutant displayed a markedly stronger chlorosis together with white areas on the cotyledons that are not observed in the *tk1b* single mutant (Figure 3A). While the cotyledons of the *tk1b* single mutant recover and accumulate more green pigment over time until they are indistinguishable from wild-type, the white areas in the *tk1b dhfr-ts2* double mutant do not recover (Supplemental Figure S6A). A pronounced fold difference between the cpDNA/nuclear DNA (ncDNA) ratios of wild-type and *tk1b* as well as *tk1b dhfr-ts2* mutants 2 days after transfer to long-day growth conditions is progressively reduced at later time points (6 and 10 days; Supplemental Figure S6B) concomitant with the visual recovery. At 10 days, the ratio of cpDNA/ncDNA in the *tk1b* mutant compared to the wild-type was not significantly different anymore while a small difference between wild-type and *tk1b dhfr-ts2* was still present (Supplemental Figure S6B). At early time points, dT salvage but not de novo T synthesis plays apparently the main role for supplying thymidylates for cpDNA synthesis. Interestingly, in dry seeds no significant differences in cpDNA can be observed, suggesting that TK1b and DHFR-TS2 are predominantly important during germination rather than embryo development (Supplemental Figure S6B). The synergistic effect of *tk1b* and *dhfr-ts2* on the chlorosis suggests that predominantly salvage in mitochondria and chloroplasts but also de novo T synthesis in mitochondria contribute to plant fitness. Rarely, but exclusively in the *tk1b dhfr-ts2* double mutant population, we also observed plants with completely albinotic sections (Supplemental Figure S6C) suggesting that abolished thymidine salvage and thymidylate de novo T synthesis together can sometimes also affect the phenotype of an adult plant. Overall, the data show that the lack of the organellar enzymes cannot be fully compensated by the cytosolic isoenzymes (TK1a and DHFR-TS1).

dNTs and dN pools increase rapidly during early seedling establishment

We hypothesized that the impact of *tk1b* and *tk1b dhfr-ts2* mutants on the phenotype of seedlings is the result of altered thymidine and thymidylate formation during germination. Since comprehensive data on the nucleotide metabolome of plants is scarce (Straube et al., 2021b, 2021a), our first objective was to monitor the (d)NT and (d)N pool formation during germination. We recently established a protocol to comprehensively monitor the nucleotide and nucleoside metabolome (Straube et al., 2021a) and used this technique to investigate germination and seedling establishment.

The treatment of the seeds was identical as described in transcriptome studies (Narsai et al., 2011; Law et al., 2014) and a study on nuclear, chloroplast, and mitochondria DNA
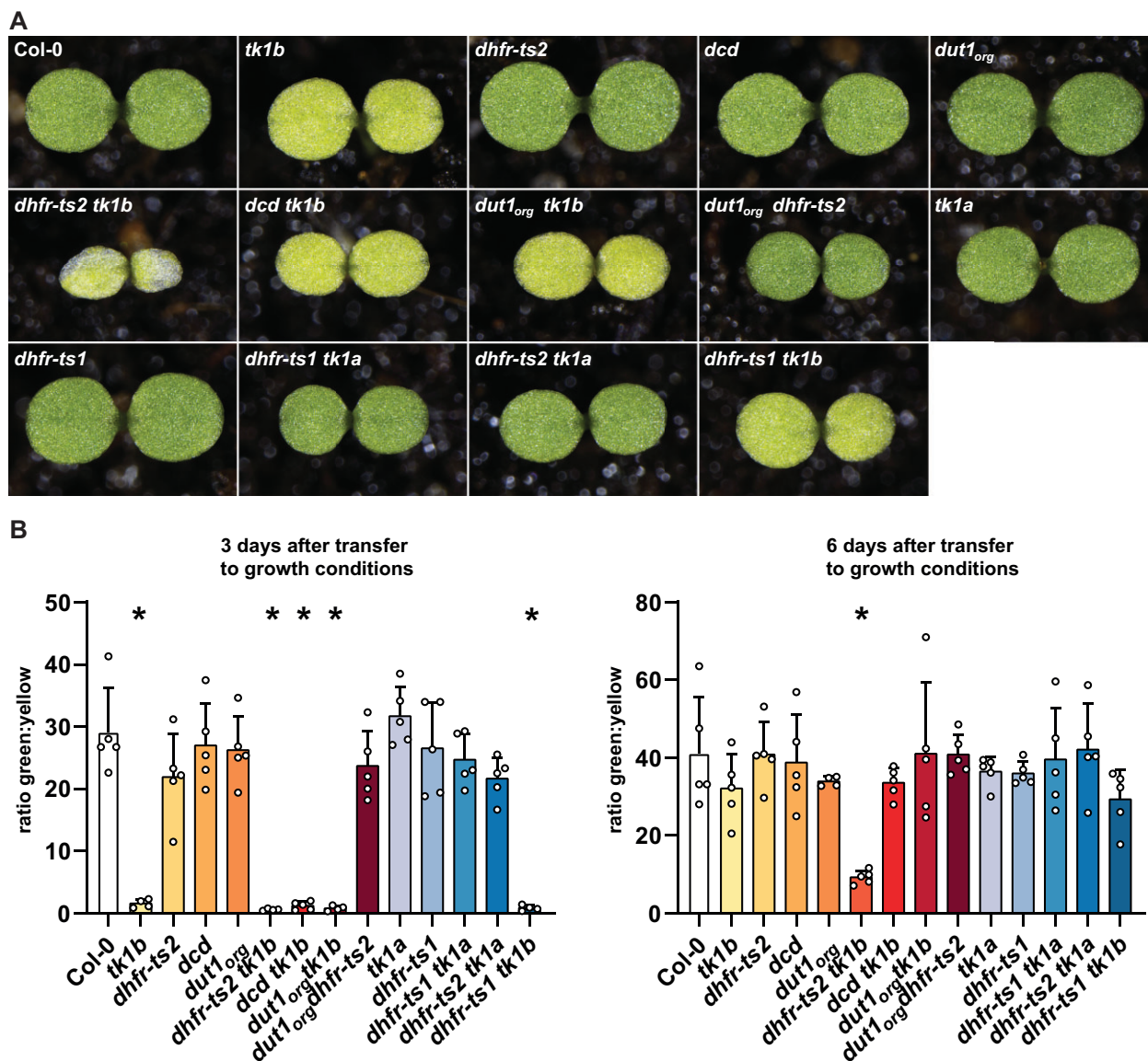
Figure 3 Phenotypes of wild-type and mutant seedlings lacking enzymes putatively involved in thymidylate homeostasis. Seedlings of all mutants and double mutants used in this study were grown on soil and photographed 3 days after transfer to long-day growth conditions. A, Images are representative for four to five individual plants. B, Quantification of the ratio between green and yellow pixels of each genetic variant. Images of seedlings 3 and 6 days after transfer to long-day growth conditions were analyzed. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Asterisks indicate significant differences ($P < 0.05$) to Col-0. Four to five biological replicates (individual plants) were analyzed. Error bars are SD. Adjusted P-values for multiple comparisons can be found in Supplemental Data Set 2.

replication (Paszkiewicz et al., 2017) to allow comparability. Seeds were imbibed at 4°C for 48 h in the dark and then transferred to long-day growth conditions (16-h light, 22°C) and monitored for 6 days (144 h). In our recent report (Straube et al., 2021a), we did not observe dT and dC in dry wild-type seeds, whereas small amounts of these nucleosides were quantified here (Figure 4B). This is likely due to an improved disruption of seed tissue and a refined metabolite separation in this study. While there was a slight increase of dNTP concentrations during the incubation at 4°C, dT, dC, and dTMP did not change. However, upon transfer to long-day growth conditions, dNTP and also dT concentrations increased rapidly within a few hours. Note that the dT pool

was in general 10 times larger than the dC pool, whereas the dTTP concentration exceeded that of dCTP only two- to three-fold. The increase of the dTMP pool was delayed by 24 h compared to the dT and dNTP pools, suggesting that initially dTMP was consumed for dTTP production as rapidly as it was made (Figure 4, A and B).

Ribonucleotide mono and triphosphates, especially ATP (adenosine triphosphate) but also UMP (uridine monophosphate) and UTP (uridine triphosphate) (Supplemental Figure S7, A and B), were rising in concentration over time, similar to dNTPs, suggesting that pyrimidine de novo biosynthesis was creating the substrates for RNR and the subsequent de novo T synthesis of thymidylates.

Interestingly, the dATP (deoxyadenosine triphosphate) and dTTP concentrations decreased after an initial burst but then rose again. This short-term reduction, more pronounced for dATP, corresponds with the onset of chloroplast DNA (cpDNA) replication at 6 h after transfer to long-day growth conditions (Paszkiewicz et al., 2017). The amount of dNTPs needed for cpDNA replication is substantial. In a conservative estimation, we calculated that at least 4.3-nmol dTTP is required per gram seed (see "Material and methods")—this amount is easily accessible with LC–MS (liquid chromatography–mass spectrometry). It is tempting to speculate that the pool size changes in this phase reflect a stronger consumption of dATP and dTTP than of dCTP and dGTP (deoxyguanosine triphosphate) because the chloroplast genome is rather A/T rich—the GC content of cpDNA is 36%.

One can calculate the GC content of the dNTPs that is the sum of dGTP and dCTP amounts as a percentage of the total amount of dNTPs. We previously suggested that this GC content of the dNTPs correlates with the GC content of genomes in plants and algae (Straube et al., 2021a). Because in germination the chloroplastic genome is replicated earlier than the nuclear genome, we wanted to assess whether the GC content of the dNTPs is specifically adjusted for the distinct phases of germination. Interestingly, we found that between 24 and 48 h after transfer to long-day growth conditions, when mainly cpDNA is made (Paszkiewicz et al., 2017), the GC content in dNTPs is higher (40%) than at later time points (30%; Figure 4C). We reported previously a content of 30% for adult Arabidopsis plants (Straube et al., 2021a). Indeed, it seems that the dNTP GC content is adjusted for cpDNA replication, but this does not relate to the GC contents of the chloroplastic and the nuclear genomes (both 36%). Maybe these adjustments of the dNTP GC content reflect different requirements of the evolutionarily unrelated polymerases (Mori et al., 2005; Parent et al., 2011) that replicate the cpDNA and ncDNA.

During germination the lack of organellar DUT1 reduces the dT pool, which is suppressed by an abolished organellar salvage of thymidine

The seedling phenotype of tk1b and tk1b dhfr-ts2 mutants (Figure 3) and the rapid formation of dT and thymidylates during germination (Figure 4, A and B) prompted us to study the nucleotide metabolome of all available mutants (Figure 3A). Previously none of these mutants had been analyzed regarding the nucleotide metabolome, which may be disturbed although most of the mutants are phenotypically inconspicuous. Based on the results of the time course (Figure 4A), we chose one time point in the germination and two time points in the seedling establishment phase (3 h "germination," 48 h "early seedling establishment," 144 h "late seedling establishment" after transfer to long-day growth conditions). These reflect an initial phase of dNTP pool increase (3 h), a strong increase

of pool sizes (48 h) and a last phase of decreasing dNTP pool sizes (144 h).

During germination (3 h), seeds from wild-type and all mutant plants contained roughly similar amounts of dNs and dNTs (Figure 5), suggesting that a lack of the investigated enzymes does not cause disturbances during embryogenesis that affect the thymidylate pools in the seed. The sink demand for dNTs is still low at 3 h because cpDNA biosynthesis is just beginning (Paszkiewicz et al., 2017). Thus, metabolic fluxes through the dNTP pools are probably small. We observed a decrease of the dT pool in the $dut1_{org}$ mutant compared to the wild-type. Assuming that $DUT1_{org}$ provides the substrate dUMP for dTMP biosynthesis and thus the lack of $DUT1_{org}$ interferes with dTMP production, the cytosolic and organellar dT salvage by the TKs compensate for the reduced dTMP synthesis resulting in depletion of the dT pool in the mutant. In fact, dT pools were restored to wild-type levels and even surpassed these in the $dut1_{org}$ $tk1b$ double mutant, suggesting that dT salvage in organelles and the cytosol is indeed responsible for depletion of the dT pool in the $dut1_{org}$ mutant. Interestingly, no differences in dT pools were observed in mutants lacking de novo dTMP biosynthesis (dhfr-ts1 and dhfr-ts2) compared to the wild-type, although DHFR-TS1 and DHFR-TS2 directly form dTMP from dUMP.

The observation that wild-type dT levels are surpassed in the $dut1_{org}$ $tk1b$ double mutant suggests that yet another process might be involved—possibly reduced dT degradation. We showed recently that dT can be degraded by the NSH1 in seedlings (Straube et al., 2021a) and it is tempting to speculate that compromised dTMP biosynthesis, caused by a putatively reduced dUMP pool in the $dut1_{org}$ mutant, results in reduced dT degradation by inhibition of NSH1. We attempted to monitor xanthosine accumulation as an indicator for NSH1 inhibition; however, no effect in the $dut1_{org}$ mutant compared to the wild-type was observed (Supplemental Figure S8). Regarding dTTP, no significant effects other than a slightly increased level in the dhfr-ts1 mutant were observed (Figure 5C). Furthermore, there is no contribution of C-to-U conversion by DCD to the formation of thymidylates during germination.

We also monitored the ratio of chloroplast/nuclear and mitochondrial/ncDNA in the wild-type and all mutants by qPCR (quantitative polymerase chain reaction). We observed a reduction of cpDNA, but not mitochondrial DNA (mtDNA; Figure 6; Supplemental Figure S9A), only in the tk1b mutant compared to the wild-type, indicating that already at this early time point in germination organellar dT salvage is crucial for the synthesis of cpDNA. A lack of the respective enzyme cannot be fully compensated by the cytosolic counterpart (TK1a). One needs to bear in mind that at this early time point only little cpDNA synthesis occurs, thus metabolic fluxes are probably small, but apparently thymidine salvage already contributes to the production of thymidylates.
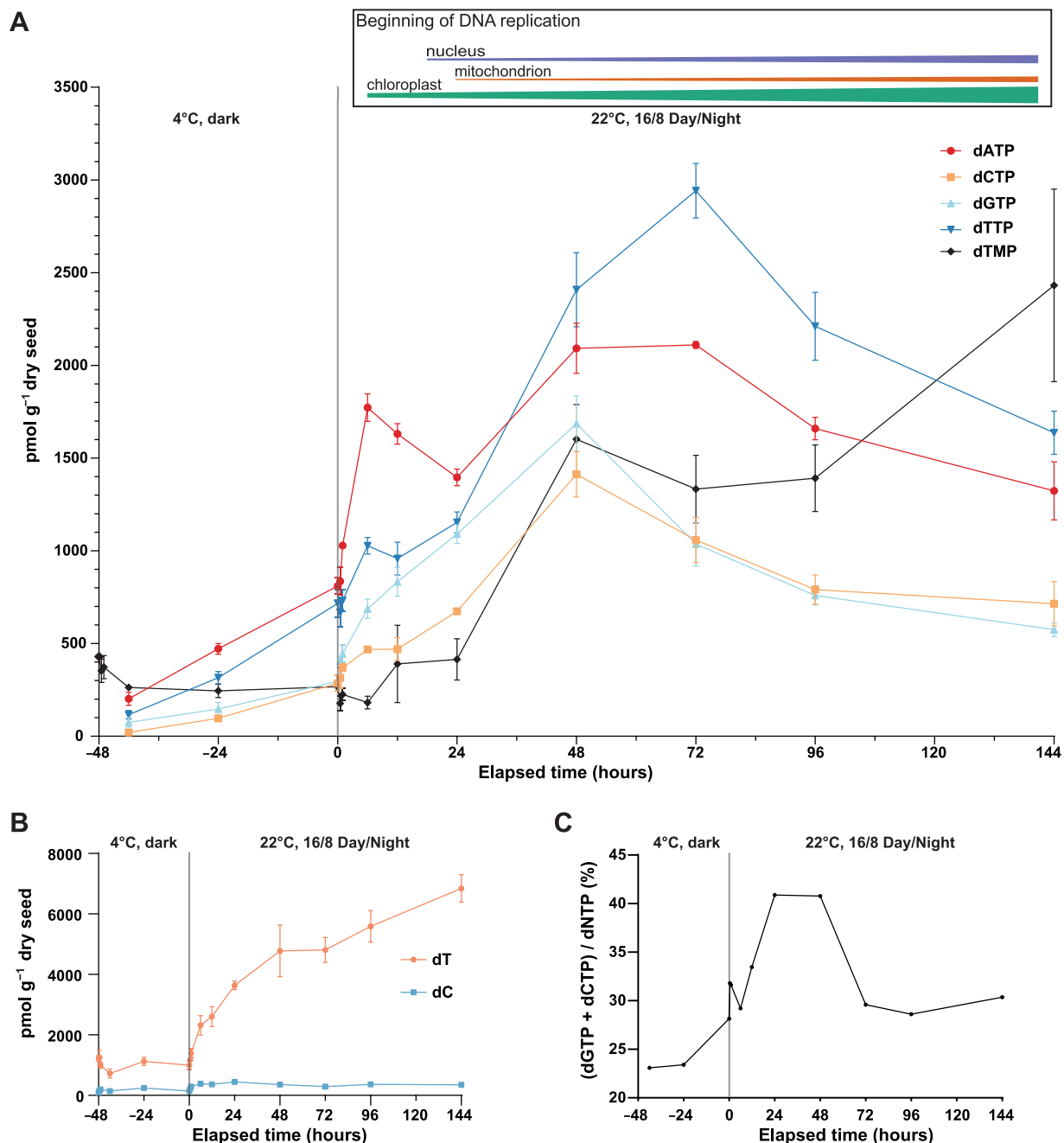
A

Figure 4 Concentration of dNTs and dNs at different time points during seed germination and in the establishing seedling. Concentrations of (A) all canonical dNTPs as well as dTMP and (B) dT and dC were quantified with LC–MS before transfer to long-day growth conditions (−48 h, −47.5 h, −42 h, −24 h; dark and 4°C), at transfer to long-day growth conditions (0 h, light and 22°C) and after the transfer to long-day growth conditions (0.5, 1, 6, 12, 24, 48, 72, 96, and 144 h). While seeds at the −48 h time point were dry all other samples were imbibed in water. At the −48, −47.5, and −42 h time points no reliable signal for dNTPs was detected. The onset of nuclear, plastid, and mtDNA replication, as reported by Paszkiewicz et al. (2017), Masubelele et al. (2005); and Sliwinska et al. (2009) is displayed by purple, green, and orange bars respectively in (A). C, Time course of the sum of dGTP and dCTP amounts as a percentage of the total amount of dNTPs. Four biological replicates (pooled seeds/seedlings) were analyzed. Error bars are SD.

During early seedling establishment, a lack of organellar dT salvage strongly interferes with chloroplast genome replication and globally affects the nucleotide metabolome

We tested the impact of mutants in the salvage, de novo T synthesis, and C-to-U conversion on the dN and dNT pools during early seedling establishment (48 h after transfer to long-day growth conditions) and observed the most severe effects at this time point (Figure 7). Mutants in the organellar salvage (tk1b; P = 0.052) and more pronounced in the cytosolic salvage (tk1a) showed an accumulation of dT, suggesting that both metabolic pathways are active

Figure 6 Abundance of the cpDNA relative to the ncDNA in wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis during germination. Seeds were imbibed in water for 48 h at 4°C in the dark then transferred for 3 h to long-day growth conditions and harvested. qPCR was performed with primers amplifying one gene from cpDNA (*RBCL*) and ncDNA (*UBC21*), respectively. The cpDNA/ncDNA ratios were calculated from Ct values for the respective genes. Ratios for the *dcd* mutant and the wild-type were analyzed in a separate experiment due to a technical error. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Asterisks indicate significant differences ($P < 0.05$) to Col-0. Three biological replicates (pooled seedlings) were analyzed. Error bars are SD, n.a. indicates that the value is not available. Adjusted *P*-values for multiple comparisons can be found in Supplemental Data Set 2.

(Figure 7A). These observations suggest that not only the organellar but also the cytosolic salvage pathway are active to form thymidylates at this developmental stage.

However, in the *tk1b* and *dhfr-ts2* single mutants, the amounts of dTMP were reduced (Figure 7B), demonstrating that during early seedling establishment not only organellar salvage but also mitochondrial de novo biosynthesis of dTMP contributes to the thymidylate pool. A reduction of dTMP was also observed in all double mutants with abolished *tk1b* function, while a *tk1b dhfr-ts2* double mutant has similar amounts of dTMP compared with the *tk1b* single mutant, suggesting that the impact of *dhfr-ts2* and *tk1b* on dTMP are not additive. While the *dhfr-ts2* single mutant has a reduced dTMP level compared to the wild-type, we do not observe this effect in the *dut1org dhfr-ts2* and *tk1a dhfr-ts2* double mutants for unknown reasons. We hypothesize that the observed reduction of total dTMP in the *tk1b* background reflects mainly a reduction of the chloroplast and mitochondrial dTMP pools because TK1b is located in these organelles. Total dTMP levels are neither altered in the *dhfr-ts1* and *tk1a* single mutants nor in the *dhft-ts1 tk1a* double mutant (Figure 7B). Apparently, dTMP generation in the organelles can compensate for a defect of dTMP production in the cytosol but not vice versa. Especially the salvage of dT in the organelles cannot be compensated by cytosolic dT salvage, which is nonetheless clearly active at this time point (see effect of *tk1a* mutation on dT in Figure 7A). This may
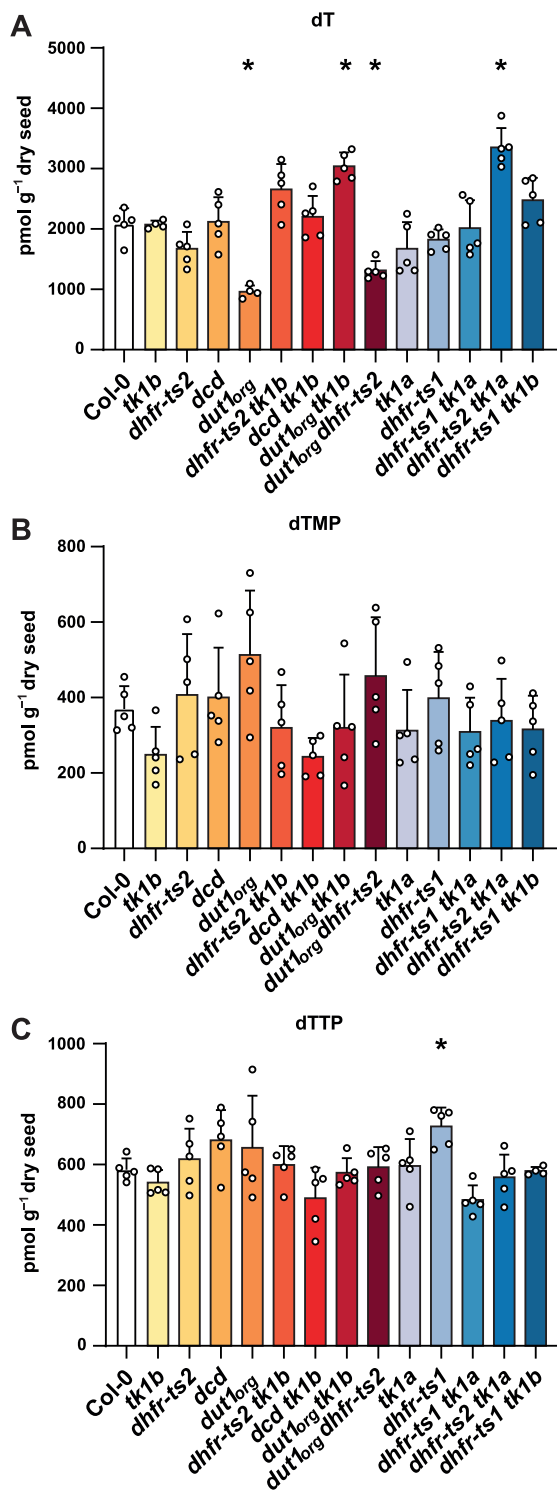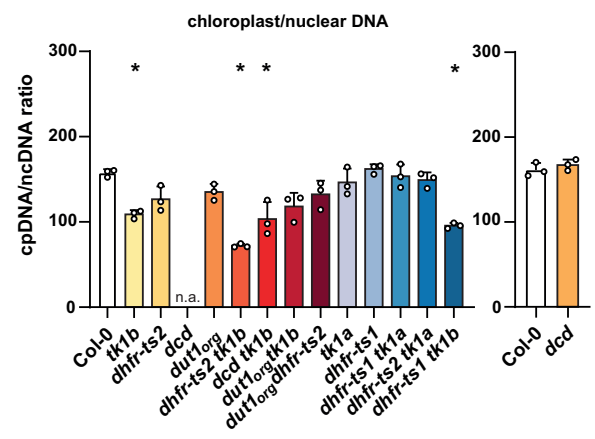
Figure 5 Concentration of thymidine and thymidylates in wild-type and mutant seeds lacking enzymes putatively involved in thymidylate homeostasis during germination. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 3 h to long-day growth conditions and harvested for LC–MS analysis. Concentrations of (A) dT, (B) dTMP, and (C) dTTP in seeds are shown. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Asterisks indicate significant differences ($P < 0.05$) to Col-0. Four to five biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted *P*-values for multiple comparisons can be found in Supplemental Data Set 2.
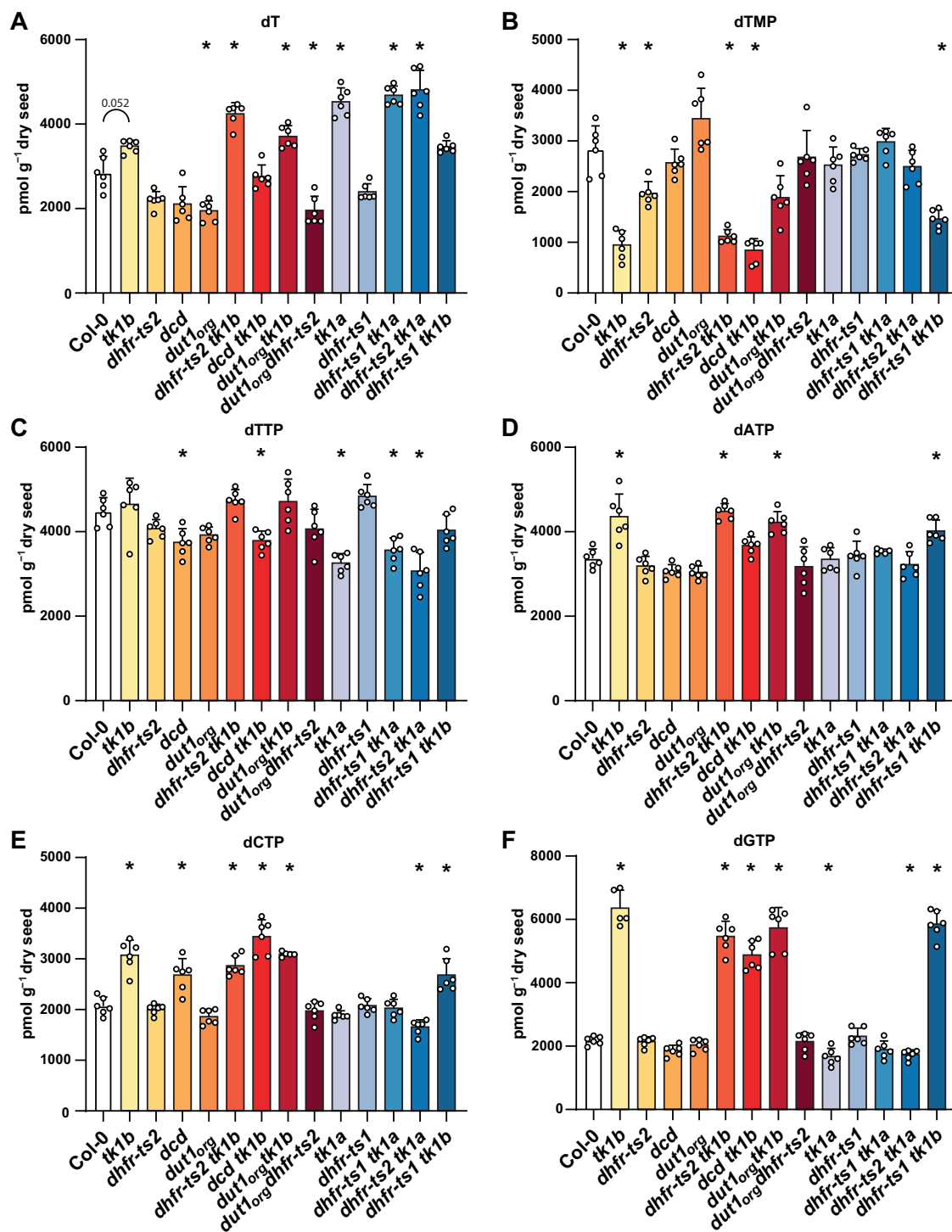
Figure 7 Concentrations of dT, dTMP, and dNTPs in wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis during early seedling establishment. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 48 h to long-day growth conditions and harvested for LC–MS analysis. Concentrations of (A) dT, (B) dTMP, (C) dTTP, (D) dATP, (E) dCTP, and (F) dGTP in germinating seed are shown. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Asterisks indicate significant differences ($P < 0.05$) to Col-0; $P > 0.05$ are indicated if value comparisons are made in the text. Five to six biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted P-values for multiple comparisons can be found in Supplemental Data Set 2.

be explained by a low capacity for thymidylate import into the organelles.

Interestingly, the amounts of dTTP were not as different as those of dTMP between any of the mutants and the

wild-type (Figure 7C). We showed that in the *tk1b* mutant, due to the lack of thymidylates, the replication of the chloroplast genome severely slowed down (Supplemental Figure S6B), probably because the plastidic dTTP pool is depleted.
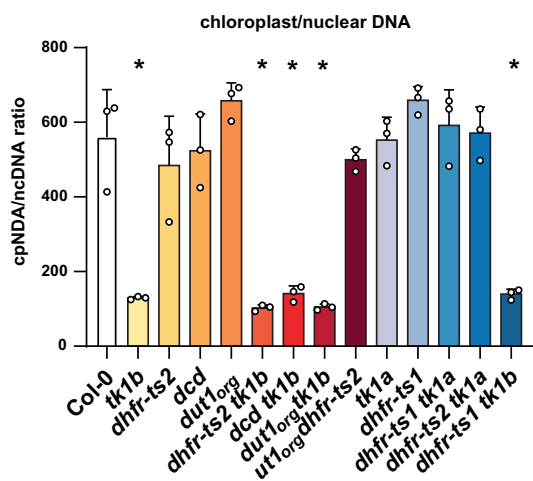
Figure 8 Abundance of the cpDNA relative to the ncDNA in wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis during early seedling establishment. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 48 h to long-day growth conditions and harvested. qPCR was performed with primers amplifying one gene from cpDNA (*RBCL*) and ncDNA (*UBC21*), respectively. The cpDNA/ncDNA ratios were calculated from Ct values for the respective genes. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Asterisks indicate significant differences ($P < 0.05$) to Col-0. Three biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted *P*-values for multiple comparisons can be found in Supplemental Data Set 2.

The measured dTTP would, in this case, reflect the cytosolic and mitochondrial pools, but these are not consumed since ncDNA and mtDNA replication is negligible at this time point in germination (Paszkiewicz et al., 2017). In line with this hypothesis, no differences in dTTP concentrations were observed between *tk1b* and the wild-type. Consistent with a probable arrest of cpDNA replication in the *tk1b* mutant, we observed significantly elevated concentrations of the other DNA building blocks (dGTP, dATP, and dCTP) compared to the wild-type (Figure 7, D–F). These probably accumulate in *tk1b* because they are not consumed for cpDNA replication. Interestingly, the amount of building blocks that accumulate is about the same as that required for cpDNA synthesis during germination (see "Materials and methods" for an estimation). Furthermore, we observed elevated levels of NTPs (Supplemental Figure S10) in the *tk1b* background compared to the wild-type, suggesting that an arrest of cpDNA replication reduces the template required for transcription, resulting in the accumulation of RNA building blocks.

We also assessed the cpDNA copy number in mutant and wild-type plants at 48 h after transfer to long-day growth conditions. The ratio of cpDNA to ncDNA did not change significantly between the 3 and the 48-h time points in the *tk1b* mutant, whereas it increased almost 3.5-fold in the wild-type (Figures 6 and 8). The halt of cpDNA synthesis in the *tk1b* mutant combined with the ongoing synthesis in the wild-type resulted in a 4.3-fold difference of the cpDNA/
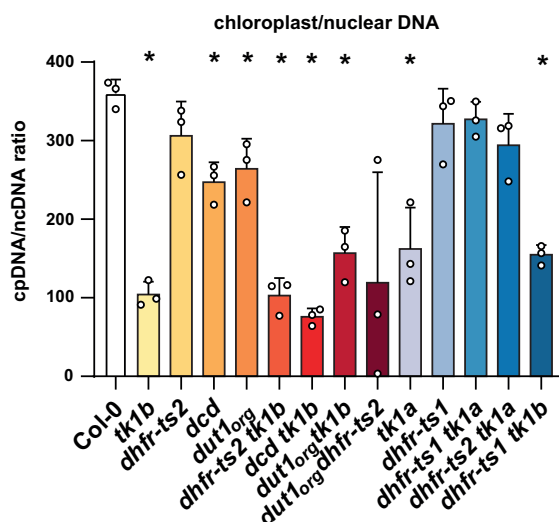
Figure 9 Abundance of the cpDNA relative to the ncDNA in wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis during late seedling establishment. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 144 h to long-day growth conditions and harvested. qPCR was performed with primers amplifying one gene from cpDNA (*RBCL*) and ncDNA (*UBC21*), respectively. The cpDNA/ncDNA ratios were calculated from Ct values for the respective genes. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Asterisks indicate significant differences ($P < 0.05$) to Col-0. Three biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted *P*-values for multiple comparisons can be found in Supplemental Data Set 2.

ncDNA ratio between wild-type and *tk1b* and also all double mutants with *tk1b* background at 48 h, indicative of cpDNA replication arrest in *tk1b*. No differences were observed in the mtDNA/ncDNA ratios between the wild-type and mutants (Supplemental Figure S9B), which is in line with previous observations that mtDNA and ncDNA replication is negligible in this phase of germination (Paszkiewicz et al., 2017). We supported this result using additional loci on the chloroplast and mitochondrial genomes and confirmed our initial observation (Supplemental Figure S11).

In established seedlings, organellar salvage, and mitochondrial de novo T synthesis are still active but less important

Next, we investigated (d)N and (d)NT concentrations as well as cpDNA/ncDNA ratios in 6-day-old seedlings (144 h after transfer to long-day growth conditions). This time point was chosen, because the *tk1b* mutant appeared to have recovered as indicated by green pigmentation comparable to that of the wild-type (Supplemental Figure S6A). Nonetheless, a reduced cpDNA/ncDNA ratio was still detectable in *tk1b* compared to wild-type plants but it was only 3.5-fold smaller than that of the wild-type (Figure 9) in contrast to a 4.3-fold difference at the 48-h time point (Figure 8). Note that the ratios between the 48 and 144 h time points cannot be compared since ncDNA is

synthesized at 144 h but not at 48 h (Barroco et al., 2005), causing a general decrease of the ratios. During late seedling establishment, mtDNA is likely replicated as well. Salvage by the organellar TK1b but also the cytosolic TK1a seems to support cpDNA and mtDNA synthesis because the respective mutants (tk1a and tk1b) tend to have a reduced cpDNA/ncDNA and mtDNA/ncDNA (tk1b; $P = 0.054$, tk1a; $P = 0.078$) ratio compared to the wild-type (Figure 9; Supplemental Figure S9C).

The accumulation of dATP, dCTP, and dGTP observed in the tk1b mutant during early seedling establishment (48 h) was not observed anymore in the established seedling (Figure 10, A–C), suggesting that cpDNA synthesis was now operative. Consistent with the greening of the tk1b seedlings (Supplemental Figure S6A), similar concentrations of NTs in wild-type compared to mutant seedlings were observed (Supplemental Figure S12), which suggests that transcription in the chloroplast was not halted anymore. It is tempting to speculate that in the established seedling (144 h), in contrast to the seedling during the early establishment phase (48 h), the transport of thymidylates to the chloroplast has sufficient capacity to allow cytosolic de novo T synthesis and salvage to partially compensate for the loss of these pathways in organelles. Interestingly, a reduction of dTMP and also dTTP was observed in the seedling not only in the tk1b but also in the tk1a mutant compared to the wild-type. This suggests that both corresponding enzymes contribute to the thymidylate pools in seedlings (Figure 10D). Loss-of-function of any of these genes cannot be fully compensated by the respective counterpart and causes an alteration of the nucleotide metabolome that does not affect the visible appearance. A small effect of DHFR-TS1 but not DHFR-TS2 on the dTMP concentration was also observed (Figure 10E), suggesting that additionally cytosolic de novo T synthesis contributes to the thymidylate pools. Similarly, an accumulation of dCTP and a reduction of dTMP was detected in the dcd mutant compared to the wild-type (Figure 10, C and E), suggesting that DCD has a role in equilibrating the dNTP pools, especially by reducing the dCTP concentration (already observed during early seedling establishment; Figure 7E), in addition to a role in thymidylate synthesis.

Chloroplast TK1b can complement the dhfr-ts2 tk1b double mutant

Native TK1b is localized in mitochondria and chloroplasts. We wondered whether an engineered targeting of TK1b in the tk1b background exclusively localized to either the mitochondria or the chloroplasts can complement the pigmentation phenotype of the mutant. We expressed TK1b fused to either a chloroplast transit peptide (CTP) or a mitochondrial transit peptide (MTP; Kohler et al., 1997; Shen et al., 2017), in the tk1b dhfr-ts2 background because this mutant shows the strongest phenotype. The constructs were driven by a nopaline synthase promoter (Pro$_{NOS}$:CTP:TK1b, Pro$_{NOS}$:MTP:TK1b). To assess whether TK1b activity is required for complementation, transgenic plants expressing the

corresponding inactive variants were also generated. Here the catalytically important amino acid Glu163 was exchanged to Gln (TK1b$_{E163Q}$), rendering the enzyme inactive (Welin et al., 2004). A quantification of green and yellow pixels in images from cotyledons of 5-day-old seedlings showed a reduced ratio of green/yellow pixels in the dhfr-ts2 tk1b double mutant compared to wild-type, as observed previously. The CTP:TK1b transgene fully restored the wild-type phenotype, whereas the MTP:TK1b transgene complemented the lack of TK1b only partially (Figure 11). The inactive TK1b version TK1b$_{E163Q}$ did not restore the wild-type phenotype independent of its localization. When targeted to the chloroplast, it even aggravated the phenotype, indicative of a dominant negative effect (Figure 11). The in vivo function of TK1b thus clearly requires its enzymatic activity. The observation that only the chloroplastic TK1b variant restores the wild-type phenotype fully further suggests that transport of thymidylates over the chloroplast membrane is limited. Nonetheless, mitochondrial TK1b can partially support the chloroplasts, indicating that transport of thymidylates occurs to some extent in the already established seedling.

Is global genome repair the source of dT during germination and seedling establishment?

The dT concentration is comparatively high and rises constantly during germination and seedling establishment (Figure 4B) but the source of this dT is unclear. Since this study highlights the importance of dT salvage for cpDNA replication in the process of germination, it is of interest to identify the source of dT. The phenotype of the tk1b mutant is aggravated by the application of genotoxic chemicals, suggesting that salvage of dT supports DNA repair processes by supplying building blocks (Pedroza-García et al., 2015, 2019). However, dNMPs produced during global DNA repair may be dephosphorylated and the resulting dNs require salvage; thus, DNA repair could also be a substantial source of dT. DNA repair is a prominent process during germination even in the absence of genotoxic stress (Bray and West, 2005; Weitbrecht et al., 2011). Therefore, we used five mutants crucial for global genome repair, that is, rad1, atm-1, atr-2, cen2, and ddb1-A (Garcia et al., 2003; Molinier et al., 2004, 2008; Yoshiyama et al., 2009), to assess whether they show reduced dT concentrations in seeds 48 h after transfer to long-day growth conditions. This was not the case for any of the mutants and some even contained more dT (Figure 12), suggesting that DNA repair is not a significant source for dT.

Discussion

Formation of the thymidylate pool during seedling establishment is crucial for cpDNA replication

Nucleotide metabolism, especially with respect to deoxynucleotides, is poorly understood compared to other metabolic pathways in plants, although it is of central importance (Ashihara et al., 2020; Straube et al., 2021a). In this study, we monitored dNT and dN concentrations at different time
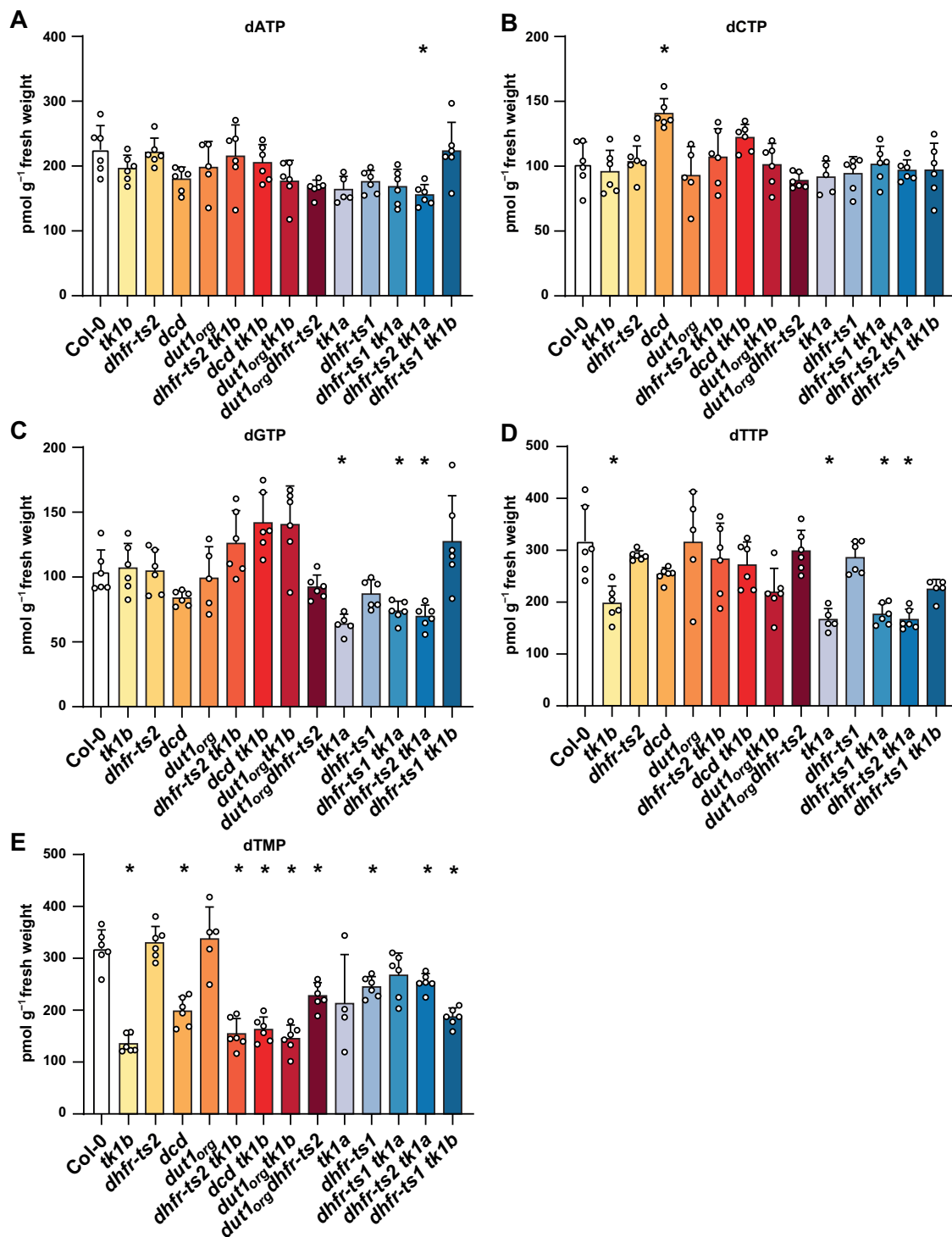
Figure 10 Concentrations of dNTPs and dTMP in wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis during late seedling establishment. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 144 h to long-day growth conditions, and harvested for LC–MS analysis. Concentrations of (A) dATP, (B) dCTP, (C) dGTP, (D) dTTP, and (E) dTMP in germinating seed are shown. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Asterisks indicate significant differences ($P < 0.05$) to Col-0. Five to six biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted P-values for multiple comparisons can be found in Supplemental Data Set 2.

points during seed germination and observed a strong increase of dNTPs but also dT and with some delay dTMP (Figure 4, A and B) within the first hours of transferring

seeds to long-day growth conditions. This coincides with the induction of nucleotide biosynthesis genes, which are among the first to be activated in germination (Law et al., 2014).
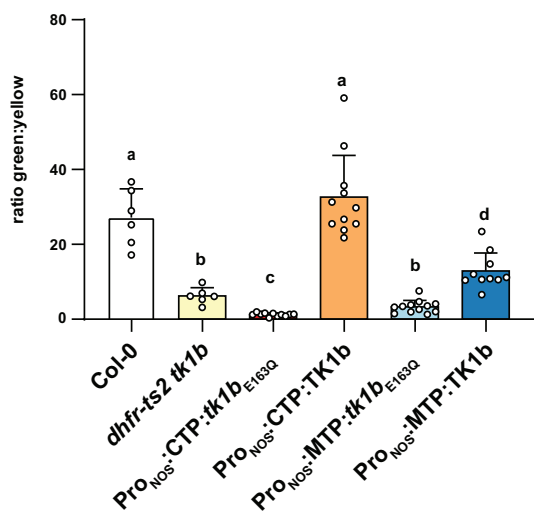
Figure 11 Phenotypic analysis of the chloroplast- or mitochondria-specific complementation with *TK1b* in the *dhfr-ts2 tk1b* background. TK1b fused to a CTP or an MTP was expressed under the control of a nos promoter (Pro$_{NOS}$). The wild-type version of TK1b and a catalytically inactive mutant version (TK1b$_{E163Q}$) were expressed in transgenic plants with *dhfr-ts2 tk1b* mutant background. Images of cotyledons were taken 5 days after transfer to long-day growth conditions. The ratio of green to yellow pixels was quantified for each individual plant. For all transgenic plants, two independent lines were selected and analyzed. Six to twelve biological replicates (individual plants) were analyzed. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Error bars are SD. Different letters indicate significant differences ($P < 0.05$) between genotypes. Adjusted *P*-values for multiple comparisons can be found in Supplemental Data Set 2.

With the onset of cpDNA replication (6 h after transfer to long-day growth conditions; Paszkiewicz et al., 2017), the concentrations of dATP and dTTP decrease transiently, which is probably due to their consumption by the chloroplast replication machinery. Only the cpDNA genome is replicated during germination (Paszkiewicz et al., 2017) and the cpDNA amount increases 5 times during maturation of the proplastid (Fujie et al., 1994), creating a high demand for dNTPs (see "Material and Methods" for an estimation of the required amounts). The cpDNA content is thus particularly affected by a genetic block of thymidylate formation, whereas the mtDNA is not (Figures 6 and 8; Supplemental Figure S9, A and B). Together these data suggest that cpDNA is the main sink for dNTPs during germination.

Deoxy-UMP for DHRF-TS2 is not mainly supplied by DUT1$_{org}$ in mitochondria of germinating seeds

In this study, we identified a variant of DUT1 that is located in mitochondria and chloroplasts (DUT1$_{org}$; Figure 2) and we were able to create a dut1$_{org}$ mutant. We speculated that the main role of DUT1$_{org}$ might be to supply dUMP as substrate for the mitochondrial dTMP-synthesizing enzyme DHFR-TS2. For cytosolic DUT1 such a role has been demonstrated in other organisms (Vértessy and Tóth, 2009; Martínez-Arribas et al., 2020). If DUT1$_{org}$ has this role, effects

on the thymidylate pools should be similar in the *dut1$_{org}$* mutant and the *dhfr-ts2* mutant. However, while the *dhfr-ts2* mutant has less dTMP compared to the wild-type during early seedling establishment, the dTMP pool in the *dut1$_{org}$* mutant is unaltered compared to the wild-type (Figure 7B). Additionally, the strong *dhfr-ts2 tk1b* phenotype is not observed for the *dut1$_{org}$ tk1b* double mutant (Figure 3), although this would be expected if DUT1$_{org}$ supplies substantial amounts of dUMP for DHFR-TS2. Yet, DUT$_{org}$ may nonetheless be involved in de novo T synthesis because its absence appears to stimulate salvage, maybe to compensate for reduced biosynthesis. The induction of salvage is indicated by less dT in the *dut1$_{org}$* mutant compared to the wild-type during germination (Figure 5A).

In summary, we discovered that DUT1 is also located in the mitochondria and the chloroplasts where its main function is probably to sanitize the dNTP pools and not to generate dUMP as DHFR-TS2 substrate in mitochondria. However, we cannot exclude that DUT1$_{org}$, DCD or so far unknown enzymes contribute to the dUMP pool in a redundant manner so that a lack of DUT1$_{org}$ or DCD alone is insufficient to affect dUMP and consequently thymidylate pools similarly to what was observed for DHFR-TS2.

A loss of DCD function increases the concentration of dCTP and has limited influence on the formation of thymidylates

In rice, it was suggested that the formation of thymidylates depends greatly on DCD catalyzing the deamination of dCMP to dUMP, the DHFR-TS substrate (Niu et al., 2017). Here we were able to determine the absolute concentrations of nucleotides and observed clearly increased levels of dCTP in the *dcd* mutant compared to the wild-type but only a slight decrease in dTTP in the early and dTMP in the late seedling establishment phase. This is consistent with our observation that the *dcd tk1b* double mutant is phenotypically similar to the *tk1b* mutant and not to the *dhfr-ts2 tk1b* double mutant that is impaired in salvage and de novo T biosynthesis in mitochondria (Figure 3). In contrast to rice, it appears that DCD plays only a limited role in thymidylate biosynthesis in Arabidopsis. Instead DCD might have an important role in balancing the dNTP pools by limiting the dCTP pool size. Nonetheless, in a genetic background over accumulating deoxycytidilates (in a mutant of *CYTIDINE DEAMINASE*) we previously observed more dTMP and dTTP (Straube et al., 2021a) indicating that under these conditions DCD contributes to thymidylate synthesis in Arabidopsis.

Limited transport capacity for thymidylates during germination may cause the dependency on dT salvage in chloroplasts

During early seedling establishment (48 h), the salvage of dT in the chloroplast by TK1b is predominantly responsible for providing dTMP as precursor for cpDNA replication. This was demonstrated by (1) severely reduced cpDNA, (2) reduced dTMP concentration, and (3) an accumulation of
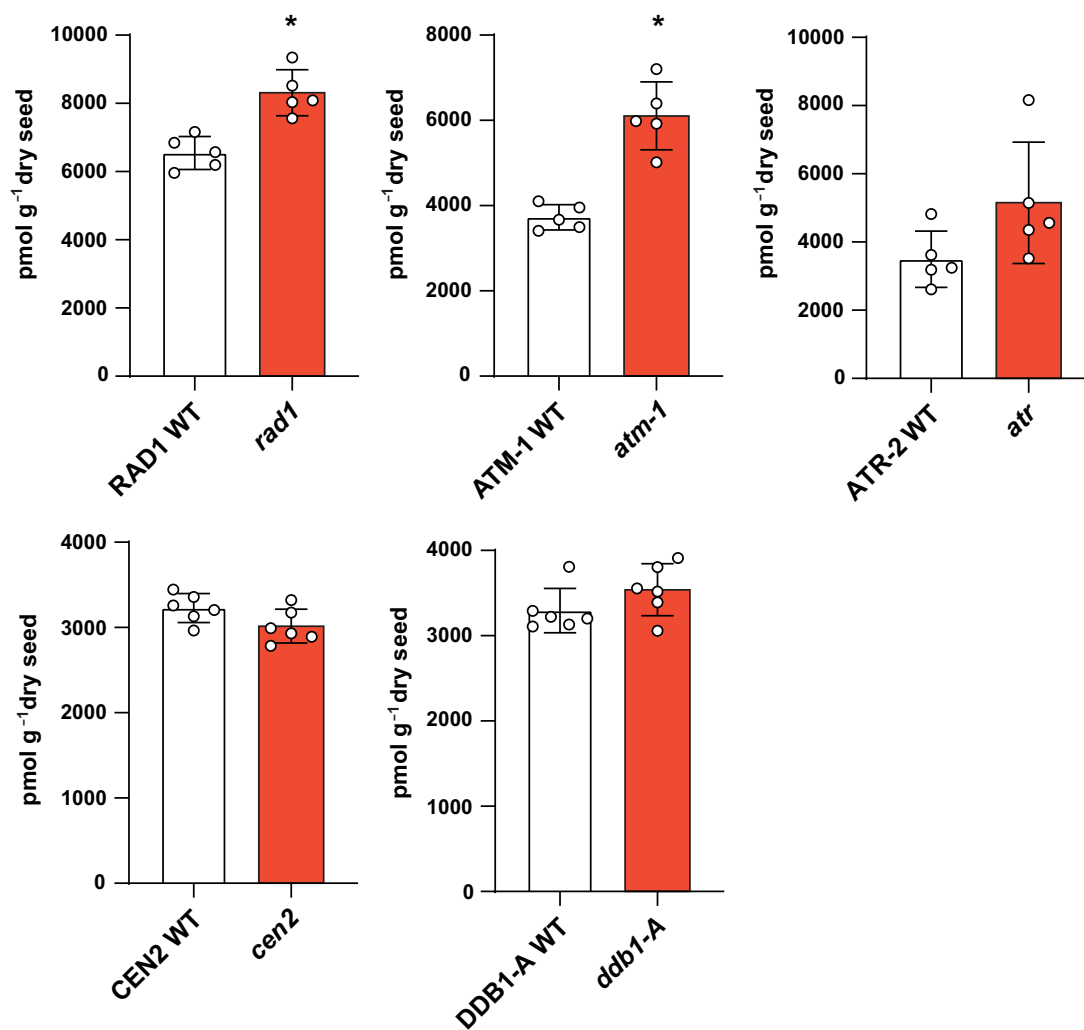
Figure 12 Concentrations of dT in wild-type and mutants affected in global genome repair during early seedling establishment. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 48 h to long-day growth conditions, and harvested. Concentrations of dT in plants lacking RAD1, ATM-1, ATR-2, CEN2, and DDB1-A and their corresponding wild-types segregated from the respective mutant populations are shown. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were used for statistical analysis. Asterisks indicate significant differences ($P < 0.05$) to the wild-type. Six biological replicates (pooled seedlings) were analyzed. Error bars are sd. Adjusted P-values for statistical comparisons can be found in Supplemental Data Set 2.

other dNTPs and rNTPs in the *tk1b* mutant compared to the wild-type. A reduction of cpDNA replication in the *tk1b* mutant was observed previously in Arabidopsis (Le Ret et al., 2018). However, here we show that the impact of defective organellar dT salvage is by far more severe than previously reported. This can be explained by different sampling time points (48 h here and 6 days in Le Ret et al., 2018). The strong impact on cpDNA, which we observed early for the *tk1b* mutant alone (48 h), was previously only observed later for the *tk1a tk1b* double mutant (6 days) (Le Ret et al., 2018), which is in line with our interpretation that at later time points a loss-of-function in *TK1b* can be compensated by *TK1a*. Here we identified a time point during early seedling establishment (48 h) when such a compensation is very limited or not at all possible. Is this simply because TK1a is not active at this time point? This seems unlikely because the transcript for TK1a is clearly present

during germination and early seedling establishment (Law et al., 2014; Supplemental Figure S1) and more importantly dT is accumulating at this time point in the *tk1a* mutant suggesting that in the wild-type TK1a is phosphorylating dT to dTMP in the cytosol (Figure 7A). If TK1a is active at this time point, why can it not compensate a *tk1b* mutation? A possible explanation is that the import capacity for thymidylates into plastids is limited during germination and early seedling establishment (up to 48 h) and only becomes sufficient to support cpDNA synthesis in the established seedling (144 h). In contrast, import of the TK substrate dT appears to be sufficient at all times. Possible candidates for organellar thymidylate transporters are NDT1 and NDT2, which transport NAD^+ but with lower efficiency also thymidylates (Palmieri et al., 2009) and whose genes are induced during early seedling establishment (Narsai et al., 2011). The onset of thymidylate transport post 48 h allows the cytosolic

formation of thymidylates by DHFR-TS1 and TK1a to compensate for a lack of dT salvage in chloroplasts, which explains the reversion of the pigmentation and cpDNA phenotypes in established seedlings of the *tk1b* single and *tk1b dhfr-ts2* double mutants (Supplemental Figure S6, A and B). Interestingly, only in the *tk1b dhfr-ts2* double mutant variegated adult plants were obtained (although quite rarely) in which whole sections of the plant were albinotic (Supplemental Figure S6C). The lesions observed in the cotyledons might in rare cases also occur in the adult plant for unknown reasons.

In summary, there is no redundancy between TK1a and TK1b at early time points of germination, whereas later they do seem partially redundant probably because thymidylates are then sufficiently exchanged between compartments. This shows that apparently functionally redundant genes can have a unique role depending on the particular developmental stage of the plant. At least in eudicots, an organellar and a cytosolic TK are conserved (Le Ret et al., 2018), also suggesting that both isoenzymes have nonredundant roles.

Although at early time points, TK1a-mediated thymidylate production cannot compensate the lack of thymidylates in the plastids in *tk1b* mutants, mitochondrial thymidylate production by DHFR-TS2 can partially support dTMP formation during early seedling establishment (Figure 7). Additionally, the *dhfr-ts2 tk1b* mutant has a more severe phenotype compared to the *tk1b* single mutant. It is puzzling that cytosolic thymidylates cannot support the chloroplast at all in germination, whereas mitochondrial thymidylates can.

The prominent role of dT is highlighted by its strong accumulation and the absolute requirement for its salvage

We observed a strong accumulation of dT that is continuously rising throughout germination (Figure 4B). Interestingly, the absolute amount of dT is similar to the combined amounts of dTMP and dTTP and it remains unclear why so much dT is generated and what is its source. Maybe because the plastids rely on dT import but apparently cannot import thymidylates during germination, the cellular dT concentration must be high. The dT may be generated by dephosphorylation of dTMP originating from de novo T synthesis. In this model dT salvage, that is, the re-phosphorylation to dTMP, would be placed downstream of the de novo T synthesis (Figure 1). Apart from serving to provide dT for the plastids, this energy consuming process may also improve the purity of the thymidylate pool. The substrate specificity of TK1 may represent a molecular filter excluding noncanonical or damaged nucleotides from the dTMP pool.

Although de novo T synthesis would be an obvious source for the dT accumulating during germination, it is unclear how RNR activity could lead to excess thymidylates over the other deoxynucleotides, because the allosterically regulated enzyme is known to produce dNDPs in stochiometric amounts (Nordlund and Reichard, 2006). If de novo T synthesis is the main source of thymidylates during germination, it is also curious that first dTMP is made, then it is

dephosphorylated and then it is salvaged to dTMP again, which appears to be a futile cycle.

Because of these unsolved issues, we also followed up an alternative hypothesis regarding the origin of dT. We tested whether DNA strand degradation caused by global genome repair may be a source of dT. Global genome repair must produce dG and dC as well as dA and dT in stoichiometric amounts as a result of the fixed ratio of nucleotides in the DNA. However, dA was not detected, whereas dT was readily detectable in our analysis and its concentration was always at least one order of magnitude higher than that of dC (Figure 4B). Additionally, several mutants with a defect in global genome repair did not have reduced amounts of dT. It is thus unlikely that genome repair is a quantitatively significant source of dT during germination, although we cannot exclude the possibility that other mutants with an impact on DNA repair such as *sog1* (Yoshiyama et al., 2009), potentially combined with the application of genotoxic stress, can cause altered dT concentrations during germination. The quantification of dT in seeds and during germination (Figure 4B) also disfavors the hypothesis that dT itself is a metabolite stored in seeds that is used for dTTP synthesis. The dT concentration is low in the beginning of germination and then rises continuously, whereas a storage component should be consumed over time. Thus, the question of the origin of dT remains open.

In the literature, salvage is usually discussed as a process that reintegrates nucleosides or nucleobases, which are substrates for a degradation pathway. Energy conservation is often described as the purpose of salvage (Zrenner et al., 2006; Ashihara et al., 2020; Witte and Herde, 2020). The importance of pyrimidine salvage is underlined by the fact that not only the double mutant of *tk1a* and *tk1b* is not viable (Clausen et al., 2012) but also the double mutant of *ukl1 ukl2*, in which the salvage of the ribonucleoside uridine is abolished, is not viable (Chen and Thelen, 2011). A role for dT salvage in the defense against genotoxic stress in chloroplasts and the cytosol was demonstrated previously (Pedroza-García et al., 2015, 2019; Xu et al., 2015). Here we show that dT salvage has a severe impact on thymidylate pools also in the absence of any genotoxic stress and plays a vital role during germination. The general importance of dT salvage disfavors the idea that this pathway merely serves as an energy-conserving mechanism that reintegrates "accidentally" dephosphorylated dTMP. Thymidine salvage rather represents a major pathway for providing thymidylates. TKs in other organisms are usually discussed in tumorgenesis, DNA repair, and viral infection (Bitter et al., 2020) but in plants dT salvage seems to have a more fundamental role adapted to the plant lifestyle, which involves germination of dormant seeds and chloroplast metabolism.

Material and methods

Plant materials
T-DNA insertion mutant lines of Arabidopsis (*A. thaliana*) for *DHFR-TS1* (AT2G16370; GK-010G06) and *DHFR-TS2*

(AT4G34570; GK-893C02), both previously described in Gorelova et al. (2017); *TK1a* (AT3G07800; SALK_097767C) and *TK1b* (AT5G23070; SALK_074256C), both previously described in Clausen et al. (2012); *ATM-1* (AT3G48190; SALK_006953), described in Garcia et al. (2003); *ATR-2* (AT5G40820; SALK_032841) and *DDB1-A* (AT4G05420; SALK_038757), both previously described in Molinier et al. (2008), and *CEN2* (AT4G37010; SALK_127583) described in Molinier et al. (2004); and *RAD1* (AT5G41150; SALK_096156), previously described in Yoshiyama et al. (2009); and *DCD* (AT3G48540; SALK_057582; Niu et al., 2017) were obtained from the Nottingham Arabidopsis Stock Centre and homozygous T-DNA insertion mutants were screened by PCR. The primers used are shown in Supplemental Data Set 1. The mutant line for DUT1$_{org}$ (AT3G46940) was newly obtained by applying the CRISPR/Cas9 technique (see below). The mutation was introduced in the region encoding the transit peptide. The loss of the DUT1$_{org}$ protein was confirmed by transient expression and subcellular localization of the enzyme fused to a fluorescent tag using confocal microscopy. Double mutants were obtained by crossing the respective single mutants. All single and double mutant lines were grown together in parallel in a randomized fashion with Col-0 serving as the wild-type control, or their corresponding wild-type segregated from the respective mutant population to obtain a uniform seed batch. The seed material used for preparing the time course experiment consisted of about 1-year-old Col-0 seeds originating from multiple Col-0 plants. For all experiments involving the mutant collection the uniform seed batch was used. The age of the seeds at the time of analysis was 4 months. The seeds of the mutants involved in global genome repair were 2 months old.

Growth conditions

For phenotypic analysis, seeds were sown on soil (Steckmedium, Klasmann-Deilmann, Geeste, Germany), pretreated at 4°C in the dark for 48 h, and then transferred to a growth chamber set to 16-h day/8-h night (long-day), 22°C day/20°C night, 100 μmol s^{-1} m^{-2} light (Osram Fluora 36W/77, Osram, Munich, Germany), and 70% humidity. For metabolite and qPCR analysis, seeds were surface sterilized with 70% (v/v) EtOH for 20 min, dried until all EtOH evaporated and weighed in 10 ± 0.5 mg aliquots. The seeds were transferred onto two layers of 6 × 6 cm autoclaved filter paper soaked with 2-mL cold and sterile modified liquid half strength MS media as described by Niehaus et al. (2020) (3-mM CaCl$_2$, 1.5-mM MgSO$_4$, 1.25-mM KH$_2$PO$_4$, 18.7-mM KNO$_3$, 0.1-mM FeSO$_4$, 0.1-mM Na$_2$EDTA, 0.13-mM MnSO$_4$, 0.1-mM H$_3$BO$_3$, 0.03-mM ZnSO$_4$, 1-μM Na$_2$MoO$_4$, 0.1-μM CuSO$_4$, 0.1-μM NiCl$_2$, 0.5-g L^{-1} MES, pH adjusted to 5.7 with KOH) in a closed unsealed petri dish (94 × 16 mm). Seeds were evenly spread across the filter paper to minimize the number of seeds touching each other. The plates with the seeds were pretreated at 4°C in the dark for 48 h, and then transferred to a growth chamber set to 16-h day/8-h night (long-day), 22°C day/20°C night, 100 μmol s^{-1} m^{-2} (Osram Fluora 36W/77) light, and 70% humidity. For time points

beyond 48 h after transfer to long-day growth conditions, plates were carefully watered with 2-mL sterile bi-distilled water each, and 48 h later 2-mL sterile modified half-strength liquid MS media were added.

Phenotypic analysis and quantification of chlorotic leaf areas

The images used for phenotypic analysis were taken using a Nikon SMZ25 stereo microscope equipped with a Nikon DS-Ri2 microscope camera (Nikon, Minato, Japan) set to a resolution of 4,908 × 3,264 pixels. Images were taken 3–6 days after transfer to long-day growth conditions with identical instrument and light settings to ensure comparability. The chlorotic leaf area and green leaf areas were quantified using Fiji (Schindelin et al., 2012) with ImageJ 1.53c. First, the image was loaded into Fiji, and subsequently the Color Threshold was determined using the Hue, Saturation, Brightness Color Space (Image → Adjust → Color Threshold). For the chlorotic leaf area, the Hue was set to 35–45, saturation was kept at 0–255, and the Brightness was set automatically. The Thresholding method was set to default. For the green leaf area, the Hue was set to 46–100, and all other settings were kept the same. For each individual image, the green leaf area/chlorotic leaf area ratio was calculated.

Cloning

The constructs used for generating the dut1$_{org}$ knockout mediated by CRISPR/Cas9 and the chloroplast- or mitochondria-specific complementation of tk1b were cloned using the MoClo system (Weber et al., 2011; Engler et al., 2014). For this, a variety of different intermediate vectors had to be created that were used in combination with already published ones. The vector collection used for introducing the CRISPR-mediated lesion in DUT1 was previously used in Rinne et al. (2021). The MoClo Toolkit and the Plant Parts Kit were a gift from Nicola Patron and Sylvestre Marillonnet (Addgene kit #1000000044, Addgene kit #1000000047), the pHEE-2E-TRI was a gift from Qi-Jun Chen (Addgene plasmid #71288). The guideRNA used for the dut1$_{org}$ loss-of-function was chosen by applying the criteria proposed by Doench et al. (2014). The final level_2 vector used for generating the mutant was assembled from four independent level_1 vectors. The level_1 vector mediating the phosphinothricin resistance for later plant selection was cloned by combining pICH87633, pICH43844, and pICH41421 with the recipient vector pICH47732 in a BsaI cut/ligation, resulting in pICH47732_Basta_pos1_fwd (V183). For the level_1 vector containing the Cas9WT + Intron under the control of an egg-cell specific promoter, first the Cas9 + Intron was amplified from the vector pB330p6i2xoR-UcasW-U6Os4 (DNA Cloning Service) using the primers P773 and P774. The resulting PCR product was subsequently cloned into pICH41308 in a BbsI cut/ligation resulting in pICH41308_CDS1_Cas9 + Intron (V150). The egg-cell specific promoter was amplified from the vector pHEE2E-TRI (Wang et al., 2015) using the primers P769 and P770, and the resulting PCR product was cloned into

pICH41295 in a BbsI cut/ligation resulting in pICH412 95_Pro + 5U_egg-cell-specific (V140). Next, V140, V150, and pICH41276 were cloned into the recipient vector pICH47742 in a *Bsa*I cut/ligation leading to pICH47742_Cas9 + Intron_ pos2_fwd (V182). For the level_1 vector containing the tGFP under the control of the napinA promoter, first the napinA promoter had to be amplified from gDNA originating from *Brassica napus* using the primers P767 and P768. The resulting PCR product was then cloned into pICH41295 in a BbsI cut/ ligation resulting in pICH41295_Pro + 5U_napinA (V139). The tGFP first had to be amplified from pICSL50016 using the primers P626 and P627 and was then cloned into pICH41308 leading to pICH41308_tGFP (V134). Thereafter, V139, V134, and pICH41432 were cloned into pICH47761 in a BsaI cut/ligation resulting in pICH47761_napinA_tGFP_pos4_fwd (V181). The level_1 vector containing the guideArray under the control of the AtU6-26 promoter was generated by first amplifying the two parts of the guideArray using specific primers P293/P606 and P605/P272 with the plasmid pGTR (Xie et al., 2015; V113) as template, and subsequently fused together in a BsaI cut/ligation. This fused guideArray was again amplified using P294/P274 (as described in Xie et al., 2015). Subsequently, the guideArray was cloned into the MoClo compatible gRNA shuttle vector (J. Streubel, unpublished material) downstream of the ATU6-26 promoter in a BbsI-mediated cut/ligation reaction. This vector was cloned into pICH47751 in a BsaI cut/ligation resulting in pICH4 7751_org.dut1_guide_pos3_fwd. In a last step V183, V182, V181, pICH47751_org.dut1_guide_pos3_fwd and pICH41780 were cloned into the final recipient vector pAGM4723 in a BbsI cut/ligation resulting in pAGM4723_CRISPR_org.dut1 (H410).

In total ten transgenic T1 lines were obtained of which one (line 10) showed a heterozygous editing event in the gene encoding DUT1. The progeny of this line was tested for the absence of the CRISPR construct, but still had to have the editing event.

For the chloroplast- or mitochondria-specific complementation experiment of *tk1b*, an inactive and active version of TK1b, lacking the subcellular localization sequence, were needed. Additionally, both versions had to be made compatible with the MoClo system by removing a *Bbs*I recognition site inside the CDS. By introducing a missense mutation using primer-directed mutagenesis, an E codon was changed to a Q codon at position 163 (corresponding to E98 in human TK1; Welin et al., 2004), resulting in a CDS for the inactive version of TK1b (TK1b$_{E163Q}$). The BbsI site was removed using a similar approach, but in this case, a sense mutation was introduced. To exclusively localize both versions to the chloroplast or mitochondria, the optimized CTP described by Shen et al. (2017) and the MTP of *Sc*CoxIV (Maarse et al., 1984) were added to the 5'-ends of the truncated sequences, respectively.

To generate the level_0 vector containing the inactive and truncated version of TK1b directed to the chloroplast, the following primers were used in a PCR on gDNA: P2137 and

P2138, P2139 and P2140, and P2141 and P2143. All three PCR products and the synthetic fragment of CTP were then cloned into pICH41308 in a BbsI cut/ligation resulting in pICH41308_CDS1_CTP_truncated_inactive_TK1b (H959). The level_0 vector containing the active and truncated version directed to the chloroplast was created using the following primers in a PCR on gDNA: P2143 and P2144, P2145, and P2146. The two resulting PCR products and the synthetic fragment of CTP (B9; Integrated DNA Technologies) were cloned into pICH41308 in a BbsI cut/ligation resulting in pICH41308_CDS1_CTP_truncated_active_TK1b (H960).

The level_0 vectors needed for the mitochondrial complementation were generated by first amplifying the ScCoxIV transit peptide with the primers P2316 and P2317 with pAGM1482 as template. The inactive and active truncated versions of TK1b were amplified using P2318 and P2319 on H959 or H960, respectively. The level_0 vector containing the inactive and truncated version directed to the mitochondrion was created by combining both PCR products in a BbsI cut/ligation with pICH41308, resulting in pICH41308_CDS1_ScCoxIV_truncated_inactive_TK1b (H1007). In case of the level_0 vector containing the active and truncated version of TK1b directed to the mitochondrion, both PCR products were combined with pICH41308 in a BbsI cut/ligation, resulting in pICH41308_CDS1_Sc CoxIV_truncated_active_TK1b (H1008).

For creating the level_1 vectors, the vectors H959, H960, H1007 and H1008 were individually combined with pICH42211, pICH44179, and pICH44300 and cloned into pICH47732 in a BsaI cut/ligation, leading to the vectors pICH47732_CTP_inactive_TK1b_pos1_fwd (H962), pICH477 32_CTP_active_TK1b_pos1_fwd (H963), pICH47732_ScC oxIV_inactive_TK1b_pos1_fwd (H1015) and pICH47732_Sc CoxIV_active_TK1b_pos1_fwd (H1016).

In a last step, the level_1 vectors were individually combined with a level_1 vector mediating a phosphinothricin resistance for later plant selection (V166, Niehaus et al., 2020) and pICH41744 and cloned into pAGM4723 in a BbsI cut/ligation, resulting in the level_2 vectors pAGM47 23_CTP_inactive_TK1b_Basta (H965), pAGM4723_CTP_ active_TK1b_Basta (H966), pAGM4723_ScCoxIV_inactive_ TK1b_Basta (H1018), and pAGM4723_ScCoxIV_active_T K1b_Basta (H1019).

For each final construct, 12–15 T1 lines were selected for Basta resistance and subsequently checked for homozygosity in the T2 generation. For each construct two independent homozygous T2 lines were selected (H965: line 1.1 and 2.4; H966: line 1.4 and 2.1; H1018: line 1.6 and 2.1; H1019: line 1.2 and 2.1).

The vectors needed for the subcellular localization of the C-terminal eYFP tagged DHFR-TS2 (AT4G34570), TK1b (AT5G23070), DCD (AT3G48540), TMPK (AT5G59440), and DUT1 (AT3G46940) were generated by amplifying the sequences with cDNA or gDNA (only *TMPK*) as template with P149/P150 (*DHFR-TS2*), P218/P219 (*TK1b*), N398/N399 (*DCD*), P1698/P1699 (*TMPK*), and P607/P608 (*DUT1*),

respectively. The resulting PCR products and the recipient vector pXCS-eYFP (V36; Dahncke and Witte, 2013) were digested with EcoRI and XmaI and ligated, resulting in pXCS_eYFP-Thy2 (H26), pXCS_eYFP_TK1b (H76), pXCS_eY FP_dCMP-deam (X165), pXCS_eYFP_DUT1 (H239), and pXCS_eYFP_org.dut1 (H1210). In the case of TMPK, the PCR product and V36 were digested with HindIII and XmaI and ligated, resulting in pXCS_eYFP_TMPK (H875). For each construct five T1 lines were selected for Basta resistance. For confocal microscopy analysis, plants of the T2 generation were pretested for the presence of the eYFP signal and subsequently used for confocal microscopy.

Estimation of the amount of dTTP needed for cpDNA replication during germination

It is assumed that a seed has roughly as many cells as an Arabidopsis embryo 9 days after pollination ($\sim15 \times 10^3$ cells per embryo; Kiyosue et al., 1999) and contains on average of seven chloroplasts per cell as described previously for an embryo of the same age (Mansfield and Briarty, 1992). The cpDNA copy number per plastid in a mesophyll cell is 20–35 (Zoschke et al., 2007); however, this number is highly variable especially in plant development (Oldenburg and Bendich, 2015; Dobrogojski et al., 2020). Here we assumed that there is only one cpDNA copy per plastid, which is likely an underestimation. Following these assumptions, in one seed 105×10^3 cpDNA copies are present. Furthermore, we assumed that the cpDNA copy number increases 5 times during germination as reported for proplastid maturation in meristems (Fujie et al., 1994). Considering the size (154,478 bp) and the GC content (36%) of cpDNA, 9.84×10^4 molecules of dTTP are needed to replicate one plastidic genome. Thus, for one seed $\frac{(105 \times 10^3) \times (9.84 \times 10^4) \times 5}{6.022 \times 10^{23} \text{ (Avogadro constant)}}$ 0.086 pMol dTTP are required for a five-fold increase of cpDNA amount. Assuming that 1 g of seeds (the reference unit used for nucleotide quantification) contains 5×10^4 seeds (Jako et al., 2001), 0.086 pMol $\times 5 \times 10^4 = 4,300$ pmol g^{-1} dTTP and an equivalent amount of dATP are required for cpDNA synthesis. Similarly, for dGTP and dCTP $\frac{(105 \times 10^3) \times (5.61 \times 10^4) \times 5}{6.022 \times 10^{23} \text{ (Avogadro constant)}}$ = 0.049 pMol per seed and 2,450 pmol g^{-1} dGTP and dCTP are required for replicating the cpDNA.

Nucleic acid isolation and cDNA preparation

For the isolation of gDNA and total RNA, ~50 mg plant material was harvested at the respective time points and transferred to a 2-mL safe-lock tube containing five 2-mm steel beads and one 11-mm steel bead. Frozen samples were ground in a bead mill at 16 Hz for 4 min. For gDNA extraction, a CTAB-based method including an RNA-digestion step was used. Total RNA was extracted using the NucleoSpin RNA Plant (Machery-Nagel, Dueren, Germany) according to the manufacturer's protocol, but instead of the suggested RA1 buffer, the RAP buffer was used for the cell lysis. Synthesis of cDNA was done with the RevertAid First

Strand cDNA Synthesis Kit (Thermo Fisher, Waltham, MA, USA) according to the manufacturer's protocol. Total RNA (2 µg) per sample was transcribed using an oligo dT primer.

qPCR analysis

All qPCR experiments were performed using QuantStudio3 (Thermo Fisher, Waltham, MA, USA) in combination with qPCRBIO SyGreen Mix (PCR Biosystems, London, UK), both used according to the manufacturer's protocol. Each reaction had a total volume of 10 µL. The following program was used: 3 min at 95°C, 40 cycles of 5 s at 95°C followed by 30 s at 60°C. After the PCR Stage the Melt Curve Stage was done by heating the sample to 95°C for 15 s, holding it at 60°C for 1 min and slowly heating it again to 95°C at a rate of 0.1°C s^{-1}. Melting curves showed a single peak for all amplified PCR products. For each experiment three technical replicates per sample and three biological replicates were analyzed. To determine the relative genome abundance of the mtDNA and cpDNA, four different primer pairs binding specifically to the plastid or mitochondrial genomes were used. The amplifications were compared using Ct values to that obtained from primer pair binding to the nuclear genome. The signals shown in Figures 6, 8, and 9 and Supplemental Figure S9 were obtained by amplifying RBCL (AtCG00490) in the case of the plastid genome, COX1 (AtMG01360) for the mitochondrial genome, and UBC21 (At5G25760) as a reference for the nuclear genome. The primers used for qPCR were already published and are listed in Supplemental Data Set 1 with their corresponding citations. Results were analyzed according to Livak and Schmittgen (2001), and for determination of the relative genome abundance the $2^{-\Delta Ct}$ method was used.

Confocal microscopy and image analysis

For subcellular localization, all constructs except pXCS_eYFP_TMPK (H875), were stably transformed into A. thaliana. Samples were analyzed by confocal microscopy as previously described by Dahncke and Witte (2013). Mitochondria of seedling roots were labeled in vivo with tetramethylrhodamine methyl ester (TMRM) as described in Niehaus et al. (2020). Both TMRM and the ScCOXIV:mCherry show unspecific labeling, TMRM also stains the cell wall, whereas ScCOXIV:mCherry is also localized to the nucleus, most likely due to high overexpression. Microscopy was performed using a Leica SP8 confocal laser microscope with an HC PL APO 40x/1.10 water immersion objective (Leica Microsystems, Wetzlar, Germany). The settings for the respective fluorescent proteins/chemicals were as following. eYFP: excitation 514 nm, collection bandwidth 519–539 nm, laser intensity 1%–2%, gain 30%. mCherry: excitation 552 nm, collection bandwidth 600–620 nm, laser intensity 0.5%–1%, and gain 50%. TMRM: excitation 552 nm, collection bandwidth 570–620 nm, laser intensity 1%, and gain 30%.

The confocal microscopy images of the co-localized candidates DCD and DHFR-TS2 were analyzed using the "Just Another Colocalisation Plugin" in ImageJ (Bolte and Cordelieres, 2006). Here the Pearson's coefficient and Van

Steensel's crosscorrelation function (Van Steensel et al., 1996) were calculated for several images using a dx range from −150 to 150 pixels. For images taken in leaves, the signal of the eYFP fusion proteins was compared to the chloroplastic autofluoresence. In case of images taken in root tissue, incubated with TMRM, the signal of the eYFP fusion protein was compared to the mitochondrial TMRM signal. For the analysis of the dual-localized proteins TK1b, DUT1$_{org}$, and TMPK, the "Plot Profile" function of ImageJ was used (Sharma et al., 2018). The analyzed region of the respective candidates was indicated by a dotted white line in Supplemental Figure S3.

Metabolome analysis

Seedlings of the respective time points were completely transferred from the soaked filter paper using a flat spatula and forceps and transferred to a 2-mL safe-lock tube containing five 2-mm steel beads and one 11-mm steel bead and rapidly cooled down in liquid nitrogen. Subsequently, samples could be stored at −80°C for several weeks or were further processed by grinding them in a frozen state in a bead mill for 4 min at 16 Hz twice with cooling in liquid nitrogen between grinding steps. Further steps were performed according to Straube et al. (2021a) with the exception that the bead mill step after addition of the extraction buffer was replaced with two rounds of vortexing at maximum speed for 12 s. For the quantification of nucleotides, a modified gradient for the Hypercarb chromatography was used: 0 min, 100% A; 3 min, 100% A; 18 min, 70% A; 19 min, 0% A; 22 min, 0% A; 22.5 min, 100% A until 30 min. The mobile phase A consisted of 5-mM ammonium acetate, pH 9.5, in deionized water and the mobile phase B consisted of 100% acetonitrile. The chromatography was performed employing a flowrate of 0.6 mL min^{-1} and a column temperature of 35°C. The injection volume was increased to 20 µL.

Nucleosides were analyzed using an Agilent 1290 Infinity II LC System coupled with an Agilent 6470 triple quadrupole mass spectrometer. Chromatographic separation was achieved using a 4.6 × 50 mm Polaris C18-A column with 3-µm particle size (Agilent, Palo Alto, CA, USA). The flowrate was 0.6 mL min^{-1}, the column temperature was set to 30°C and the injection volume was 10 µL. Mobile phase A was 0.1% (v/v) formic acid in water and mobile phase B was 0.1% (v/v) formic acid in methanol. The following gradient was used: 0 min, 96% A; 8.00 min, 35% A; 8.20 min, 100% A; 10.00 min, 100% A; 10.10 min, 96% A until 12.50 min. In-source parameters used are: Gas temperature 250°C, gas flow 12 L min^{-1}, nebulizer pressure 35 psi, sheath gas temperature 395°C, sheath gas flow 12 L min^{-1}, capillary voltage 4,000 V, and nozzle voltage 500 V. All nucleosides were measured in positive mode using the multiple-reaction-monitoring mode. Transitions (precursor ions and product ions) and fragmentor energies for nucleoside analysis are the same as described in Straube et al. (2021a).

The isotope standard for dC was generated by complete dephosphorylation of a dCTP isotope standard (Eurisotope, Saint-Aubin, France) using Shrimp Alkaline Phosphatase (New England BioLabs, Ipswich, MA, USA). Full conversion was confirmed by LC–MS analysis.

The metabolite quantification refers to the weight of the starting material (dry seeds). For each sample a similar amount (10 ± 0.5 mg) of dry seeds was weighed and the weight was noted and used for calculation. For each time point all material was harvested and used for metabolite extraction. Thus, the values in the time course reflect the total amount of the indicated metabolite present in seedlings or germinated seeds that originate from one gram of dry seeds. This procedure eliminates a bias due to the different uptake of water in the individual phases of germination and seedling establishment.

Only for the comparison of genotypes at the late seedling establishment time point (144 h; Figure 10 and Supplemental Figure S12) ca. 100 mg of seedlings (fresh weight) were used for metabolite extraction. Here, the metabolite concentration refers to the fresh weight of the respective seedlings.

Statistical analysis

Statistical analysis was performed as described by Heinemann et al. (2021). Here, R Software version 4.1.1 was used together with RStudio version 1.4.1717 and CRAN packages multcomp and sandwich to perform two-sided Tukey's pairwise comparisons. Heteroscedasticity of the dataset was considered with the sandwich variance estimator (Herberich et al., 2010; Pallmann and Hothorn, 2016). The adjusted P-values of the individual datasets are listed in Supplemental Data Set 2.

Accession numbers

Information regarding used mutants can be found in the GenBank/EMBL data libraries under the following accession numbers: *DHFR-TS1* (AT2G16370), *DHFR-TS2* (AT4G34570), *TK1a* (AT3G07800), *TK1b* (AT5G23070), *DCD* (AT3G48540), *DUT1* (AT3G46940), *TMPK* (AT5G59440), *ATM-1* (AT3G48190), *ATR-2* (AT5G40820), *CEN2* (AT4G37010), *DDB1-A* (AT4G05420), and *RAD1* (AT5G41150)

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. Gene expression profiles of genes putatively involved in thymidylate formation.

Supplemental Figure S2. Generation of the *dut1$_{org}$* mutant line.

Supplemental Figure S3. Subcellular localization of DHFR-TS2, TK1b, DCD, and TMPK.

Supplemental Figure S4. Quantitative image analysis for colocalization studies.

Supplemental Figure S5. Determination of the relative cotyledon growth in wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis in established seedlings.

Supplemental Figure S6. Recovery of *dhfr-ts2*, *tk1b*, and *dhfr-ts2 tk1b* mutants in comparison to wild-type.

Supplemental Figure S7. Concentrations of ribonucleotides and ribonucleosides at different time points during seed germination and in the establishing seedling.

Supplemental Figure S8. Relative amounts of xanthosine in wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis during germination.

Supplemental Figure S9. Abundance of the mtDNA relative to the ncDNA in wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis during germination and early and late seedling establishment.

Supplemental Figure S10. Concentrations of ribonucleotide triphosphates in the wild-type and mutant seeds lacking enzymes putatively involved in thymidylate homeostasis during early seedling establishment.

Supplemental Figure S11. Relative abundances of cpDNA and mtDNA during early seedling establishment determined by the amplification of four genes, respectively.

Supplemental Figure S12. Concentrations of ribonucleotide triphosphates in the wild-type and mutants lacking enzymes putatively involved in thymidylate homeostasis during late seedling establishment.

Supplemental Table S1. Abbreviations.

Supplemental Data Set 1. List of all oligonucleotides used in this study

Supplemental Data Set 2. Adjusted *P*-values for all multiple comparisons

References

Anderson DD, Quintero CM, Stover PJ (2011) Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. Proc Natl Acad Sci USA **108**: 15163–15168

Ashihara H, Crozier A, Ludwig IAA (2020) Plant Nucleotide Metabolism: Biosynthesis, Degradation, and Alkaloid Formation. Wiley Blackwell, Chichester

Barroco RM, van Poucke K, Bergervoet JH, Veylder L, de Groot SP, Inze D, Engler G (2005) The role of the cell cycle machinery in resumption of postembryonic development. Plant Physiol **137**: 127–140

Baumann H, Hofmann R, Lammers M, Schimpff-Weiland G, Follmann H (1984) Aurintricarboxylic acid and polynucleotides as novel inhibitors of ribonucleotide reductases. Zeitschrif Naturforsch C Biosci **39**: 276–281

Bitter EE, Townsend MH, Erickson R, Allen C, O'Neill KL (2020) Thymidine kinase 1 through the ages: a comprehensive review. Cell Biosci **10**: 1–16

Bolte S, Cordelieres FP (2006) A guided tour into subcellular colocalization analysis in light microscopy. J Microsc **224**: 213–232

Bray CM, West CE (2005) DNA repair mechanisms in plants: crucial sensors and effectors for the maintenance of genome integrity. New Phytologist **168**: 511–528

Bryant JA (1980) Biochemical aspects of DNA replication with particular reference to plants. Biol Rev Cambridge Philos Soc **55**: 237–284

Buckland RJ, Watt DL, Chittoor B, Nilsson AK, Kunkel TA, Chabes A (2014) Increased and imbalanced dNTP pools symmetrically promote both leading and lagging strand replication infidelity. PLoS Genet **10**: e1004846

Castroviejo M, Tharaud D, Mocquot B, Litvak S (1979) Factors affecting the onset of deoxyribonucleic acid synthesis during wheat embryo germination. Study of the changes in DNA polymerases A, B and C and the pool of DNA precursors. Biochem J **181**: 193–199

Chen M, Thelen JJ (2011) Plastid uridine salvage activity is required for photoassimilate allocation and partitioning in Arabidopsis. Plant Cell **23**: 2991–3006

Clausen AR, Girandon L, Ali A, Knecht W, Rozpedowska E, Sandrini MPB, Andreasson E, Munch-Petersen B, Piškur J (2012) Two thymidine kinases and one multisubstrate deoxyribonucleoside kinase salvage DNA precursors in *Arabidopsis thaliana*. FEBS J **279**: 3889–3897

Corral MG, Haywood J, Stehl LH, Stubbs KA, Murcha MW, Mylne JS (2018) Targeting plant DIHYDROFOLATE REDUCTASE with antifolates and mechanisms for genetic resistance. Plant J **95**: 727–742

Dahncke K, Witte CP (2013) Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in Arabidopsis. Plant Cell **25**: 4101–4109

Dellaquila A, Lioi L, Scarascia I (1980) Deoxyribonucleic-acid synthesis and deoxyribonucleic-acid polymerase-activity during early germination of wheat embryos at high and low viability. Biol Plant **22**: 287–293

Dobrogojski J, Adamiec M, Lucinski R (2020) The chloroplast genome: a review. Acta Physiol Plant **42**: 98

Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol **32**: 1262-U130

Dubois E, Cordoba-Canero D, Massot S, Siaud N, Gakiere B, Domenichini S, Guerard F, Roldan-Arjona T, Doutriaux MP (2011) Homologous recombination is stimulated by a decrease in dUTPase in Arabidopsis. PLoS One **6**: e18658

Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JDG, Patron NJ, Marillonnet S (2014) A golden gate modular cloning toolbox for plants. ACS Synth Biol **3**: 839–843

Franzolin E, Pontarin G, Rampazzo C, Miazzi C, Ferraro P, Palumbo E, Reichard P, Bianchi V (2013) The deoxynucleotide triphosphohydrolase SAMHD1 is a major regulator of DNA precursor pools in mammalian cells. Proc Natl Acad Sci USA **110**: 14272–14277

Fujie M, Kuroiwa H, Kawano S, Mutoh S, Kuroiwa T (1994) Behavior of organelles and their nucleoids in the shoot apical meristem during leaf development in Arabidopsis-thaliana l. Planta **194**: 395–405

Garcia V, Bruchet H, Camescasse D, Granier F, Bouchez D, Tissier A (2003) AtATM is essential for meiosis and the somatic response to DNA damage in plants. Plant Cell **15**: 119–132

Gorelova V, Lepeleire J, de van Daele J, Pluim D, Mei C, Cuypers A, Leroux O, Rebeille F, Schellens JHM, Blancquaert D, et al. (2017) Dihydrofolate reductase/thymidylate synthase fine-tunes the folate status and controls redox homeostasis in plants. Plant Cell **29**: 2831–2853

Guillet M, van der Kemp PA, Boiteux S (2006) dUTPase activity is critical to maintain genetic stability in Saccharomyces cerevisiae. Nucleic Acids Res **34**: 2056–2066

Heinemann KJ, Yang SY, Straube H, Medina-Escobar N, Varbanova-Herde M, Herde M, Rhee S, Witte CP (2021) Initiation of cytosolic plant purine nucleotide catabolism involves a monospecific xanthosine monophosphate phosphatase. Nat Commun **12**: 6846

Herberich E, Sikorski J, Hothorn T (2010) A robust procedure for comparing multiple means under heteroscedasticity in unbalanced designs. PLoS One **5**: e9788

Jako C, Kumar A, Wei YD, Zou JT, Barton DL, Giblin EM, Covello PS, Taylor DC (2001) Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. Plant Physiol **126**: 861–874

Kiyosue T, Ohad N, Yadegari R, Hannon M, Dinneny J, Wells D, Katz A, Margossian L, Harada JJ, Goldberg RB, et al. (1999) Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. Proc Natl Acad Sci **96**: 4186–4191

Kohalmi SE, Glattke M, Mcintosh EM, Kunz BA (1991) Mutational specificity of dna precursor pool imbalances in yeast arising from deoxycytidylate deaminase deficiency or treatment with thymidylate. J Mol Biol **220**: 933–946

Kohler RH, Zipfel WR, Webb WW, Hanson (1997) The green fluorescent protein as a marker to visualize plant mitochondria in vivo. Plant J **11**: 613–621

Law SR, Narsai R, Whelan J (2014) Mitochondrial biogenesis in plants during seed germination. Mitochondrion **19 Pt B**: 214–221

Le Ret M, Belcher S, Graindorge S, Wallet C, Koechler S, Erhardt M, Williams-Carrier R, Barkan A, Gualberto JM (2018) Efficient replication of the plastid genome requires an organellar thymidine kinase. Plant Physiol **178**: 1643–1656

Lincker F, Philipps G, Chabouté ME (2004) UV-C response of the ribonucleotide reductase large subunit involves both E2F-mediated gene transcriptional regulation and protein subcellular relocalization in tobacco cells. Nucleic Acids Res **32**: 1430–1438

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, CA) **25**: 402–408

Maarse AC, Adolphus Vanloon PG, Riezman H, Gregor I, Schatz G, La Grivell (1984) Subunit-iv of yeast cytochrome-c oxidase - cloning and nucleotide sequencing of the gene and partial amino-acid sequencing of the mature protein. EMBO J **3**: 2831–2837

Mansfield SG, Briarty LG (1992) Cotyledon cell development in *Arabidopsis thaliana* during reserve deposition. Can J Bot **70**: 151–164

Marciniak B, Bucholc M, Buchowicz J (1987) Early DNA-synthesis during the germination of wheat embryos. Phytochemistry **26**: 331–334

Martínez-Arribas B, Requena CE, Pérez-Moreno G, Ruíz-Pérez LM, Vidal AE, González-Pacanowska D (2020) DCTPP1 prevents a mutator phenotype through the modulation of dCTP, dTTP and dUTP pools. Cell Mol Life Sci **77**: 1645–1660

Masubelele NH, Dewitte W, Menges M, Maughan S, Collins C, Huntley R, Nieuwland J, Scofield S, Murray JA (2005) D-type cyclins activate division in the root apex to promote seed germination in Arabidopsis. Proc Natl Acad Sci USA **102**: 15694–15699

Molinier J, Lechner E, Dumbliauskas E, Genschik P (2008) Regulation and role of Arabidopsis CUL4-DDB1A-DDB2 in maintaining genome integrity upon UV stress. PLoS Genet **4**: e1000093

Molinier J, Ramos C, Fritsch O, Hohn B (2004) CENTRIN2 modulates homologous recombination and nucleotide excision repair in Arabidopsis. Plant Cell **16**: 1633–1643

Mori Y, Kimura S, Saotome A, Kasai N, Sakaguchi N, Uchiyama Y, Ishibashi T, Yamamoto T, Chiku H, Sakaguchi K (2005) Plastid

DNA polymerases from higher plants, *Arabidopsis thaliana*. Biochem Biophys Res Commun **334**: 43–50

Nagy GN, Leveles I, Vertessy BG (2014) Preventive DNA repair by sanitizing the cellular (deoxy) nucleoside triphosphate pool. FEBS J **281**: 4207–4223

Narsai R, Law SR, Carrie C, Xu L, Whelan J (2011) In-depth temporal transcriptome profiling reveals a crucial developmental switch with roles for RNA processing and organelle metabolism that are essential for germination in Arabidopsis. Plant Physiol **157**: 1342–1362

Nelson BK, Cai X, Nebenfuhr A (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. Plant J Cell Mol Biol **51**: 1126–1136

Niehaus M, Straube H, Kuenzler P, Rugen N, Hegermann J, Giavalisco P, Eubel H, Witte CP, Herde M (2020) Rapid affinity purification of tagged plant mitochondria (Mito-AP) for metabolome and proteome analyses. Plant Physiol **182**: 1194–1210

Niu M, Wang Y, Wang C, Lyu J, Wang Y, Dong H, Long W, Di Wang, Kong W, Wang L, et al. (2017) ALR encoding dCMP deaminase is critical for DNA damage repair, cell cycle progression and plant development in rice. J Exp Bot **68**: 5773–5786

Nordlund P, Reichard P (2006) Ribonucleotide reductases. Ann Rev Biochem **75**: 681–706

Oldenburg DJ, Bendich AJ (2015) DNA maintenance in plastids and mitochondria of plants. Front Plant Sci **6**: 883

Osaki Y, Kodama Y (2017) Particle bombardment and subcellular protein localization analysis in the aquatic plant Egeria densa. PeerJ **5**: e3779

Pallmann P, Hothorn LA (2016) Analysis of means: a generalized approach using R. J Appl Stat **43**: 1541–1560

Palmieri F, Rieder B, Ventrella A, Blanco E, Do PT, Nunes-Nesi A, Trauth AU, Fiermonte G, Tjaden J, Agrimi G, et al. (2009) Molecular identification and functional characterization of *Arabidopsis thaliana* mitochondrial and chloroplastic NAD(+) carrier proteins. J Biol Chem **284**: 31249–31259

Parent JS, Lepage E, Brisson N (2011) Divergent roles for the two poll-like organelle DNA polymerases of Arabidopsis. Plant Physiol **156**: 254–262

Paszkiewicz G, Gualberto JM, Benamar A, Macherel D, Logan DC (2017) Arabidopsis seed mitochondria are bioenergetically active immediately upon imbibition and specialize via biogenesis in preparation for autotrophic growth. Plant Cell **29**: 109–128

Pedroza-García JA, Nájera-Martínez M, de la Paz Sanchez M, Plasencia J (2015) *Arabidopsis thaliana* thymidine kinase 1a is ubiquitously expressed during development and contributes to confer tolerance to genotoxic stress. Plant Mol Biol **87**: 303–315

Pedroza-García JA, Nájera-Martínez M, Mazubert C, Aguilera-Alvarado P, Drouin-Wahbi J, Sánchez-Nieto S, Gualberto JM, Raynaud C, Plasencia J (2019) Role of pyrimidine salvage pathway in the maintenance of organellar and nuclear genome integrity. Plant J Cell Mol Biol **97**: 430–446

Rajjou L, Duval M, Gallardo K, Catusse J, Bally J, Job C, Job D (2012) Seed germination and vigor. Ann Rev Plant Biol **63**: 507–533

Rinne J, Witte CP, Herde M (2021) Loss of MAR1 function is a marker for co-selection of CRISPR-induced mutations in plants. Front Genome Ed **3**: 723384

Ronceret A, Gadea-Vacas J, Guilleminot J, Lincker F, Delorme V, Lahmy S, Pelletier G, Chabouté ME, Devic M (2008) The first zygotic division in Arabidopsis requires de novo transcription of thymidylate kinase. Plant J Cell Mole Biol **53**: 776–789

Schimpff G, Müller H, Follmann H (1978) Age-dependent DNA labeling and deoxyribonucleotide synthesis in wheat seeds. Biochim Biophys Acta **520**: 70–81

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods **9**: 676–682

148

Schmidt TT, Sharma S, Reyes GX, Kolodziejczak A, Wagner T, Luke B, Hofer A, Chabes A, Hombauer H (2019) Inactivation of folylpolyglutamate synthetase Met7 results in genome instability driven by an increased dUTP/dTTP ratio. Nucleic Acids Res **48**: 264–277

Sharma M, Bennewitz B, Kloesgen RB (2018) Dual or not dual? Comparative analysis of fluorescence microscopy-based approaches to study organelle targeting specificity of nuclear-encoded plant proteins. Front Plant Sci **9**: 1350

Shen BR, Zhu CH, Yao Z, Cui LL, Zhang JJ, Yang CW, He ZH, Peng XX (2017) An optimized transit peptide for effective targeting of diverse foreign proteins into chloroplasts in rice. Sci Rep **7**: 46231

Sliwinska E, Bassel GW, Bewley JD (2009) Germination of *Arabidopsis thaliana* seeds is not completed as a result of elongation of the radicle but of the adjacent transition zone and lower hypocotyl. J Exp Bot **60**: 3587–3594

Stasolla C, Loukanina N, Ashihara H, Yeung EC, Thorpe TA (2002) Pyrimidine nucleotide and nucleic acid synthesis in embryos and megagametophytes of white spruce (Picea glauca) during germination. Physiol Plant **115**: 155–165

Stillman B (2013) Deoxynucleoside triphosphate (dNTP) synthesis and destruction regulate the replication of both cell and virus genomes. Proc Natl Acad Sci USA **110**: 14120–14121

Straube H, Niehaus M, Zwittian S, Witte CP, Herde M (2021a) Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. Plant Cell **33**: 270–289

Straube H, Witte CP, Herde M (2021b) Analysis of nucleosides and nucleotides in plants: an update on sample preparation and LC-MS techniques. Cells **10**: 689

Strugala K, Buchowicz J (1984) The use of [deoxyadenosine-h-3] to measure the rate of dna-synthesis in germinating wheat embryos. Plant Sci Lett **34**: 17–23

Thornton JM, Collins ARS, Powell AA (1993) The effect of aerated hydration on DNA synthesis in embryos, of Brassica oleracea L. Seed Sci Res **3**: 195–199

Van Steensel B, van Binnendijk EP, Hornsby CD, van der Voort HT, Krozowski ZS, de Kloet ER, van Driel R (1996) Partial colocalization of glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons. J Cell Sci **109**: 787–792

Vértessy BG, Tóth J (2009) Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. Account Chem Res **42**: 97–106

Wang C, Liu Z (2006) Arabidopsis ribonucleotide reductases are critical for cell cycle progression, DNA damage repair, and plant development. Plant Cell **18**: 350–365

Wang ZP, Xing HL, Dong L, Zhang HY, Han CY, Wang XC, Chen QJ (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biol **16**: 144

Waterworth WM, Bray CM, West CE (2015) The importance of safeguarding genome integrity in germination and seed longevity. J Exp Bot **66**: 3549–3558

Weber E, Engler C, Gruetzner R, Werner S, Marillonnet S (2011) A modular cloning system for standardized assembly of multigene constructs. PLoS One **6**: e16765

Weitbrecht K, Mueller K, Leubner-Metzger G (2011) First off the mark: early seed germination. J Exp Bot **62**: 3289–3309

Welin M, Kosinska U, Mikkelsen NE, Carnrot C, Zhu C, Wang L, Eriksson S, Munch-Petersen B, Eklund H (2004) Structures of thymidine kinase 1 of human and mycoplasmic origin. Proc Natl Acad Sci USA **101**: 17970–17975

Witte CP, Herde M (2020) Nucleotide metabolism in plants. Plant Physiol **182**: 63–78

Xie K, Minkenberg B, Yang Y (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci USA **112**: 3570–3575

Xu J, Deng Y, Li Q, Zhu X, He Z (2014) STRIPE2 encodes a putative dCMP deaminase that plays an important role in chloroplast development in rice. J Genet Genom **41**: 539–548

Xu J, Zhang L, Yang DL, Li Q, He Z (2015) Thymidine kinases share a conserved function for nucleotide salvage and play an essential role in *Arabidopsis thaliana* growth and development. New Phytologist **208**: 1089–1103

Yoshiyama K, Conklin PA, Huefner ND, Britt AB (2009) Suppressor of gamma response 1 (SOG1) encodes a putative transcription factor governing multiple responses to DNA damage. Proc Natl Acad Sci USA **106**: 12843–12848

Zoschke R, Liere K, Boerner T (2007) From seedling to mature plant: Arabidopsis plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. Plant J **50**: 710–722

Zrenner R, Stitt M, Sonnewald U, Boldt R (2006) Pyrimidine and purine biosynthesis and degradation in plants. Ann Rev Plant Biol **63–57**: 805–836
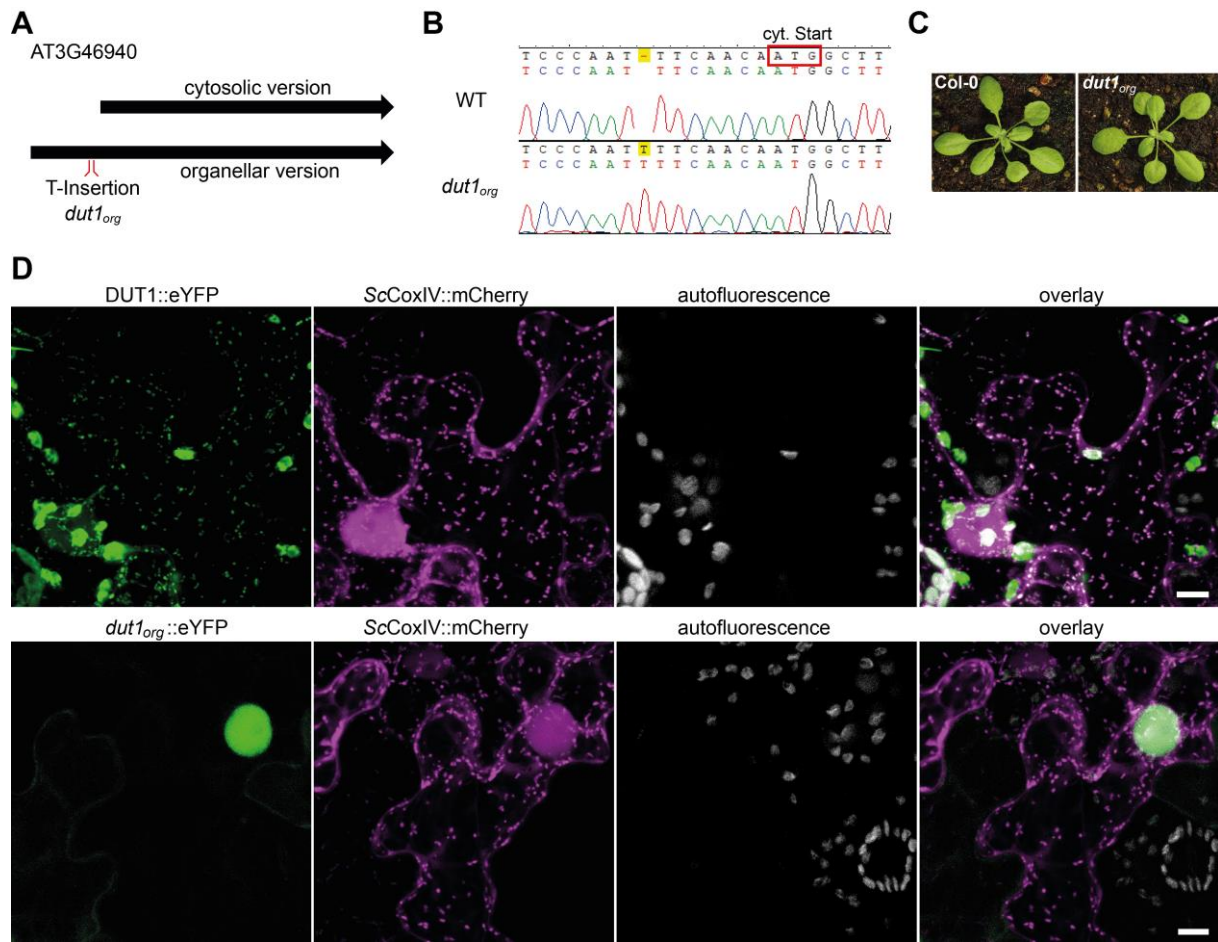
Figure S1. Generation of the *dut1org* mutant line. A mutant line lacking the organellar version of DUT1 was created with CRISPR/Cas. (**A**) The short and long ORFs coding for the cytosolic and the organellar DUT1 variant are schematically shown including the guideRNA targeting site and the editing event. (**B**) Confirmation of the editing event by cDNA amplification and subsequent sequencing of the PCR product. (**C**) Phenotypic comparison of a *dut1org* and a wildtype plant 20 days after transfer to growth conditions. (**D**) Confocal microscopy images showing the subcellular localization of DUT1 preduced from the long intact ORF (upper panels as in Figure 2) and the respective edited ORF (lower panels). Images were taken three days after agrobacteria infiltration of *N. benthamiana*. From left to right are shown: DUT1::eYFP, mitochondrial *Sc*COXIV::mCherry, chloroplast autofluorescence and an overlay of all three channels. Scale bars = 10 µm.

A   DHFR-TS2::eYFP   autofluorescence   overlay

B   DHFR-TS2::eYFP   TMRM   overlay

C   TK1b::eYFP   autofluorescence   overlay

D   TK1b::eYFP   TMRM   overlay

E   DCD::eYFP   autofluorescence   overlay

F   DCD::eYFP   TMRM   overlay

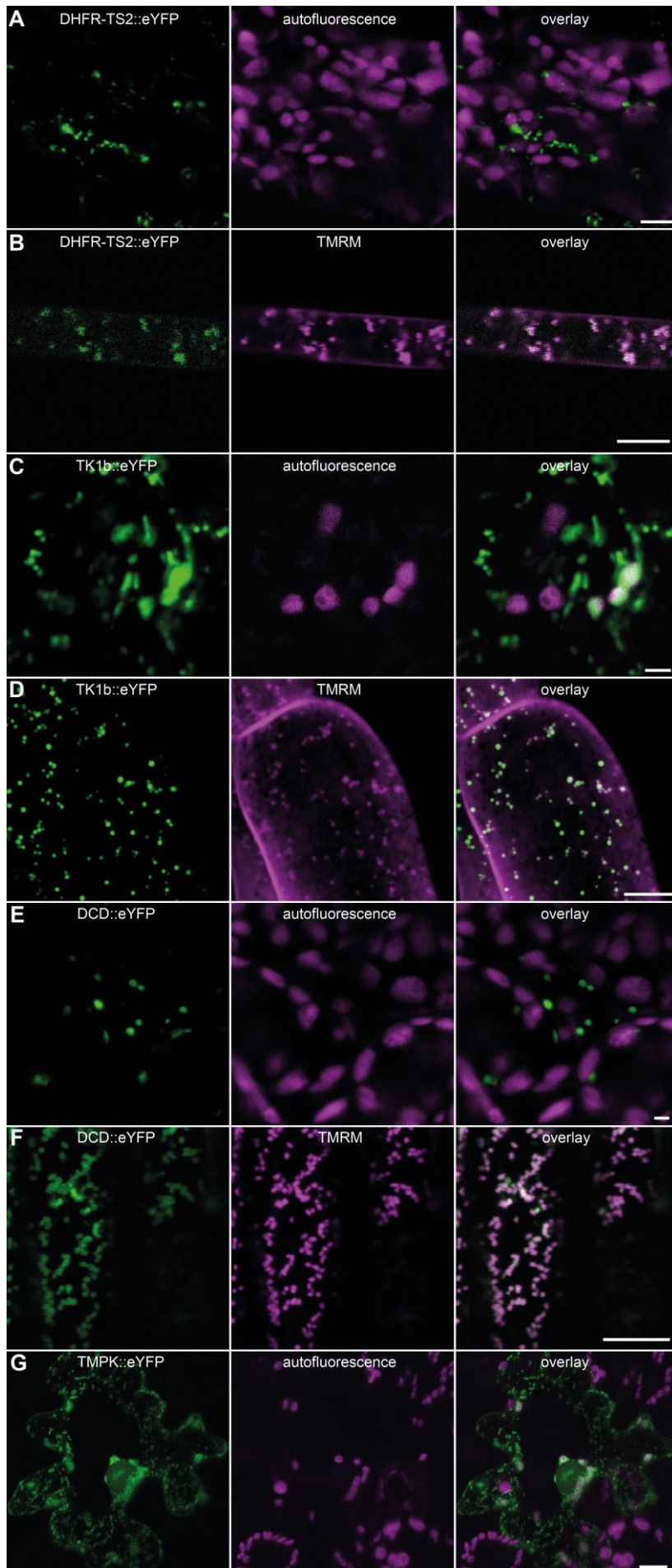G   TMPK::eYFP   autofluorescence   overlay

Figure S2. Subcellular localization of DHFR-TS2, TK1b, DCD and TMPK. Stable transformed Arabidopsis seeds expressing C-terminal eYFP fusions with DHFR-TS2, TK1b and DCD, respectively, were imbided in water for 48 h at 4°C in the dark and transferred for 48 h to growth conditions. Confocal microscopy images of cotyledon tissues were taken for (**A**) DHFR-TS2 (**C**) TK1b and (**E**) DCD. Images show (from left to right): the eYFP signal, chloroplast autofluorescence and an overlay of both channels. Confocal microscopy images of root tissue were taken for (**B**) DHFR-TS2 (**D**) TK1b and (**F**) DCD. Images show (from left to right): eYFP signal, signal from TMRM staining indicating the presence of mitochondria and an overlay of both channels. (**G**) Confocal microcopy images of a *N. benthamiana* leaf transiently expressing the TMPK::eYFP construct. Images were taken three days after agrobacteria infiltration. Images represent (from left to right): eYFP signal, chloroplast autofluorescence and the overlay of both channels. Scale bars = (**A**) 5 µm, (**B**) 8 µm, (**C**) 3 µm, (**D**) 8 µm, (**E**) 3 µm, (**F**) 10 µm and (**G**) 10 µm.
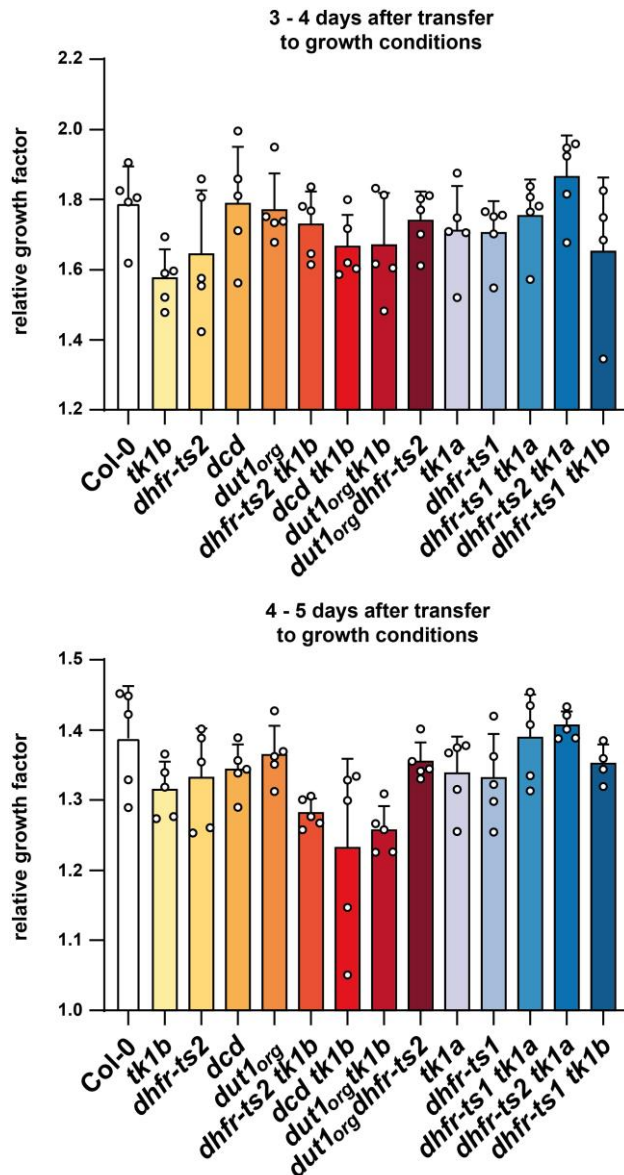
Figure S3. Determination of the relative cotyledon growth in wildtype and mutants lacking enzymes putatively involved in thymidylate homeostasis in established seedlings. Seeds were imbibed in water for 48 h at 4°C in the dark and transferred to growth conditions. Images were taken at three, four and five days and the sum of yellow and green pixels was determined and used for the calculation of the relative growth rate of the cotyledons from individual plants for the time period of three to four days and four to five days. Four to five biological replicates (pooled seedlings) were analyzed. Error bars are SD. No significant differences ($p < 0.05$) to Col-0 were observed. Adjusted p-values for multiple comparisons can be found in Table S2.
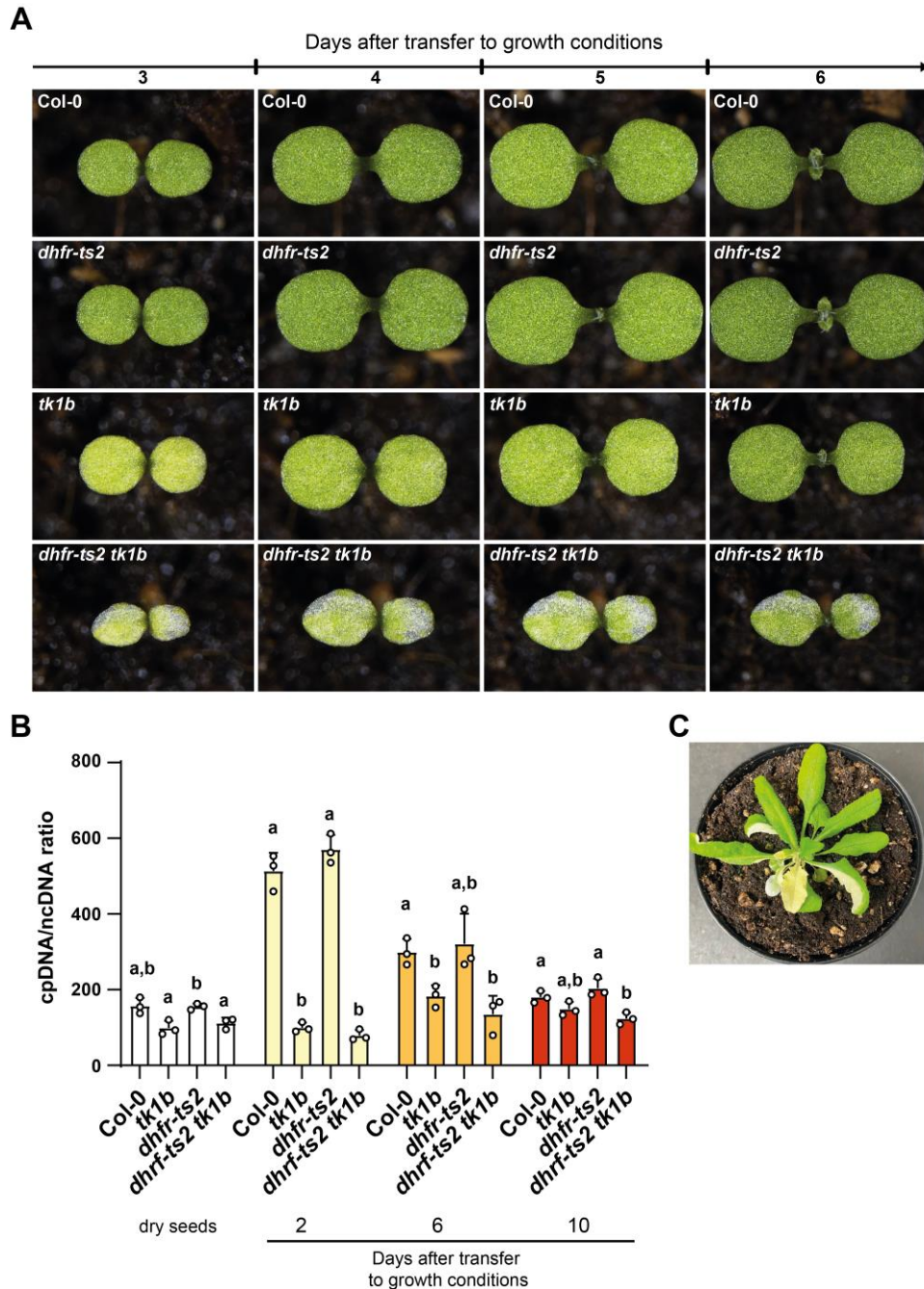
Figure S4. Recovery of *dhfr-ts2*, *tk1b*, *dhfr-ts2 tk1b* mutants in comparison to wildtype. Seeds were imbibed in water for 48 h at 4°C in the dark and transferred to growth conditions. (**A**) Images were taken of one specific plant at day three to six and are representative for 4 to 5 documented individuals. (**B**) One gene from cpDNA and ncDNA was amplified by q-PCR, respectively, and cpDNA/ncDNA ratios were calculated. **(C)** Image of the rarely occurring albinotic section phenotype of *dhfr-ts tk1b*. Different letters indicate significant differences ($p < 0.05$) between genotypes. Three biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted *p*-values for multiple comparisons can be found in Table S2.
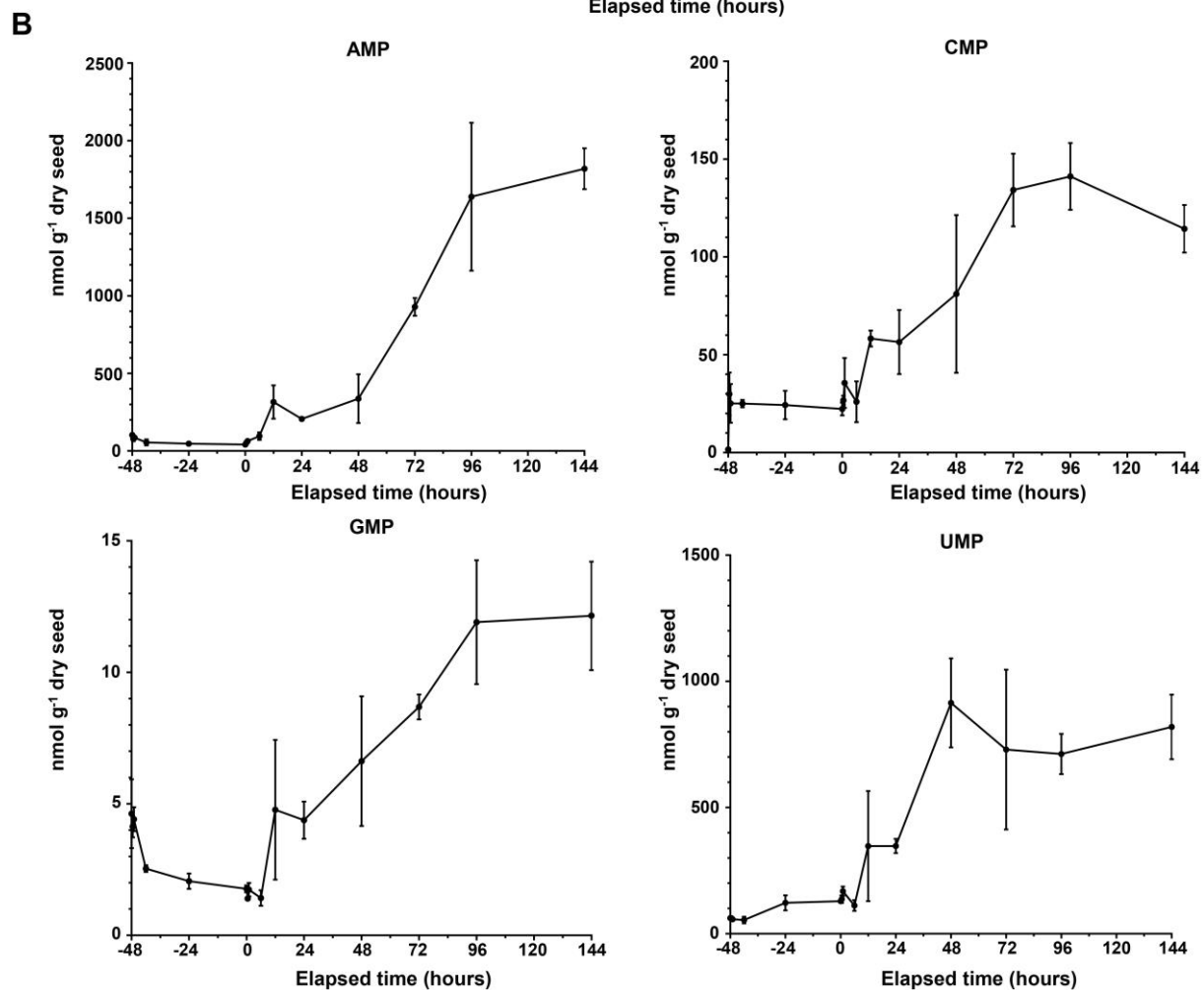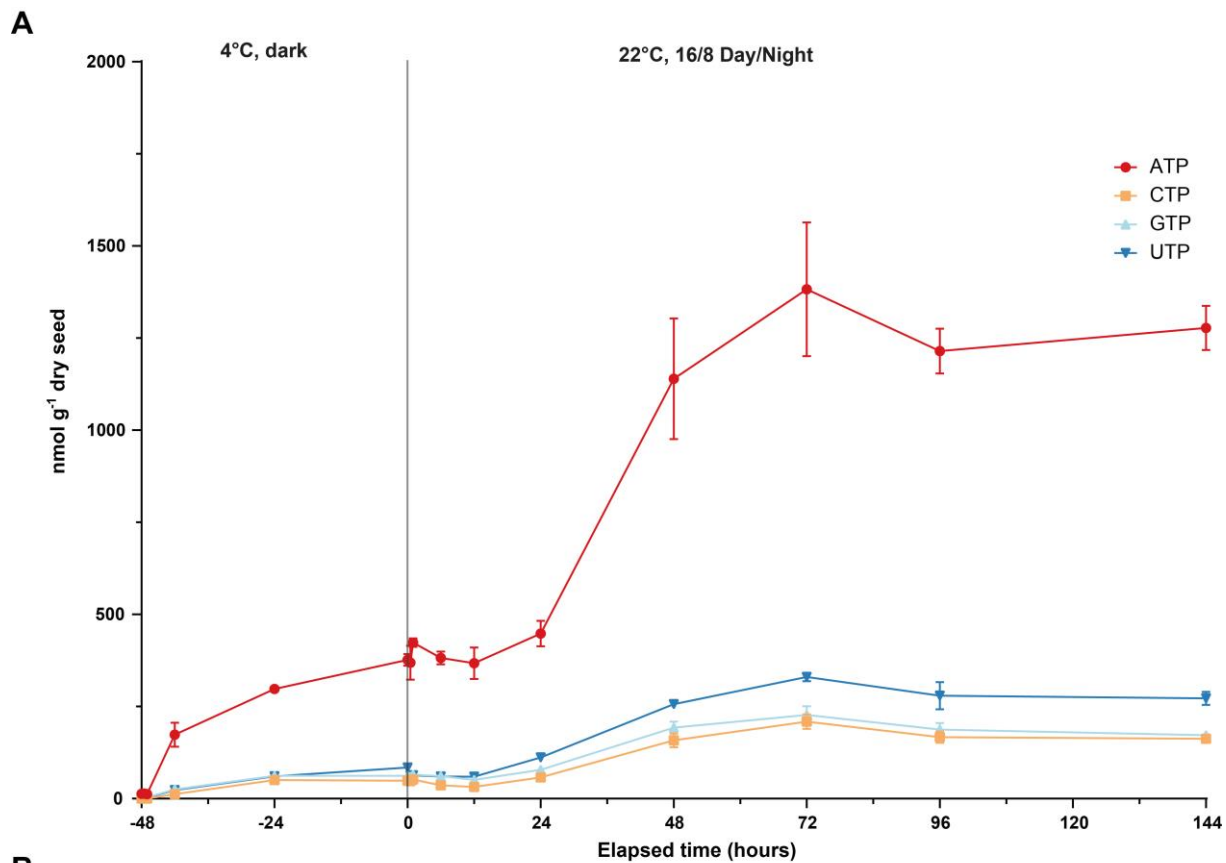
Figure S5. Concentrations of ribonucleotides and ribonucleosides at different time points during seed germination and in the establishing seedling. Concentrations of (**A**) ATP, CTP, GTP, UTP, (**B**) AMP, CMP, GMP and UMP were quantified with LC-MS before transfer to growth conditions (-48 h, -47.5 h, -42 h, -24 h; dark and 4°C), at transfer to growth conditions (0 h) and after transfer to growth conditions (0.5 h, 1 h, 6h, 12 h, 24 h, 48 h, 72 h, 96 h, 144 h). While seeds at the -48 h time point were dry all other samples were imbibed in water. Four biological replicates (pooled seeds/seedlings) were analyzed. Error bars are SD.
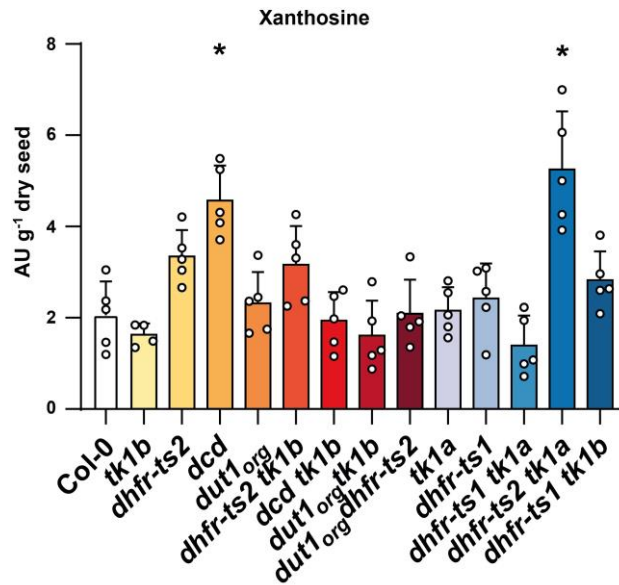
Figure S6. Relative amounts of xanthosine in wildtype and mutants lacking enzymes putatively involved in thymidylate homeostasis in early germination. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 3 h to growth conditions and harvested for LC-MS-analysis. Due to the unavailability of a xanthosine isotope standard, the response on column for xanthosine was normalized to the response on column of the dT isotope standard, resulting in an arbitrary unit (AU). Asterisks indicate significant differences ($p < 0.05$) to Col-0. Error bars are the standard deviation for $n = 5$ biological replicates. Adjusted p-values for multiple comparisons can be found in Table S2.

In the manuscript we speculated that xanthosine accumulation is caused by a reduced activity of NSH1 to avoid further depletion of the thymidylate pool in the *dhfr-ts2* mutant, which is partially compromised in thymidylate synthesis. A speculative model explaining also the impact of other genotypes on xanthosine is provided here: Lack of DHFR-TS2 probably results in accumulation of its substrate dUMP. In the cytosol dUMP can likely be dephosphorylated to dU which may act as a competitive inhibitor of NSH1. Consistently, xanthosine accumulation is stronger in the *dhfr-ts2 tk1a* mutant, because here cytosolic dU salvage is also blocked, exerting an even stronger effect on NSH1. In the *dut1org* mutant the concentration of dUMP is probably lower because it is not generated from dUTP in the organelles. In consequence also the dU content will be lower. Therefore, *dut1org* does not show xanthosine accumulation and even suppresses xanthosine accumulation in the *dhfr-ts2* background.

Fig. S7: Concentrations of ribonucleotide triphosphates in the wildtype and mutant seeds lacking enzymes putatively involved in thymidylate homeostasis in late germination. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 48 h to growth conditions and harvested for LC-MS-analysis. Concentrations of ATP, CTP, GTP and UTP in seeds are shown. Asterisks indicate significant differences ($p < 0.05$) to Col-0. Five to six biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted p-values for multiple comparisons can be found in Table S2.
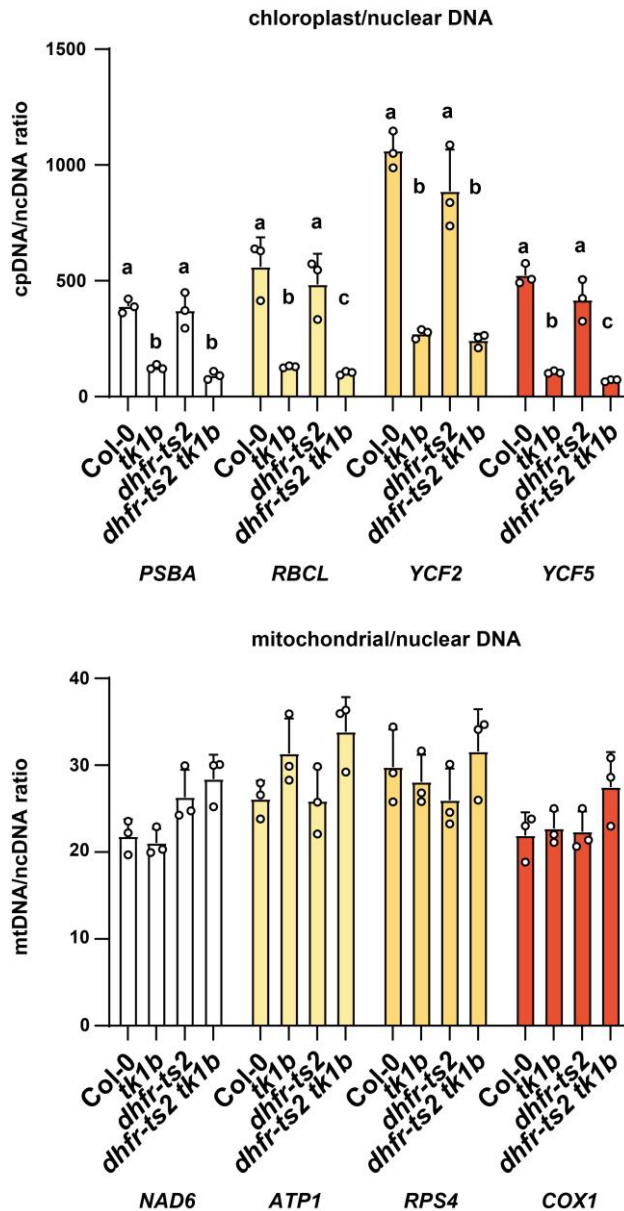
Figure S8. Relative abundances of cpDNA and mtDNA during late germination determined by the amplification of four genes, respectively. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 48 h to growth conditions and harvested. Q-PCR was performed with primers amplifying four genes from cpDNA (*PSBA*, *RBCL*, *YCF2* AND *YCF5*), mtDNA (*NAD6*, *ATP1*, *RPS4* and *COX1*) located at distant positions on the plastid and mitochondrial genomes and one gene from ncDNA (*UBC*21). The cpDNA/ncDNA and mtDNA/ncDNA ratios were calculated from Ct values for the respective genes. Different letters represent significant differences ($p <$ 0.05) between genotypes. Three biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted *p*-values for multiple comparisons can be found in Table S2.
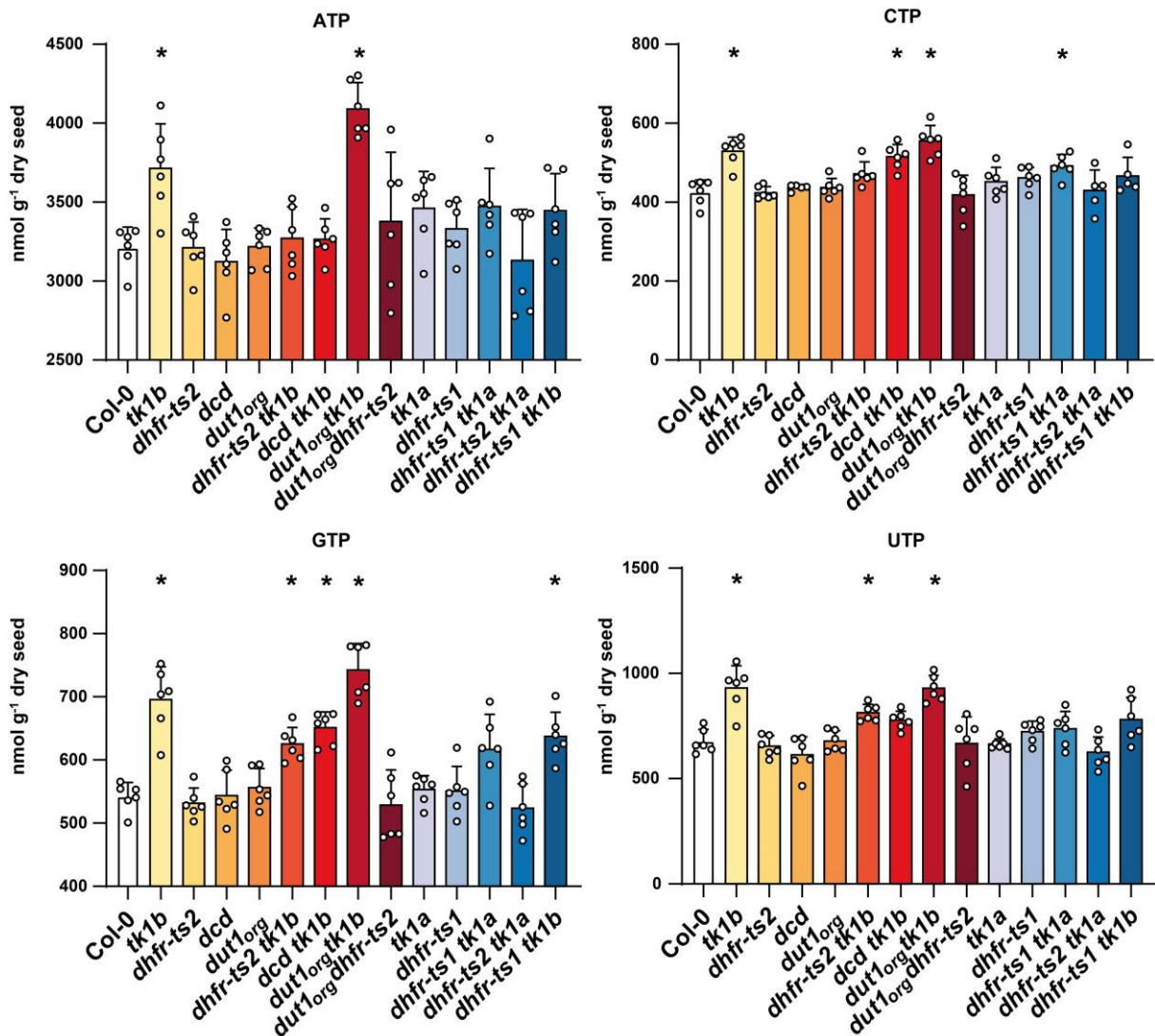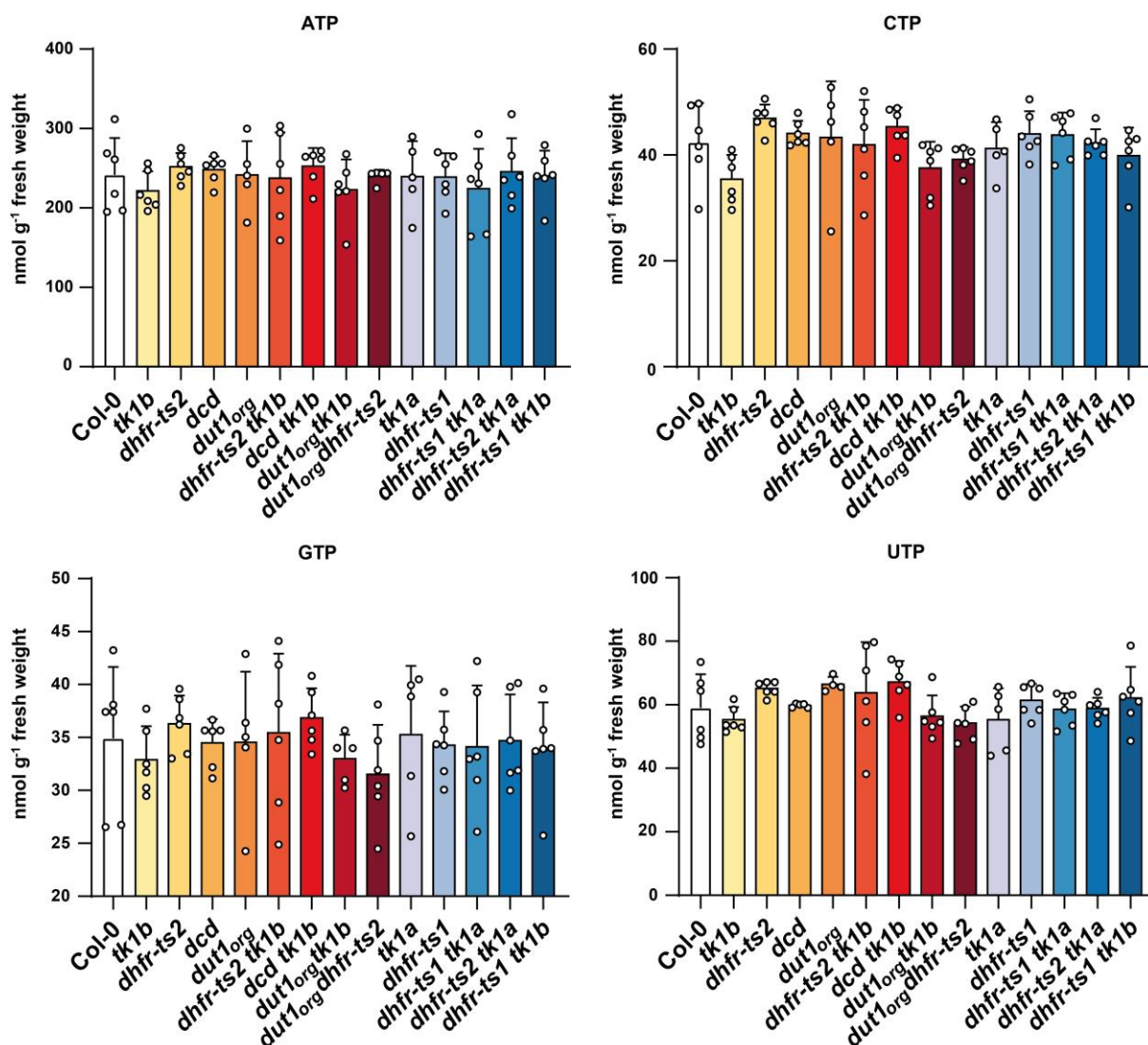
Figure S9. Concentrations of ribonucleotide triphosphates in the wildtype and mutants lacking enzymes putatively involved in thymidylate homeostasis in established seedlings. Seeds were imbibed in water for 48 h at 4°C in the dark, transferred for 144 h to growth conditions and harvested for LC-MS-analysis. Concentrations of ATP, CTP, GTP and UTP in seeds are shown. No significant differences ($p < 0.05$) between wildtype and mutants were observed. Five to six biological replicates (pooled seedlings) were analyzed. Error bars are SD. Adjusted p-values for multiple comparisons can be found in Table S2

2.6 Initiation of cytosolic plant purine nucleotide catabolism involves a monospecific xanthosine monophosphate phosphatase

Katharina J. Heinemann*[1], Sun-Young Yang*[2], Henryk Straube[1], Nieves medina-Escobar[1], Marina Varbanova-Herde[1], Marco Herde[1], Sangkee Rhee and Claus-Peter Witte[1]

[1] Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany

[2] Seoul National University, Department of Agricultural Biotechnology, 151-921 Seoul, Republic of Korea

* Authors contributed equally

Initiation of cytosolic plant purine nucleotide catabolism involves a monospecific xanthosine monophosphate phosphatase

Katharina J. Heinemann[1,3], Sun-Young Yang[2,3], Henryk Straube [1], Nieves Medina-Escobar[1], Marina Varbanova-Herde[1], Marco Herde[1], Sangkee Rhee [2✉] & Claus-Peter Witte [1✉]

In plants, guanosine monophosphate (GMP) is synthesized from adenosine monophosphate via inosine monophosphate and xanthosine monophosphate (XMP) in the cytosol. It has been shown recently that the catabolic route for adenylate-derived nucleotides bifurcates at XMP from this biosynthetic route. Dephosphorylation of XMP and GMP by as yet unknown phosphatases can initiate cytosolic purine nucleotide catabolism. Here we show that *Arabidopsis thaliana* possesses a highly XMP-specific phosphatase (XMPP) which is conserved in vascular plants. We demonstrate that XMPP catalyzes the irreversible entry reaction of adenylate-derived nucleotides into purine nucleotide catabolism in vivo, whereas the guanylates enter catabolism via an unidentified GMP phosphatase and guanosine deaminase which are important to maintain purine nucleotide homeostasis. We also present a crystal structure and mutational analysis of XMPP providing a rationale for its exceptionally high substrate specificity, which is likely required for the efficient catalysis of the very small XMP pool in vivo.

[1] Leibniz Universität Hannover, Department of Molecular Nutrition and Biochemistry of Plants, Herrenhäuser Strasse 2, 30419 Hannover, Germany. [2] Seoul National University, Department of Agricultural Biotechnology, 151-921 Seoul, Republic of Korea. [3]These authors contributed equally: Katharina J. Heinemann, Sun-Young Yang. ✉email: srheesnu@snu.ac.kr; cpwitte@pflern.uni-hannover.de

Purine metabolism in plants and other eukaryotes differs in several aspects. For example, in plants, purine nucleotide biosynthesis generating AMP is localized in chloroplasts, whereas in mammals or in yeasts it is located in the cytosol. However, plant GMP biosynthesis, starting from AMP, occurs in the cytosol (Fig. 1a and Supplementary Fig. 1)[1,2]. In contrast to mammals, plants are able to degrade purine nucleotides completely, disintegrating the purine ring via intermediates like urate

and allantoate into glyoxylate, carbon dioxide and ammonium[3,4]. Recently, it was shown in *Arabidopsis thaliana* that there are two independent entry reactions into purine nucleotide catabolism[5]. One is leading from xanthosine monophosphate (XMP) to xanthosine presumably catalyzed by an XMP phosphatase (XMPP). This reaction may serve to degrade the adenylate-derived purine nucleotides. The other also generates xanthosine but via the dephosphorylation of GMP to guanosine by a so far unknown
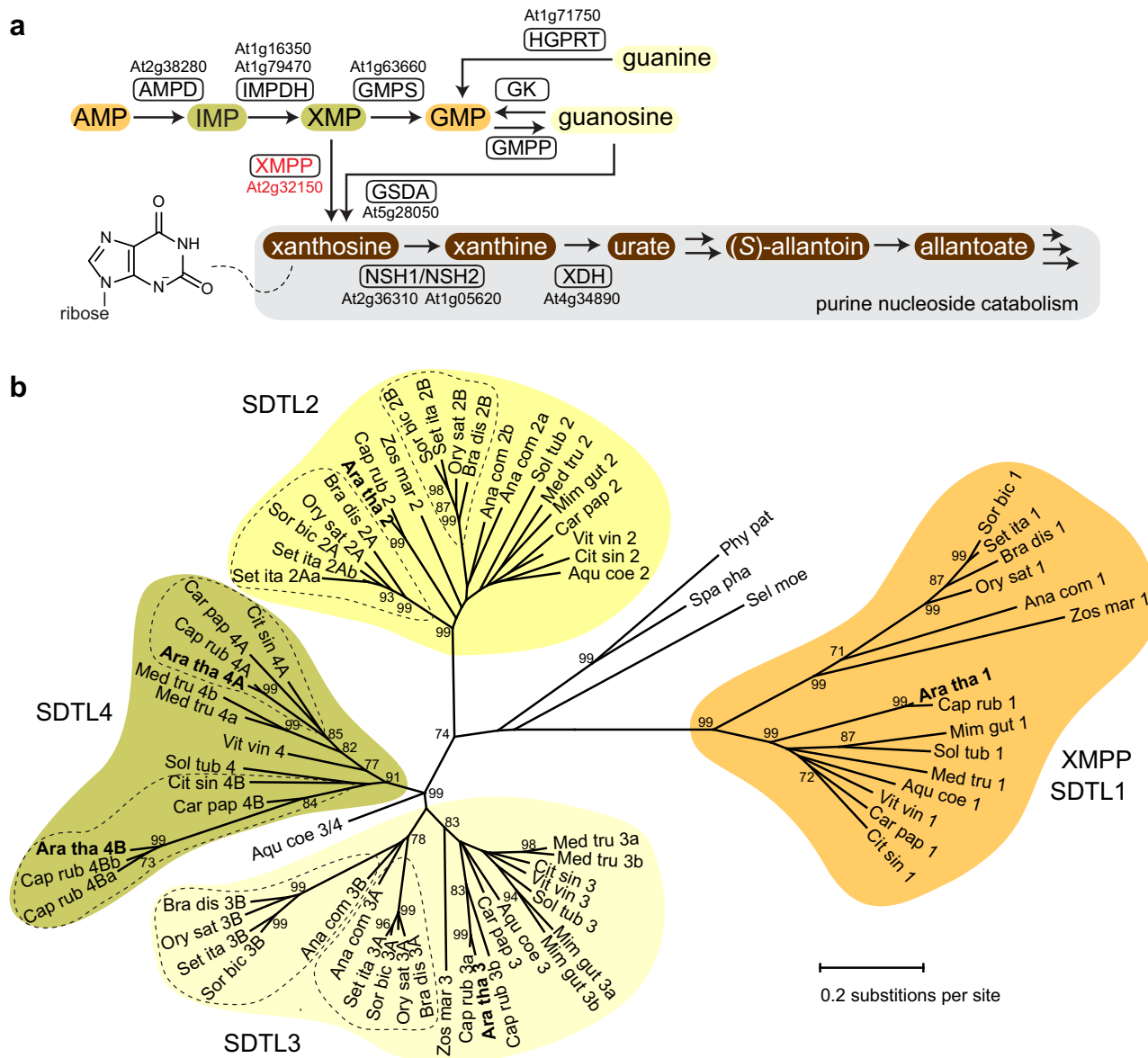
Fig. 1 Scheme of GMP biosynthesis and purine nucleotide catabolism and phylogenetic analysis of SDTL proteins. a GMP biosynthesis and purine nucleotide catabolism in the cytosol are initiated from AMP, which is deaminated to inosine monophosphate (IMP) by AMP deaminase (AMPD). IMP is oxidized by IMP dehydrogenase (IMPDH) to xanthosine monophosphate (XMP). For biosynthesis of GMP, XMP is aminated by GMP synthetase (GMPS). GMP and XMP can be dephosphorylated by XMP phosphatase (XMPP) and GMP phosphatase (GMPP) yielding guanosine and xanthosine, respectively. Both enzymes have not yet been described. Guanosine is deaminated to xanthosine by guanosine deaminase (GSDA). In plants, xanthosine can only be degraded and not be salvaged into nucleotides. By contrast, guanosine can be salvaged to GMP by a so far unknown guanosine kinase (GK) and guanine, which is not a purine nucleotide catabolism intermediate in Arabidopsis, can be salvaged by hypoxanthine guanine phosphoribosyltransferase (HGPRT). Xanthosine is hydrolyzed to ribose and xanthine by the nucleoside hydrolase heterocomplex (NSH1/NSH2) which is not functional in the absence of NSH1. Xanthine is oxidized by xanthine dehydrogenase (XDH) to urate. Via the intermediates (*S*)-allantoin and allantoate, urate can be fully degraded to glyoxylate, carbon dioxide, and ammonium. The gene locus identifiers are given where known. A scheme including chemical formulas is shown in Supplemental Fig. 1. **b** Maximum Likelihood tree constructed with SDTL sequences from 15 phylogenetically distant vascular plant and three moss species. The tree with the highest log likelihood is shown. In total, 1000 bootstraps were performed and only bootstrap values over 70% are shown. Species names and accession numbers are given in Supplementary Table 1. Supplementary Data 1 contains the multiple protein alignment from which the tree was calculated.

GMP phosphatase (GMPP) and the subsequent deamination of guanosine by guanosine deaminase (GSDA), an enzyme found only in plants (Fig. 1a and Supplementary Fig. 1)[6]. Because xanthosine cannot be converted into other nucleosides and cannot be re-phosphorylated (salvaged) to XMP in Arabidopsis[7,8], the XMPP and GSDA reactions both lead irreversibly into purine catabolism.

Although several enzymes with nucleotide monophosphate (NMP) phosphatase activity have been described in eukaryotes, there are only few examples where their physiological role could be clearly demonstrated[9,10]. Why is the functional assignment of NMP phosphatases so difficult? Possible reasons are (i) that NMPs are central metabolites and their pool sizes are affected by a multitude of reactions so that single mutations hardly change the pools at all; (ii) eukaryotes have many enzymes with NMP phosphatase activity often with a broad substrate spectrum complicating the assignment of specific physiological roles; (iii) NMP phosphatases may act partially redundant in vivo.

In this work, we were able to identify a plant-specific NMP phosphatase, the XMP phosphatase. We have characterized the enzyme biochemically and report its crystal structure with bound substrate elucidating its specificity determinants. Most importantly, we demonstrate that XMPP is involved in purine nucleotide catabolism in vivo.

Results and discussion

Identification and molecular characterization of XMPP. In *Saccharomyces cerevisiae* two homologous enzymes, Sdt1p (Suppressor of disruption of *TFIIS*) and Phm8p (Phosphate metabolism protein 8), were shown to hydrolyze NMPs in vitro. Phm8p but not Sdt1p is required in vivo for purine and pyrimidine mononucleotide degradation[10].

Because Phm8p also shows activity with XMP and GMP, we searched for genes in the Arabidopsis genome encoding homologous proteins to Phm8p/Sdt1p with the aim to identify XMP- and GMP-specific plant phosphatases (Fig. 1a and Supplementary Fig. 1). Five proteins homologous to Phm8p/Sdt1p were found which we called SDTL for Sdt1p-like as they were more similar to Sdt1p than to Phm8p. Because enzymes of primary metabolism are usually highly conserved in plants, we assumed that only SDTLs that are present in all vascular plants are good XMP/GMP phosphatase candidates. To identify such SDTLs, a maximum likelihood tree was constructed using the full complement of SDTL sequences from 15 phylogenetically distant vascular plants and 3 moss species (Fig. 1b and Supplementary Table 1). Four major clades of SDTL proteins were found. The clades SDTL1, SDTL2, and SDTL3 each contain at least one representative sequence from every analyzed vascular plant species, whereas the SDTL4 clade lacks sequences from several plant species, in particular from the monocots. Each of the mosses only has a single SDTL not associated with any of the clades. The clades likely represent groups of orthologous proteins with identical or at least similar functions. The SDTL1 clade is more distant from the others and contains only a single protein from each of the analyzed vascular plants, whereas in other clades also sub-clades are found indicating some functional diversification.

We transiently expressed cDNAs coding for C-terminal Strep-tagged variants of the SDTL proteins from the clades 1, 2, and 3 of Arabidopsis in leaves of *Nicotiana benthamiana*, affinity purified the proteins, and screened their activity with XMP and GMP as substrates. GMP phosphatase activity was not observed for any of the enzymes but SDTL from clade 1 showed activity with XMP. We extended the substrate survey for this enzyme to other mononucleotides and found that it is highly specific for
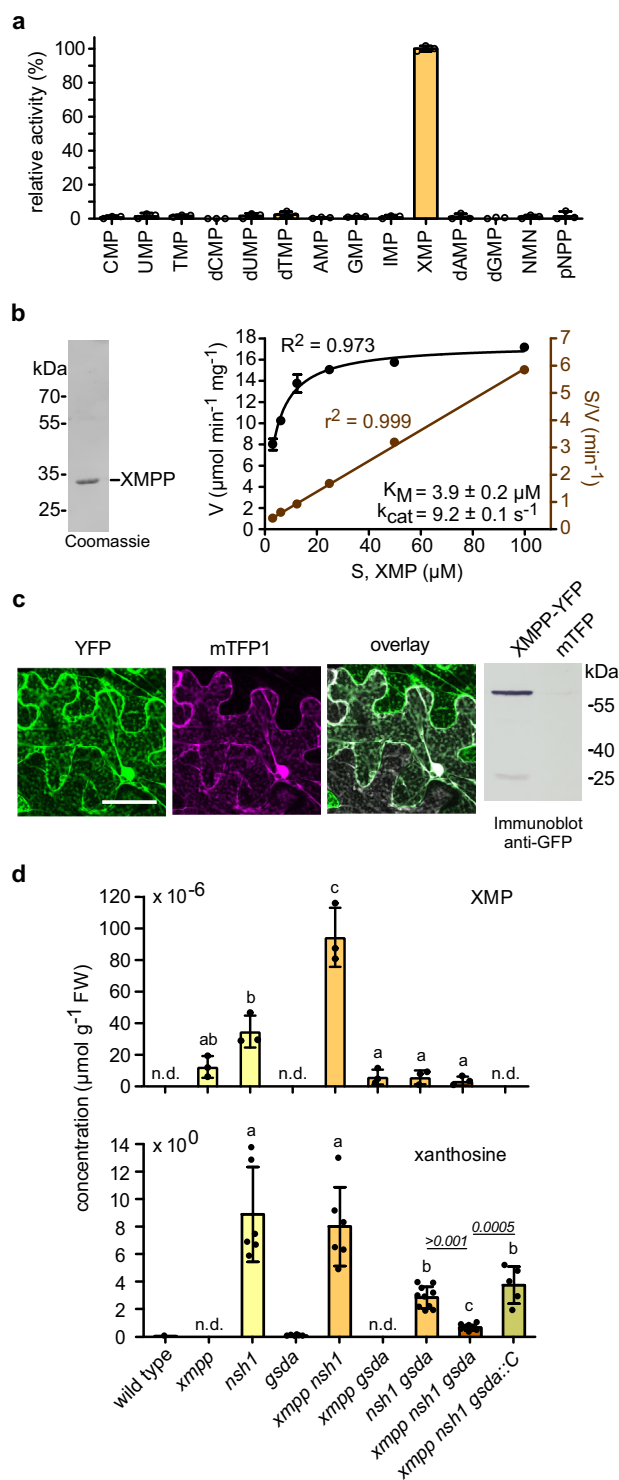
XMP (Fig. 2a). Such high specificity is unusual for NMP phosphatases. The enzyme has a K_M value of $3.9 \pm 0.2 \, \mu M$ and a k_{cat} value of $9.2 \pm 0.1 \, s^{-1}$ for XMP (Fig. 2b). An N-terminal Strep-tagged variant has similar kinetic constants (Supplementary Fig. 2) indicating that the tags and the tag position do not interfere with the activity. Based on these data, we named the enzyme XMP phosphatase (XMPP). The subcellular localization of XMPP was investigated by confocal laser scanning microscopy in the leaves of *N. benthamiana* transiently expressing the enzyme as C-terminal yellow fluorescent protein (YFP)-tagged variant. XMPP is present in the cytosol and the nucleus (Fig. 2c) as are

Fig. 2 Characterization of the XMPP protein and of *XMPP* genetic variants in the context of other mutants of purine catabolism genes in seeds. a Relative enzymatic activity of XMPP-HAStrep with 100 μM (deoxy)mononucleotides including nicotinamide mononucleotide (NMN) and the general phosphatase substrate *para*-nitrophenyl phosphate (pNPP). Error bars are SD, $n = 3$. An XMP conversion rate of $12.1 \pm 0.2\ \mu mol\ mg^{-1}\ min^{-1}$ was set to 100%. **b** XMPP-HAStrep affinity purified from leaf extracts of *Nicotiana benthamiana* after transient expression. Left panel, Coomassie-stained SDS-gel with the purified enzyme. Right panel, determination of the kinetic constants at 22 °C with the data fitted according to Michaelis Menten (left axis) or Hanes (right axis). Error bars are SD, $n = 3$. **c** Confocal fluorescence microscopy images of a lower epidermis cell of *N. benthamiana* co-expressing XMPP-YFP and the cytosolic cyan fluorescent protein mTFP1. From left to right, YFP channel, mTFP1 channel, overlay of both channels. Scale bar = 50 μm. Immunoblot of *N. benthamiana* leaf extracts expressing either XMPP-YFP or free mTFP1 probed with anti-GFP antibody demonstrating the high stability of the fusion protein. The immunoblot was repeated twice with similar results. **d** Quantification of XMP and xanthosine in seed extracts of the indicated genotypes. The *xmpp-1* allele (Supplementary Fig. 3a, b) was used in all instances. The complementation line (*xmpp nsh1 gsda::C*) contains a transgene encoding an N-terminal TwinStrep-tagged XMPP expressed from a native promoter (Supplementary Fig. 3c, d). Error bars are SD, $n = 3$ for XMP analysis and for xanthosine analysis $n = 6$ but for *xmpp nsh1 gsda::C* $n = 5$ and for *nsh1 gsda* and *xmpp nsh1 gsda* $n = 10$. A repeat (*n*) is a seed sample from an independent mother plant grown in parallel with all other plants of the experiment. Statistical analysis with two-sided Tukey's pairwise comparisons using the sandwich variance estimator. Different letters indicate *p* values < 0.05. Some *p* values are indicated above the columns in italic numbers; all *p* values can be found in the Source Data file.

other enzymes upstream and downstream of XMPP, i.e. IMP dehydrogenase (IMPDH), GMP synthetase (GMPS)[2], guanosine deaminase (GSDA)[6], and the nucleoside hydrolase heterocomplex (NSH1/NSH2)[11–13] (Fig. 1a).

The function of XMPP in purine nucleotide catabolism in vivo. To investigate whether XMPP is involved in degrading XMP in vivo, we used a series of single (*xmpp, nsh1, gsda*), double (*xmpp nsh1, xmpp gsda, nsh1 gsda*), and triple (*xmpp nsh1 gsda*) mutants as well as two complementation lines containing a *TwinStrep-XMPP* transgene driven from the native promoter (Supplementary Fig. 3). The pools of XMP, xanthosine, guanosine, and guanine in these genetic variants were quantified in fresh seeds, because it is known that purine catabolism is active during seed development. The XMP concentration was close to the detection limit in most genotypes but was slightly elevated in the *xmpp* line (Fig. 2d). In *nsh1* seeds, more XMP was detected and in the *xmpp nsh1* seeds the largest XMP pool was observed, suggesting that XMPP is involved in XMP catalysis in vivo. The absence of NSH1, which plays a central role for xanthosine hydrolysis[5], has a positive effect on the XMP concentration possibly because the strong accumulation of xanthosine in *nsh1* seeds[5,11,13] (Fig. 2d) leads to a partial inhibition of the two XMP converting enzymes: XMPP and GMPS (Fig. 1a). For XMPP, we confirmed an inhibition by xanthosine (Supplementary Fig. 4). XMP accumulation in *xmpp* seeds is comparatively low probably because XMP is easily aminated by GMPS to GMP, which can be degraded by an unknown GMP phosphatase (GMPP) and GSDA via guanosine to xanthosine (Fig. 1a). This hypothesis is supported by hyperaccumulation of guanosine and guanine in seeds with *xmpp gsda* background compared to *gsda* seeds (Supplementary Fig. 5). GMP can also be phosphorylated leading in tendency to GTP hyperaccumulation in this double mutant

background, which is prevented and even over-compensated by the introduction of an XMPP transgene (Supplementary Fig. 6). It has been noted before that guanosine accumulation in *gsda* plants results in higher concentrations of guanylates and in consequence adenylates[14]. The GMPP and GSDA reactions seem to guard guanylate homeostasis.

Xanthosine is the first common product of direct XMP dephosphorylation by XMPP and of GMP degradation by GMPP and GSDA (Fig. 1a). The xanthosine pool size in *nsh1* background may therefore serve as a proxy for assessing if both routes are operative. Whereas in *nsh1 gsda* seeds the xanthosine concentration is reduced strongly compared to the *nsh1* background, this is not the case in *xmpp nsh1* seeds (Fig. 2d), suggesting at first sight that the XMPP route plays no role. However, one needs to bear in mind that XMP can easily be converted to GMP by GMPS, especially if XMPP is not functional, thus only in *xmpp nsh1 gsda* seeds one can assess whether XMPP contributes to the xanthosine pool in vivo. This is indeed the case because in the triple mutant the xanthosine concentration is strongly reduced and this effect is reversed if the triple mutant expresses an *XMPP* transgene. These data demonstrate that XMPP catalyzes the dephosphorylation of XMP in seeds representing an entry point of the adenylate-derived nucleotides into purine catabolism in vivo. The XMPP reaction is likely the main entry point for the adenylates into catabolism, whereas the GMPP reaction serves for the degradation of the guanylates. Note, that xanthosine cannot be salvaged in Arabidopsis whereas guanosine can[8]—thus the XMPP reaction leads irreversibly into purine nucleotide degradation.

Purine catabolism is enhanced by prolonged darkness[5,15]. Therefore, we assessed the role of XMPP in seedlings exposed to a night prolonged by 48 h in comparison to seedlings in a 16 h day/ 8 h night regime. Dark treatment led to a marked increase of the steady-state pool sizes of urate and allantoate (Fig. 3a), which are intermediates of purine nucleoside catabolism[4,16] (Fig. 1a and Supplementary Fig. 1). In *gsda* seedlings under long-day conditions, the urate pool was already smaller than in the wild type or in *xmpp* seedlings, while in *xmpp gsda* plants, the urate concentration was even more reduced and allantoate could not be detected anymore. In the prolonged night, urate and allantoate concentrations rose but not in the *xmpp gsda* seedlings, demonstrating that the degradation of XMP by XMPP and of GMP via GSDA feeds these pools. The dark-exposure experiment was repeated including a line expressing an *XMPP* transgene in *xmpp gsda* background (Supplementary Fig. 3c), which complemented metabolic changes observed in the *xmpp gsda* line compared with the *gsda* line (Supplementary Fig. 7). Together these results show that XMP and GMP dephosphorylation can initiate cytosolic purine nucleotide catabolism, but it is not possible to quantify the relative contributions of XMP versus GMP degradation from the data. Reasons are that the XMPP reaction can be bypassed in *xmpp* background via GMP and guanosine (Fig. 1a), and that in *gsda* seedlings guanosine accumulates so strongly that it partially inhibits the NSH1/ NSH2 complex, which leads to xanthosine buildup in the dark (Fig. 3a and Supplementary Fig. 7), and consequently to partial inhibition of XMPP (Supplementary Fig. 4). The dark-induced xanthosine accumulation in *gsda* background is not observed in *xmpp gsda* seedlings, again showing that the XMPP reaction directly contributes to the xanthosine pool in vivo. Because XMP cannot be channeled into degradation in *xmpp gsda* plants, these double mutants accumulate more guanosine than *gsda* seedlings in the dark (Fig. 3a). In an independent experiment we observed that the *xmpp gsda* plants also over-accumulate GTP and in consequence ATP in tendency, which is prevented by the introduction of an *XMPP* transgene (Fig. 3b and Supplementary

a

b

Fig. 3 Characterization of metabolic alterations in *XMPP* genetic variants in the context of other mutants of purine catabolism genes in seedlings under long-day or prolonged-night conditions. Two-sided Tukey's pairwise comparisons using the sandwich variance estimator were employed for statistical analyses. Different letters indicate *p* values < 0.05. Samples from long day and prolonged darkness were independently analyzed indicated by small and capital letters. Some *p* values are indicated above the columns in italic numbers, all *p* values can be found in the Source Data file. **a** Urate, allantoate, xanthosine, and guanosine pool sizes in the indicated genotypes in seedlings. The seedlings were grown in a constant 16 h day/8 h night regime up to day 5 after germination. At the end of the night of day 5, they were either grown the same way for another 48 h (orange bars) or exposed to darkness for 48 h (dark bars). All genotypes were grown on the same plate for one biological replicate (*n*). Error bars are SD, *n* = 6. **b** GTP and ATP pool sizes of seedlings from an independent experiment grown as described in **a**. Error bars are SD, *n* = 5 but for *gsda* in the light and the wild type in the dark *n* = 4.

The mutation of *XMPP* does not create an obvious macroscopic phenotype at any developmental stage (Supplementary Fig. 9) and also not during dark stress. This is not surprising because (1) many purine catabolism mutants of Arabidopsis do not show phenotypes under standard growth conditions except those that accumulate toxic intermediates like guanosine[15] or uric acid[16] and because (2) it is possible to bypass the XMPP reaction via GMP and guanosine as shown above. However, this bypass is energetically costly, because the GMPS reaction requires the conversion of ATP to AMP and uses glutamine for the amination of XMP. Subsequently, this amino group is released as ammonia by GSDA and must be re-assimilated into glutamine which is energized by the conversion of ATP to ADP. Thus in total the bypass requires the hydrolysis of three phosphoanhydride bonds that are spared when the XMPP reaction is used.

The molecular basis of substrate recognition by XMPP. XMPP is highly substrate specific (Fig. 2a). To elucidate the molecular basis for this specificity, we determined the crystal structure of XMPP with and without XMP bound (Fig. 4, Supplementary Fig. 10a and Supplementary Table 2). Crystals could only be obtained for a shortened XMPP variant lacking 13 amino acids at the C-terminus (Met1 to Ser250). Consistent with the monomeric state of the enzyme in solution (Supplementary Fig. 10b), crystalized XMPP is monomeric with a two-domain architecture (Fig. 4a). In the α/β-folded central domain with five repeating β-α units, the central parallel six-stranded β-sheet is flanked by five α-helices on both sides. A 20-residue-long loop (Pro142 to Ser161) protrudes from the β3–α8 unit, and a cap domain (Leu20 to Ser89) with a four α-helices bundle-like architecture is inserted in the β1–α5 unit. The loop probably stabilizes the orientation of the cap domain (Supplementary Fig. 11). XMPP is a member of the Haloacid Dehalogenase (HAD) protein superfamily with unique characteristics since HAD enzymes typically have either a protruding loop or a cap structure[17], but not both, like XMPP.

At the junction of the cap and the central domains, a Mg^{2+} ion coordinates six ligands within 2.1 Å in a square bi-pyramidal geometry (Fig. 4b) involving Asp12 and Asp14. These aspartates are conserved (Supplementary Fig. 10a) and important for the catalytic activity of HAD enzymes. In the complex with XMP, the 5′-monophosphate moiety of XMP provides an equatorial ligand across Asp184 in the Mg^{2+}-coordination shell, and the ribose and xanthine moieties face toward the cap domain (Fig. 4c, d and Supplementary Fig. 12). Unlike the ribose moiety, xanthine embedded on the concave side of the cap domain interacts extensively with the enzyme via (i) hydrophobic interactions with
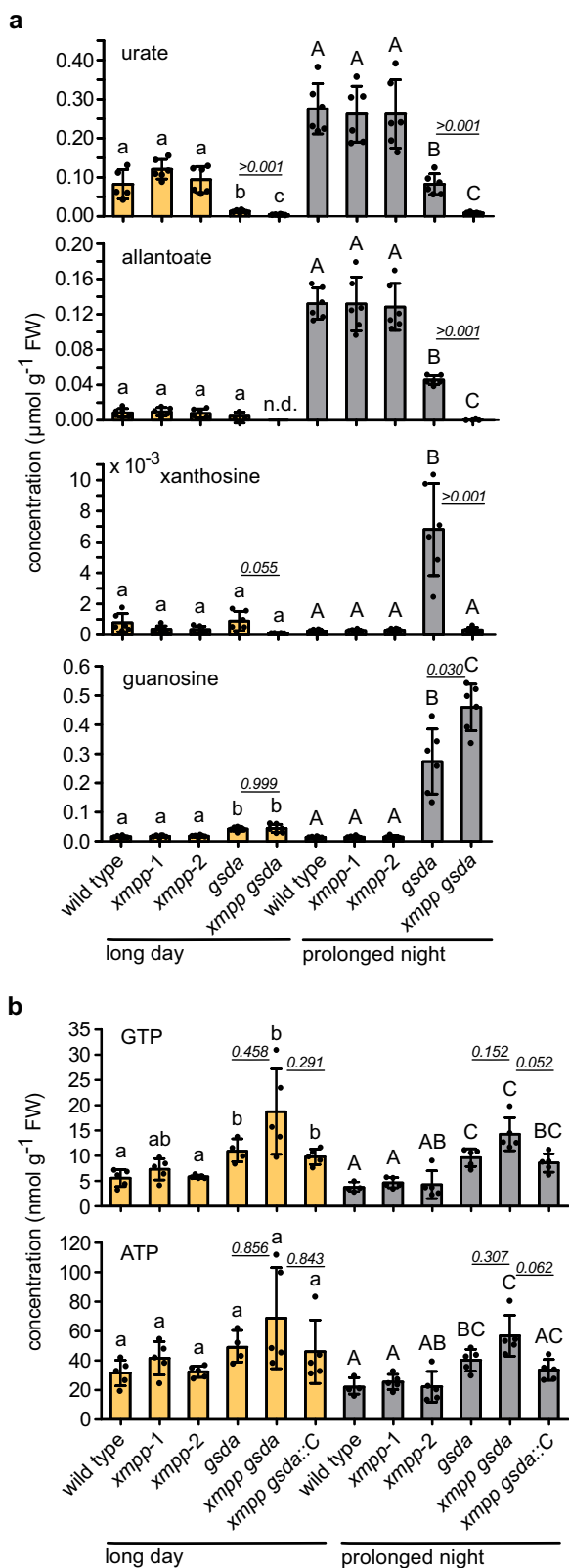
Fig. 8). These tendencies were repeatedly observed in the long day and in the prolonged night. It is clear that GSDA, probably in concert with the unknown GMPP, is necessary to keep guanylate and adenylate homeostasis, and that XMPP prevents adenylate-derived nucleotides destined for degradation to spill into the guanylate pool.
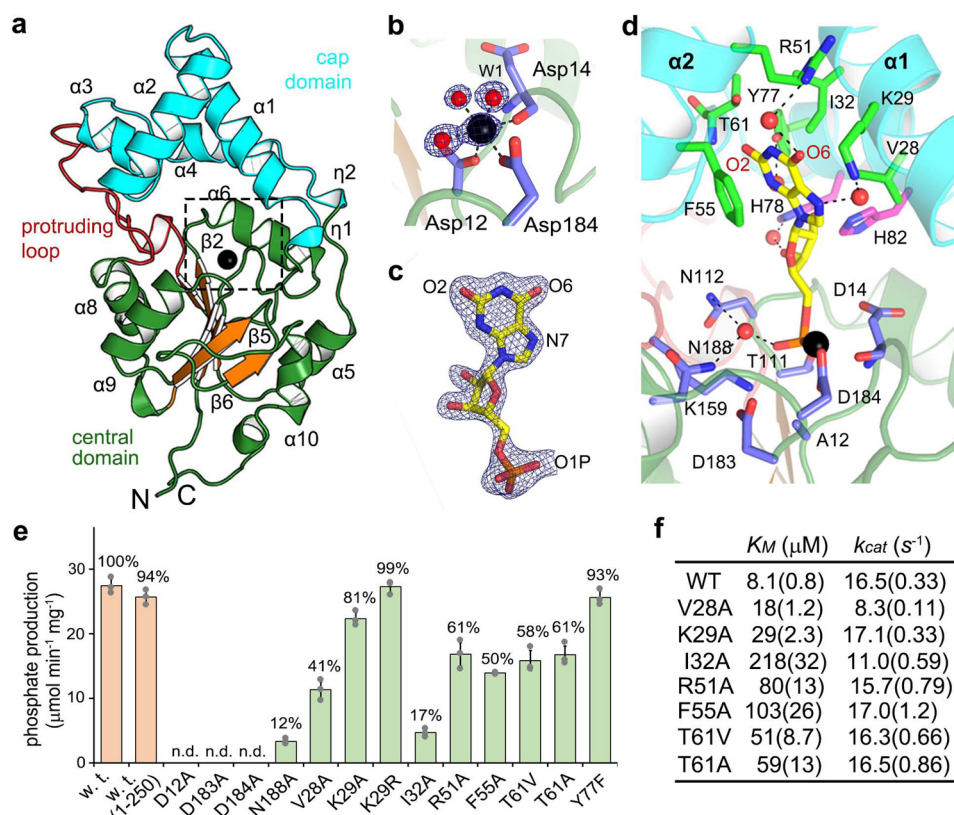
Fig. 4 Structural and mutational analysis of XMPP. a Structure of the unliganded XMPP. The repeating β–α units (organge, green) in the central domain include β1–α5, β2–α6, β3–α8, β4–α9, and β6–α10. Metal-binding site, dashed box. Mg^{2+} ion, black sphere; cap domain, cyan; protruding loop, red. **b** A zoomed-in view of the proposed Mg^{2+}-binding site. Dashed lines indicate the six coordination bonds. Asp12 and water W1, axial ligands. Asp184, the main chain carbonyl oxygen of Asp14 and two water, equatorial ligands. Water (red) and metal ion (black) are overlaid with a $2Fo-Fc$ electron density map contoured at 1.0σ and 3.0σ, respectively. **c** XMP as oriented in the structure of the complex overlaid with an $Fo-Fc$ electron density map contoured at 3.5σ. **d** An enlarged view of the XMP-binding site oriented as in **a**. The complex with XMP was obtained with an inactive mutant, in which the putative catalytic Asp12 was replaced with alanine. Blue, residues interacting with phosphate and Mg^{2+}; magenta, ribose moiety-interacting residues; green, xanthine-binding residues. All residues are within 4.0 Å of XMP. Water-mediated hydrogen bonds, dashed lines. The phosphate moiety interacts with residues that are highly conserved in the HAD superfamily, including Asp14, Thr111, Asn112, Lys159, and Asn188. **e** Specific activities of XMPP variants with 125 μM XMP as substrate. Full-length enzymes with distinct point mutations expressed in *Escherichia coli* and purified via a C-terminal His tag were used. Relative activities compared to the wild-type enzyme (27.4 nmol min^{-1} μg^{-1} set to 100%) are given. Error bars are SD, $n = 3$ repeated assays using the same enzyme preparation. n.d. for not detected. Note that the C-terminally truncated XMPP (residues 1–250) used for structure determination exhibited an almost equal specific activity to that of the full-length wild-type enzyme suggesting that the C-terminal region beyond Ser250 is not crucial for in vitro phosphatase activity. **f** Kinetic constants of various mutants of xanthine-interacting residues. SE is indicated in a parenthesis.

Val28 and Ile32 and (ii) by base stacking with Phe55, and (iii) via direct or water-mediated indirect hydrogen bonds of O_2 with Thr61, O6 indirectly with Arg51, and N3 and N7 indirectly with Tyr77 and Lys29, respectively (Fig. 4d).

To validate the structure-based functional assignments, enzymatic activities and kinetic constants of XMPP variants were determined. The specific activities of the XMPP variants closely reflect the assigned functional roles of the respective residues (Fig. 4d, e). Mg^{2+}-coordination by Asp12, Asp183, and Asp184 is linked to catalysis and is of critical importance since the respective mutants are inactive. The interaction of Asn188 with the phosphate and of various residues with the xanthine moiety are required for full activity (Fig. 4e). Further kinetic analysis revealed that mutations of xanthine-interacting residues do not affect the k_{cat} but the K_M values for XMP (Fig. 4f and Supplementary Fig. 13). For variants altered in hydrophobic residues, the K_M increases 2.2- to 26.9-fold. For those changed in hydrogen bond-forming residues, the K_M increase ranges from 3.6- to 9.9-fold. The XMP specificity is thus collectively mediated by the hydrophobic residues Val28, Ile31, and Phe55 that shape

the xanthine-binding pocket in the cap domain and by direct and indirect hydrogen bonds to Thr61, Lys29, and Arg51. These amino acids are largely conserved in putative XMPPs from other plants (Supplementary Fig. 10a) but there are some slight variations. Thr61, for example, which forms a direct hydrogen bond to the characteristic O_2 of xanthine can be replaced by serine in enzymes from other plants.

We were unable to detect XMP in seedling extracts but estimate that its concentration must be at least three orders of magnitude smaller than the GMP concentration. The high specificity of XMPP together with a K_M in the low micromolar range and a high k_{cat} are therefore of physiological importance, because a less specific enzyme would constantly dephosphorylate other more abundant NMPs. Additionally, a precise adjustment of the XMPP amount is apparently required because even a moderate overexpression can strongly disturb GTP homeostasis (Supplementary Fig. 6). It will be interesting to investigate how this enzyme at the crossroad between nucleotide biosynthesis and degradation is regulated and coordinated with the activity of GMP synthetase to adjust the flux of XMP towards the guanylates or into catabolism.

Methods

Plant material and cultivation. *Arabidopsis thaliana* ecotype Columbia-0 was used throughout. The T-DNA mutants *xmpp-1* (SALK067037) and *xmpp-2* (SALK131244, At2g32150), *nsh1* (SALK083120, At2g36310)[11], and *gsda* (GK432D08, At5g28050)[6] were obtained from the SALK[18] and GABI-Kat[19] collections. Seedlings of *xmpp-1* and *xmpp-2* were tested for the presence of intact transcript using the primers N188 and N233 (Supplementary Table 3). The presence of *Actin2* (At3g18780) transcript was tested with the primers 1033 and 1034. The mutants *xmpp-1*, *nsh1*, and *gsda* were used to generate double and triple mutants by crossing. Complementation lines of *xmpp gsda* and *xmpp nsh1 gsda* were generated by transformation with construct H453 (Supplementary Fig. 3c,d).

Arabidopsis and *N. benthamiana* were cultivated under long-day conditions (16 h light of 85 µmol m^{-2} s^{-1}, 22 °C/8 h dark, 20 °C) at 60% relative humidity. For metabolite analysis of Arabidopsis seeds, six individual plants of each genotype were grown next to each other in a randomized setup. Xanthosine, guanosine and guanine were quantified directly after seed harvest. Nucleotide analysis was performed with six-month-old seeds. Seedlings were grown on half-strength Murashige and Skoog medium. For metabolite analysis, each plate contained all genotypes representing one biological replicate. To equalize germination, plates were stored in the dark at 4 °C for 2 days before transfer to the growth chamber. Five days after germination half of the seedlings were transferred into the dark while the other half was kept under long-day conditions. After 48 h, whole seedlings were harvested. For the characterization of the complementation lines, seedlings were grown on separate plates for 10 days after germination and placed into the dark for 2 days.

Cloning. The coding sequence of *XMPP* was amplified from cDNA obtained from flower RNA introducing *Nco*I and *Xma*I sites with the primers N188 and N189 or *Cla*I and *Xma*I sites with the primers N359 and N360. Via *Cla*I and *Xma*I, the coding sequence was introduced into pXCS-HAStrep[20] (vector V13, construct X130) to generate a C-terminal hemeaglutinin (HA)- and Strep-tagged protein by transient expression in *N. benthamiana*. Similarly, it was introduced into pXCS-YFP[21] (V36, X131) for producing a C-terminal yellow fluorescent protein (YFP)-tagged protein. The coding sequence of YFP in V36 was exchanged for a mTFP1 coding sequence amplified from a wave_1T vector[22] with the primers P243 and P244 introducing *Xma*I and *Sac*I sites. This generated a vector for the expression of a cytosolic cyan fluorescent protein (V108). For production of an N-terminal Strep-tagged protein, the coding sequence was cloned into pXNS2cpmv-strep[5] (V90, construct X144) via *Nco*I and *Xma*I. To obtain the complementation lines *xmpp gsda::C* and *xmpp nsh1 gsda::C*, we cloned a genomic fragment of XMPP including 999 bp upstream of the translation start codon with an introduced coding sequence for a TwinStrep tag for N-terminal tagging (Supplementary Fig. 3d). For this, pXNS2pat-Strep-sl (V28) was generated by cloning the annealed primers 1709 and 1710 into the binary vector pAMPAT-MCS (accession number AY436765) opened with *Xho*I and *Eco*RI. V28 was used to generate pXNS2patTwinStrep-sI (V163) by opening V28 with *Kas*I and *Nco*I and introducing a second Strep tag with a preceding linker sequence formed by the annealed primers P1097 and P1098. *XMPP* genomic DNA was amplified using the primers N188 and N189 and cloned into V163 via *Nco*I and *Xma*I. Into this construct the *XMPP* promoter, amplified with the primers P1094 and P1103, was cloned via *Asc*I and *Xho*I replacing a 35 S promoter cassette (construct H453). To generate untagged Arabidopsis XMPP in *Escherichia coli* strain BL21 (DE3) for antibody production, *XMPP* amplified with the primers N188 and N189 was cloned into pET30nco-CTH[23] (V48) using *Cla*I and *Sma*I for the *XMPP* cDNA and *Cla*I and *Eco*RV for the vector (H325).

Phylogentic analysis. Sequences were obtained from Phytozome V12.1 (Supplementary Table 1) using the protein sequence of XMPP from *Arabidopsis thaliana* as query. A multiple alignment (Supplementary Data 1) was generated with muscle hosted at the website of the European Bioinformatics Institute. The Maximum Likelihood method and the JTT matrix-based model[24] were employed to infer the evolutionary history within the MEGA X[25] software. Bootstrapping with 1000 repeats was performed and the results are shown at the respective branches. To start the heuristic search, the Neighbor-Join and BioNJ algorithms were applied to a matrix of pairwise distances estimated using the JTT model, and the tree topology with the best log likelihood was selected. The evolutionary rate differences among sites was modeled with a discrete Gamma (G) distribution of five categories (G = 1.1995). In all, 8.28% of sites were treated as invariable. In total, 76 amino acid sequences were analyzed excluding positions with more than 5% gaps. The final dataset comprised 234 positions. Evolutionary analyses were conducted in MEGA X[25].

Subcellular localization. *Nicotiana benthamiana* leaves were used for transient Agrobacterium (*Rhizobium radiobacter*)-mediated (co-)expression of the constructs X131 for production of XMPP-YFP and V108 for production of mTFP1 as a cyan fluorescent cytosolic marker. Abaxial leaf surfaces were analyzed using a Leica TSC SP8 microscope equipped with an HC PL APO CS2 ×40 1.10 water immersion objective (Leica, Germany). To prevent crosstalk, images were obtained by sequential scanning with an excitation of 448 nm for mTFP1 (emission 465–495 nm) and 514 nm for YFP (emission 524–539 nm). Images were processed using the Leica Application Suite Advanced Fluorescence software (Leica Microsystems).

Protein purification and determination of kinetic constants. Transient expression of Strep-tagged proteins from constructs X130 and X144 were performed by infiltrating *N. benthamiana* leaves with *Rhizobium radiobacter* at an optical density of 0.5 (at 600 nm). Three days after infiltration 0.75 g of leaves were ground in 1.5 mL buffer E containing 100 mM HEPES (pH 8.0), 100 mM NaCl, 5 mM EDTA (pH 8.0), 0.005% Triton X-100, 10 mM dithiothreitol, 1:625 diluted Biolock (IBA Life Sciences), 1:10 diluted protease inhibitor (complete protease inhibitor cocktail, Roche). After centrifugation, 40 µL of StrepTactin Macroprep (IBA Life Sciences) was added to the supernatant and incubated for 10 min on a rotation wheel at 4 °C. The matrix was washed three times with 1 mL buffer W1 (100 mM HEPES (pH 8.0), 100 mM NaCl, 0.5 mM EDTA (pH 8.0), 0.005% Triton X-100, 2 mM dithiothreitol) and centrifuged at 700 *g* for 30 s between washes. Then the matrix was washed twice with 1 mL buffer W2 (5 mM HEPES (pH 8.0), 100 mM NaCl, 0.5 mM EDTA (pH 8.0), 0.005% Triton X-100, 2 mM dithiothreitol). For batch elution 35 µL of buffer W2 containing 2.5 mM biotin was used. Elution was repeated and fractions were pooled[5,26]. Phosphatase activity was measured at 22 °C with the Enzchek Phosphate Assay Kit (Thermo Fisher Scientific) according to the manufacturer's instructions but in a total volume of 200 µL. In brief, 2-amino-6-mercapto-7-methyl-purine (MESG) and inorganic phosphate produced from the XMPP-dependent reaction serve as two substrates for purine nucleoside phosphorylase, resulting in a product exhibiting an absorbance at 360 nm. For kinetic measurements, XMPP amounts were adjusted to obtain linear initial rates at all XMP concentrations. Protein concentrations were determined with bovine serum albumin standards in a Coomassie-stained sodium dodecyl sulfate (SDS) gel using an Odyssey Fc Imager (LI-COR Biosciences). To determine the kinetic constants, the data were plotted and fitted with the GraphPad Prism 4 software. For the substrate screen and for testing whether XMPP is inhibited by guanosine or xanthosine, XMPP with a C-terminal HAStrep tag (expressed from X130) was used.

Crystallization and structure determination. Crystallization of XMPP was successful only after 13 C-terminal residues had been removed, likely because the region is disordered as predicted by the Phyre2 server[27] and the XtalPred server[28]. The corresponding cDNA (encoding Met1 to Ser250) was amplified with the primers X0001 and X0002 using a codon-optimized synthetic full-length DNA as a template (Supplementary Table 3). The product was cloned via the In-Fusion method using *Nde*I and *Xho*I into the pET41b expression vector containing a coding sequence for a C-terminal His tag and the construct was transformed into *E. coli* BL21 (DE3). Protein expression was induced after the cells had reached an optical density of 0.6–0.8 (at 600 nm) by the addition of 0.5 mM IPTG followed by continued culturing for 14–16 h at 20 °C. Collected cells were sonicated and centrifuged in buffer A (50 mM Tris, pH 8.0; 100 mM NaCl) and 1 mM MgCl$_2$. The protein was purified using a HisTrap HP column with buffer A and eluted using buffer A plus 500 mM imidazole and subsequently subjected to size exclusion chromatography using a Superdex-200 column equilibrated with buffer A plus 5% (v/v) glycerol. For solving the structure in complex with XMP, a catalytically inactive Asp12 to Ala mutant called XMPP(D12A) was purified the same way.

Purified enzymes in buffer A with 5% (v/v) glycerol were concentrated to 8–15 mg mL^{-1} and subjected to crystallization using the sitting drop vapor diffusion method at 22 °C. Crystals of XMPP in its ligand-free form were produced using a crystallization solution of 0.1 M Tris (pH 8.5) and 20% (w/v) PEG6000. For XMPP(D12A) complexed with XMP, crystals which had been grown in solution containing 0.1 M HEPES (pH 7.0) and 18% (w/v) PEG12000 were further soaked for 1 h in crystallization solution plus 5 mM XMP, 5 mM MgCl$_2$, and 20% (v/v) glycerol. Data collection was performed at 100 K on beamline 7 A at the Pohang Accelerator Laboratory (Korea) with a 0.5° oscillation angle and 20% (v/v) glycerol as cryoprotectant (Supplementary Table 2). Collected data were processed using the HKL2000 software[29] and the high-resolution cutoff was based on a CC$_{1/2}$ statistical value of approximately 0.6 (refs. [30–32]). The space group of both crystals was P2$_1$, with one monomer in the asymmetric unit. The structure of unliganded XMPP was solved by molecular replacement using the Phaser utility of PHENIX software[33]. A search model was generated by the Ensembler[34] and Sculptor utilities[35]. Specifically, the structures of the phosphatase Sdt1p from *Saccharomyces cerevisiae* (PDB: 3NUQ, 3OPX, and 3ONN) were used for the generation and truncation of a search model. The initial electron density map calculated based on a solution from the Phaser utility did not produce a high-quality map, even with 1.4-Å resolution. Extensive manual model building was required using the program COOT[36] and subsequent refinement using PHENIX[33] produced a highly reliable electron density map. During refinement, we observed unequivocal electron density for a metal ion on the $F_o - F_c$ map. Subsequently, the structure of XMPP(D12A) in complex with XMP was determined using the refined structure of unliganded XMPP as a starting model. We clearly identified electron density for XMP in the vicinity of the metal-binding site of XMPP(D12A). Details of data collection and refinement are listed in Supplementary Table 2.

Functional analysis of XMPP variants. As for the enzyme purified from plants, the Enzchek Phosphate Assay Kit (Thermo Fisher Scientific) was used for the activity assays of enzyme purified from *E. coli*.

C-terminal His-tagged full-length XMPP and various mutants were expressed and purified by affinity chromatography using a HisTrap HP column with buffer A and were eluted using buffer A containing 500 mM imidazole. A Hi-Prep 26/10 column equilibrated with buffer A was further used as final step for desalting and buffer exchange. Genes for mutant XMPPs were constructed by *Dpn*I-mediated

168

site-directed mutagenesis using mutagenic primers (Supplementary Table 3). Enzymes mutated in Asp14, His78, His82, Thr111, Asn112, and Lys159, respectively, caused problems in expression or solubility and were excluded from further analyses.

Activity and steady-state kinetic analyses were performed according to the manufacturer's instructions, but at 30 °C. We validated that all components in the reaction mixture were present at saturating concentrations and were not limiting the measured rates. In the activity assay, 18 nM XMPP was added per reaction containing 50 mM Tris (pH 7.5), 200 μM MESG, 1.25 mM MgCl$_2$, and 1 U of purine nucleoside phosphorylase. A higher XMPP concentration of 145 nM was employed for the relatively inactive variants D12A, D183A, and D184A. Activities were determined after 140 s of reaction with 125 μM XMP. For the steady-state kinetic analysis, the reaction was initiated by adding the indicated concentrations of XMP and the initial linear rate determined for the first 35 s. The K_M and V_{max} values were obtained using the SigmaPlot software (Systat Software).

Electrophoresis and immunoblotting. Proteins were extracted with buffer E (1:2 tissue to buffer ratio) and 15 μL of supernatants were separated by electrophoresis on a 12% SDS-gel and transferred by semi-dry blotting to a nitrocellulose membrane (0.45 μm pore size, Thermo Fisher Scientific). The membrane was blocked with 5% milk powder in TBS-T (20 mM TRIS-HCl, pH 7.6; 150 mM NaCl; 0.1% (v/v) Tween 20) and was incubated with custom-made rabbit polyclonal anti-XMPP antibody (1 μg mL^{-1}) in TBS-T with 0.5% (w/v) milk powder over night at 4 °C. The membrane was washed three times for 10 min with TBS-T and incubated with secondary goat anti-rabbit IgG horseradish peroxidase conjugated antibody (RABHRP1, 1:5000, Sigma-Aldrich) in TBS-T for 1 h at room temperature and washed as before. For detection, SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) was used. The chemiluminescence was detected by a Lumi-Imager F1 (Hoffmann-La Roche). For detection of YFP, membranes were incubated with monoclonal anti-GFP antibody from mouse (Roche 11814460001; clones 7.1 and 13.1; 1:15000 diluted). This antibody does not bind to mTFP1. Anti-mouse IgG conjugated to alkaline phosphatase (Sigma-Aldrich A3562; 1:10,000 diluted) was used as a secondary antibody.

Production of custom-made polyclonal anti-XMPP antibody. Untagged Arabidopsis XMPP protein was produced from construct H325 in *E. coli* strain BL21(DE3) and precipitated in inclusion bodies. Cells were grown in 0.5 L of Luria-Bertani medium to an optical density of 0.5 (at 600 nm) and then induced by 0.5 mM isopropyl-β-D-thiogalactoside. After 3 h of induction, cells were harvested, the pellet was resuspended in 27 mL of lysis buffer (50 mM Tris-HCl (pH 8.0), 0.25% (w/v) sucrose, 1 mM EDTA (pH 8.0), 7.5 mg mL^{-1} lysozyme), vortexed, and incubated on ice for 30 min. Cells were disrupted by sonication on ice and 70 mL of detergent buffer (20 mM Tris-HCl (pH 7.5), 2 mM EDTA (pH 8.0), 200 mM NaCl, 1% (w/v) deoxycholic acid, and 1% (v/v) Nonidet P-40) was added. The lysate was centrifuged at 5000 *g* for 10 min, the supernatant was removed, and the pellet was resuspended in 80 mL washing buffer (0.5% (v/v) Triton X-100 and 1 mM EDTA (pH 8.0)). Centrifugation and resuspension were repeated until a tight pellet was obtained. The pellet was washed in 80 mL of 70% ethanol (v/v), resuspended in 1 mL of freshly prepared PBS by sonication, and used for commercial antisera production in rabbit and antibody purification (immunoGlobe Antikörpertechnik GmbH).

Metabolite analysis. The extraction method was adapted from Hauck et al.[16]. For metabolite analysis, frozen plant material was homogenized with steel beads using a mixer mill MM 400 (Retsch). Tube racks were pre-cooled in liquid nitrogen. Ten milligrams of seeds were ground with one 4 mm steel bead and five 2 mm steel beads for 4.5 min at 28 s^{-1}. For seedlings, 50 mg were ground with five 2 mm steel beads for 4.5 min at 28 s^{-1}. For the extraction, 10 mM ammonium acetate buffer (pH 7.5) pre-warmed to 60 °C was used and samples were immediately transferred to a heat block at 95 °C for shaking at 1000 rpm for 10 min. For seeds and seedlings, 1 mL and 500 μL of extraction buffer were used, respectively. Following the heat treatment, samples were chilled on ice for 5 min. After centrifugation at 20,000 *g* for 10 min at 4 °C, 80% of the supernatant was transferred to a new tube to remove beads and cell debris. By centrifugation at 45,000 *g* for 15 min at 4 °C particles which might block the HPLC system were removed.

Xanthosine, xanthine, guanosine, guanine, urate and allantoate were quantified using an Agilent HPLC 1200 system with a Polaris 5 C18-A 50 ×4.6 mm column (Agilent Technologies) coupled to an Agilent 6460 C series triple quadrupole mass spectrometer. Except for urate, measurements were performed in the positive ion mode. Ammonium acetate (10 mM, pH 7.5) and 100% methanol served as solvents A and B to form the following gradient: 0 min, 5% B; 1.5 min, 5% B; 3.5 min, 15% B; 6 min, 100% B; 7 min, 100% B; 7.1 min, 5% B; and 13 min, 5% B[5]. The flow rate was 0.8 mL min^{-1} and samples of 20 μL were injected. Seed extracts were diluted 50 fold for the quantification of xanthosine, guanosine and guanine. Seedling extracts were not diluted for the quantification of xanthosine, urate and allantoate and 100-fold diluted for the quantification of guanosine. External standard dilutions spiked 1:10 into the wild-type matrix were used for quantification of xanthosine, guanosine and guanine. For the quantification of urate and allantoate, the *xmpp gsda* matrix was used. The MS parameters are listed in Supplementary Table 4. Measurements failing the quality criteria for retention time, qualifier to

quantifier ratio or with a signal to noise ratio below 10 were called "not detected". The quantification of nucleotides was performed as recently described[14,37].

Statistical analysis. Data for the determination of kinetic constants were plotted and analyzed using the GraphPad Prism 4 software. For metabolite data, the R software (version 1.2.5042) and CRAN packages multcomp and sandwich were used to perform two-sided Tukey's pairwise comparisons and to consider heteroscedasticity of the dataset with the sandwich variance estimator[38,39]. Using the same packages, two-sided Dunnett's comparisons to the reference group (XMPP + 10 μM XMP) were performed for the inhibitor test. The number of replicates, the test values and the multiplicity-adjusted *p* values are reported in the Source Data file.

Data availability
All data generated or analyzed during this study are included in this published article (and its Supplementary Information file). Mass spectrometry raw data can be supplied upon request. Atomic coordinates and structure factors of XMPP with and without XMP bound have been deposited in the Protein Data Bank (PDB) under the accession codes 7EF7 and 7EF6, respectively. Source data are provided with this paper.

References
1. Smith, P. M. C. & Atkins, C. A. Purine biosynthesis. Big in cell division, even bigger in nitrogen assimilation. *Plant Physiol.* **128**, 793–802 (2002).
2. Witte, C.-P. & Herde, M. Nucleotide metabolism in plants. *Plant Physiol.* **182**, 63–78 (2020).
3. Werner, A. K., Romeis, T. & Witte, C.-P. Ureide catabolism in Arabidopsis thaliana and Escherichia coli. *Nat. Chem. Biol.* **6**, 19–21 (2010).
4. Werner, A. K. & Witte, C.-P. The biochemistry of nitrogen mobilization: purine ring catabolism. *Trends Plant Sci.* **16**, `381–387 (2011).
5. Baccolini, C. & Witte, C.-P. AMP and GMP catabolism in Arabidopsis converge on xanthosine, which is degraded by a nucleoside hydrolase heterocomplex. *Plant Cell* **31**, 734–751 (2019).
6. Dahncke, K. & Witte, C.-P. Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in Arabidopsis. *Plant Cell* **25**, 4101–4109 (2013).
7. Yin, Y., Katahira, R. & Ashihara, H. Metabolism of purine nucleosides and bases in suspension-cultured Arabidopsis thaliana cells. *Eur. Chem. Bull.* **3**, 925–934 (2014).
8. Ashihara, H., Stasolla, C., Fujimura, T. & Crozier, A. Purine salvage in plants. *Phytochemistry* **147**, 89–124 (2018).
9. Sa, N., Rawat, R., Thornburg, C., Walker, K. D. & Roje, S. Identification and characterization of the missing phosphatase on the riboflavin biosynthesis pathway in Arabidopsis thaliana. *Plant J.* **88**, 705-716 (2016).
10. Xu, Y.-F. et al. Nucleotide degradation and ribose salvage in yeast. *Mol. Syst. Biol.* **9**, 665 (2013).
11. Jung, B., Hoffmann, C. & Möhlmann, T. Arabidopsis nucleoside hydrolases involved in intracellular and extracellular degradation of purines. *Plant J.* **65**, 703–711 (2011).
12. Jung, B. et al. Uridine-ribohydrolase is a key regulator in the uridine degradation pathway of Arabidopsis. *Plant Cell* **21**, 876–891 (2009).
13. Riegler, H., Geserick, C. & Zrenner, R. Arabidopsis thaliana nucleosidase mutants provide new insights into nucleoside degradation. *N. Phytol.* **191**, 349–359 (2011).
14. Straube, H., Niehaus, M., Zwittian, S., Witte, C.-P. & Herde, M. Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. *Plant Cell* **33**, 270–289 (2020).
15. Schroeder, R. Y., Zhu, A., Eubel, H., Dahncke, K. & Witte, C.-P. The ribokinases of Arabidopsis thaliana and Saccharomyces cerevisiae are required for ribose recycling from nucleotide catabolism, which in plants is not essential to survive prolonged dark stress. *N. Phytol.* **217**, 233–244 (2018).
16. Hauck, O. K. et al. Uric acid accumulation in an Arabidopsis urate oxidase mutant impairs seedling establishment by blocking peroxisome maintenance. *Plant Cell* **26**, 3090–3100 (2014).

17. Seifried, A., Schultz, J. & Gohla, A. Human HAD phosphatases: structure, mechanism, and roles in health and disease. *FEBS J.* **280**, 549–571 (2013).

18. Alonso, J. M. et al. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* **301**, 653–657 (2003).

19. Kleinboelting, N., Huep, G., Kloetgen, A., Viehoever, P. & Weisshaar, B. GABI-Kat simple search. New features of the Arabidopsis thaliana T-DNA mutant database. *Nucleic Acids Res.* **40**, D1211–D1215 (2012).

20. Witte, C.-P., Noël, L. D., Gielbert, J., Parker, J. E. & Romeis, T. Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope. *Plant Mol. Biol.* **55**, 135–147 (2004).

21. Chen, M. & Witte, C.-P. A kinase and a glycosylase catabolize pseudouridine in the peroxisome to prevent toxic pseudouridine monophosphate accumulation. *Plant Cell* **32**, 722–739 (2020).

22. Geldner, N. et al. Rapid, combinatorial analysis of membrane compartments in intact plants with a multicolor marker set. *Plant J.* **59**, 169–178 (2009).

23. Myrach, T., Zhu, A. & Witte, C.-P. The assembly of the plant urease activation complex and the essential role of the urease accessory protein G (UreG) in delivery of nickel to urease. *J. Biol. Chem.* **292**, 14556–14565 (2017).

24. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282 (1992).

25. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).

26. Werner, A. K., Sparkes, I. A., Romeis, T. & Witte, C.-P. Identification, biochemical characterization, and subcellular localization of allantoate amidohydrolases from Arabidopsis and soybean. *Plant Physiol.* **146**, 418–430 (2008).

27. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protocol* **10**, 845–858 (2015).

28. Slabinski, L. et al. XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics* **23**, 3403–3405 (2007).

29. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).

30. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D.* **67**, 235–242 (2011).

31. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).

32. Diederichs, K. & Karplus, P. A. Better models by discarding data? *Acta Crystallogr. D* **69**, 1215–1222 (2013).

33. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

34. Terwilliger, T. C. et al. Interpretation of ensembles created by multiple iterative rebuilding of macromolecular models. *Acta Crystallogr. D* **63**, 597–610 (2007).

35. Bunkóczi, G. & Read, R. J. Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr. D* **67**, 303–312 (2011).

36. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

37. Straube, H., Witte, C.-P. & Herde, M. Analysis of nucleosides and nucleotides in plants: an update on sample preparation and LC-MS techniques. *Cells* **10**, 689 (2021).

38. Herberich, E., Sikorski, J. & Hothorn, T. A robust procedure for comparing multiple means under heteroscedasticity in unbalanced designs. *PloS ONE* **5**, e9788 (2010).

39. Pallmann, P. & Hothorn, L. A. Analysis of means. A generalized approach using R. *J. Appl. Stat.* **43**, 1541–1560 (2016).

Author contributions

C.-P.W. devised the project and supervised the work except for the generation of the crystal structure and the biochemical analysis of enzyme mutants, which was supervised by S.R. C.-P.W., S.R., K.J.H., and S.-Y.Y. designed the experiments and interpreted the results. K.J.H. performed the biochemical analyses of XMPP produced in plants, generated the Arabidopsis mutants, and performed the metabolite analyses of nucleosides and purine catabolic products. H.S. performed the metabolite analysis of nucleotides. N.M.-E. performed the confocal microscopy and the characterization of the T-DNA lines. C.-P.W., M.V.-H., and M.H. made bioinformatic analyses and M.V.-H. performed preliminary assays leading to the discovery of XMPP. C.-P.W. performed the phylogenetic analysis. S.-Y.Y. determined the crystal structure and made the biochemical analyses of the XMPP variants. K.J.H. conducted most of the statistical analyses. K.J.H., C.-P.W., and S.-Y.Y. prepared the figures. C.-P.W. and S.R. wrote the manuscript with the contribution of K.J.H. and S.-Y.Y. The manuscript was revised by all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-27152-4.

Correspondence and requests for materials should be addressed to Sangkee Rhee or Claus-Peter Witte.

Peer review information *Nature Communications* thanks Santiago Ramón-Maiques and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Initiation of cytosolic Plant Purine Nucleotide Catabolism

involves a monospecific Xanthosine Monophosphate Phosphatase

Katharina J. Heinemann, Sun-Young Yang, Henryk Straube, Nieves Medina-

Escobar, Marina Varbanova-Herde, Marco Herde, Sangkee Rhee, Claus-Peter Witte

SUPPLEMENTARY INFORMATION
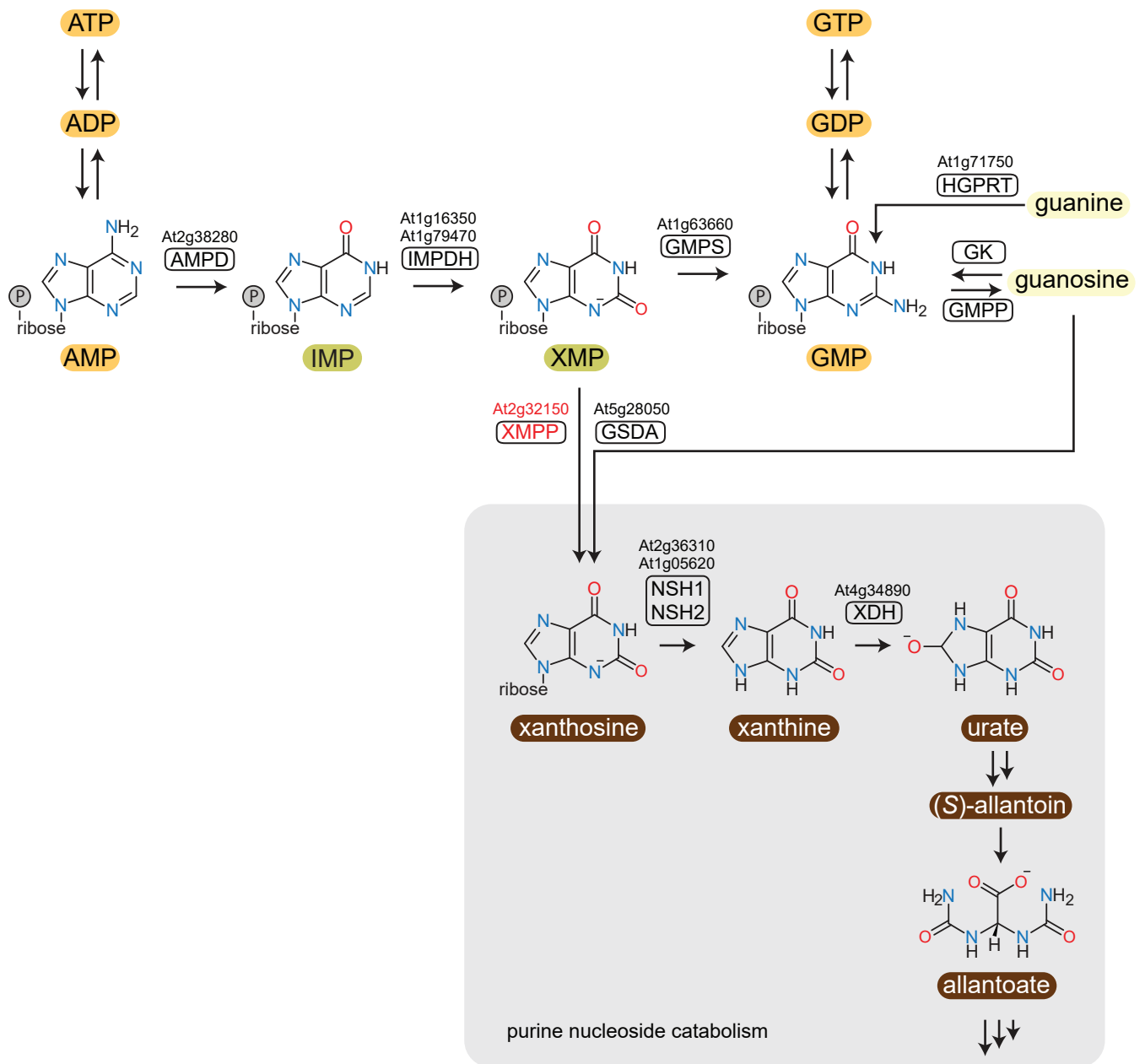
Supplementary Fig. 1. Extended scheme of GMP biosynthesis and purine nucleotide catabolism. Labels are as in Fig. 1a. The possible interconversions of the adenosine and guanosine triphosphates, diphosphates and monophosphates are additionally shown. Bases are drawn in their ionic form when they are mainly present as ions at neutral pH.
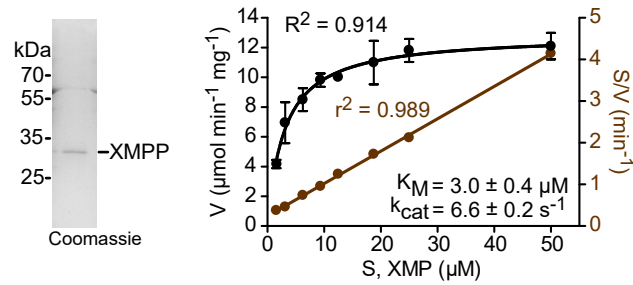
Supplementary Fig. 2. Purity and determination of kinetic constants of N-terminal tagged XMPP. Left panel, Coomassie-stained SDS-gel with the N-terminal Strep-tagged XMPP purified from leaf extracts of *Nicotiana benthamiana* after transient expression. Right panel, determination of the kinetic constants with the data fitted according to Michaelis Menten (left axis) or Hanes (right axis). Error bars are SD, n = 3 independent measurements using the same enzyme preparation.

a Locus At2g32150

SALK 131244　　　　SALK 067037
xmpp-2　　　　　　*xmpp-1*

100 bp

ATG　　　　　　　　　　　　　　　　　　　　TGA

N188　　　N232　　　　　　N233　　　　　　N189

b RT-PCR

wild type
xmpp-1
xmpp-2

bp

375 —　　　　　　　　N188
　　　　　　　　　　　N233

529 —　　　　　　　　1034　　*Actin-2* (At3g18780)
　　　　　　　　　　　1033

c Immunoblot

Col-0
xmpp gsda::C
xmpp nsh1 gsda::C
xmpp
xmpp gsda
xmpp nsh1 gsda

Twin-Strep-XMPP →
XMPP →

d Transgene encoding TwinStrep tagged XMPP

start　　　　　　　　　　　　　　StrepII

...CTTGTACTCGAGTCTCAAA**ATG**GGTTCTGCATGGTCTCATCCTCAATTTGAAAAAGGCGCCGGTT

StrepII

CCGGTGGGGGCTCCGGGGGAAGTGCATGGTCCCATCCGCAGTTCGAGAAAGGTGGTTCCATGGAT...

ATG　　　　　　　　　　　　　　　　TGA

100 bp

Supplementary Fig. 3. Characterization of plant lines with T-DNA insertions in *XMPP* and lines that express an *XMPP* transgene. a Coding region of the XMPP gene with exons shown as boxes and T-DNA insertion positions marked by triangles. Primers are indicated by arrows. **b** Semi-quantitative RT-PCR data showing the absence of intact *XMPP* mRNA in the two T-DNA insertion lines (representative result out of two repeats). The primers used for amplification are indicated at the right of the panels. The *actin2* mRNA (encoded at locus At3g18780) was used as positive control. **c** Immunoblot of seedling extracts from lines that express *XMPP* transgenes and from control lines (representative result out of two repeats). The blot was probed with a custom-made anti-XMPP antibody. The complementation lines contain a transgene (::C) encoding an N-terminal TwinStrep-tagged XMPP expressed by the native promoter. The positions of native untagged XMPP (in the wild type) and of tagged XMPP are indicated. **d** Structure of the *XMPP* transgene (exons are shown as orange boxes) with the inserted sequence including the coding sequence for the N-terminal TwinStrep tag (brown box or bar). The promoter region is shown in green.
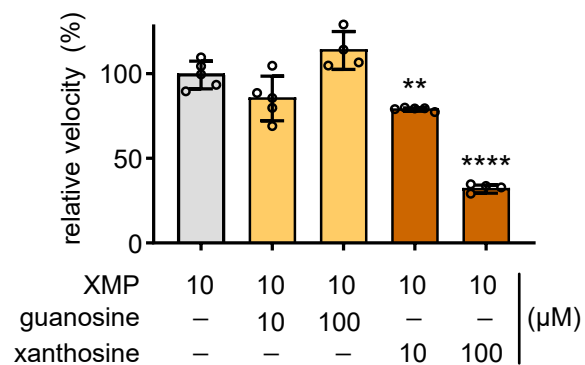
Supplementary Fig. 4. Inhibitory effect of guanosine or xanthosine on XMPP activity. The enzymatic velocity of XMPP was assessed with 10 μM XMP without any nucleosides (set to 100% at an activity of 14.0 ± 1.0 μmol mg^{-1} min^{-1}) or in the presence of 10 or 100 μM guanosine or xanthosine, respectively. Error bars are SD, n = 5 but for samples with 100 μM nucleosides n = 4 repeated assays using the same enzyme preparation. For statistical treatment, two-sided Dunnett's comparisons using the sandwich variance estimator were performed. The assay without nucleoside additions served as reference group. **, $p < 0.01$; ****, $p < 0.0001$.
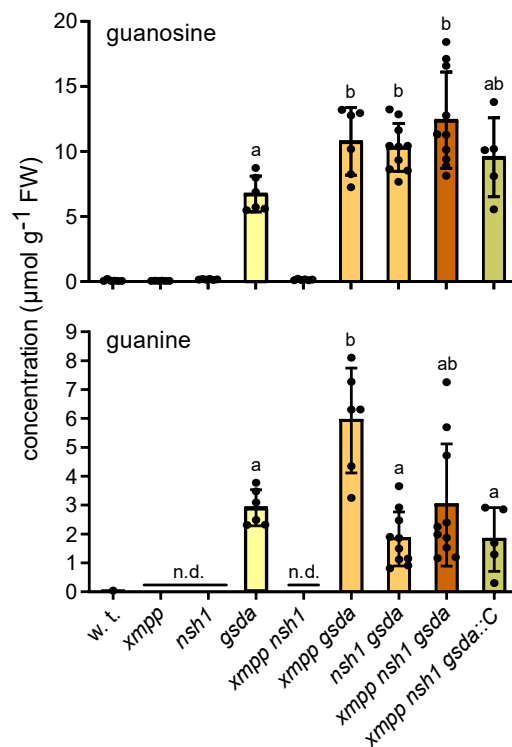
Supplementary Fig. 5. Guanosine and guanine pool sizes in *XMPP* genetic variants in the context of other mutants of purine catabolism genes in seeds. Quantification of guanosine and guanine in seed extracts of the same seeds as used in the experiment shown in Fig. 2d. The complementation line (*xmpp nsh1 gsda::C*) contains a transgene encoding an N-terminal TwinStrep-tagged XMPP expressed from a native promoter (Supplementary Fig. 3c,d). Error bars are SD; n = 6 but for *xmpp nsh1 gsda::C* n = 5 and for *nsh1 gsda* and *xmpp nsh1 gsda* n = 10. Seed samples from independent mother plants grown in parallel. Statistical analysis as in Fig. 2d.

Guanosine and guanine hyperaccumulate in *xmpp gsda* background, because additional guanosine is produced via GMP from XMP whose degradation route is blocked. Although guanine is not an intermediate of purine nucleotide catabolism in *Arabidopsis*[5], it is generated from guanosine, in part by the NSH1/NSH2 complex[2], when this metabolite accumulates in *gsda* background. A hyperaccumulation of guanosine (but not guanine) is observed also in *nsh1 gsda* seeds[2], because guanosine is less efficiently hydrolyzed to guanine and ribose due to the lack of NSH1 and because in this double mutant xanthosine accumulates (Fig. 2d) which blocks XMPP activity biochemically (Supplementary Fig. 4).
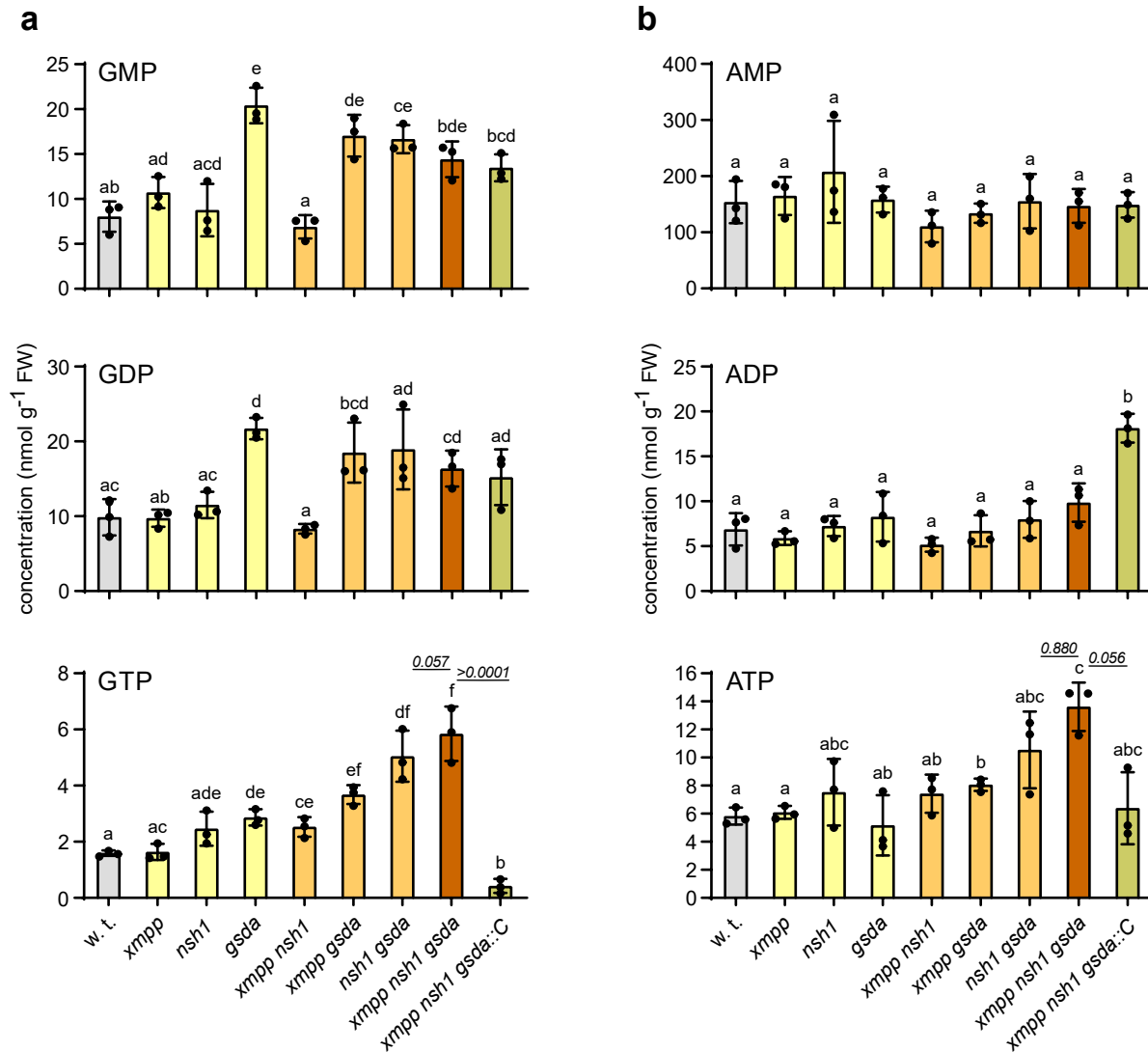
Supplementary Fig. 6. Nucleotide pool sizes in *XMPP* genetic variants in the context of other mutants of purine catabolism genes in seeds. Quantification of purine nucleotides in seed extracts of the same seeds as used in the experiment shown in Fig. 2d. **a**, GMP, GDP, GTP; **b**, AMP, ADP, ATP. Error bars are SD; n = 3 seed samples from independent mother plants grown in parallel. Statistical analysis as in Fig. 2d. Some p values are indicated above the columns in italic numbers, all p values can be found in the Source Data file.

GTP concentrations in *xmpp nsh1 gsda* seeds are even higher than in *xmpp gsda* seeds probably because guanosine hydrolysis to guanine and ribose is partially blocked in *nsh1* background (see Supplementary Fig. 5). Since guanosine degradation is compromised, more GMP is channeled into GTP.

In *nsh1 gsda* seeds, GTP concentrations are also very high, because also here guanosine hydrolysis is partially compromised. Additionally, the high xanthosine accumulation in *nsh1* (Fig. 2d) blocks XMPP biochemically (Supplementary Fig. 4) resulting in a partial molecular mimic of the *xmpp nsh1 gsda* background.
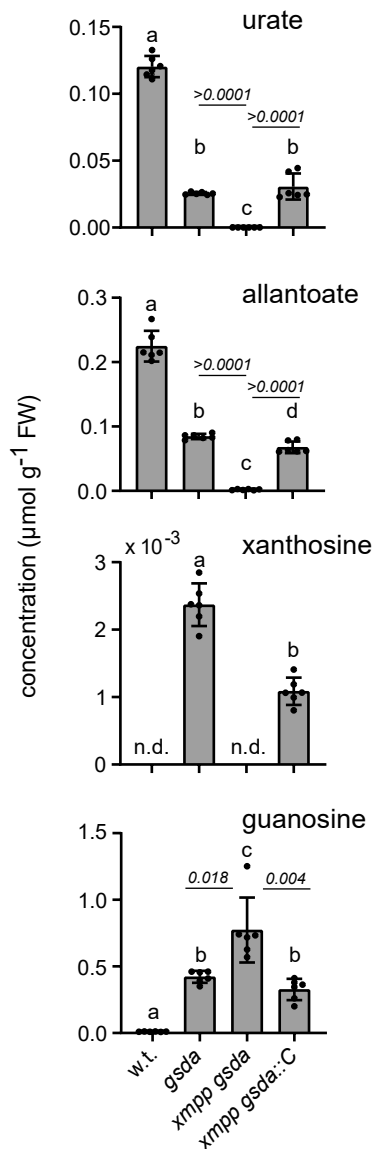
Supplementary Fig. 7. Characterization of metabolic alterations in XMPP genetic variants in the context of other mutants of purine catabolism genes in seedlings under long-day or prolonged-night conditions. This dataset includes the *xmpp gsda::C* line and is independent from similar data shown in Fig. 3. The seedlings were grown in a constant 16 hour day / 8 hour night regime up to day 5 after germination. At the end of the night of day 5, they were exposed to darkness for 48 hours. All genotypes were grown on the same plate for one biological replicate (n). Error bars are SD, n = 6.

In the complementation line *xmpp gsda::C*, urate, allantoate and guanosine pools are of similar size as in the *gsda* line, showing that the complementation by the *XMPP* transgene is effective. Comparatively small amounts of xanthosine accumulated in *gsda* background probably because the NSH1 /NSH2 nucleoside hydrolase complex is partially inhibited by the strong accumulation of guanosine in that line. This xanthosine accumulation is not observed in *xmpp gsda*, because the additional mutation of *XMPP* prevents any influx into the xanthosine pool. In the *xmpp gsda::C* complementation line, xanthosine accumulation is observed again, although the pool size is smaller than in *gsda*. The reason for this difference is unclear and might lie in the slightly different expression level of the *XMPP* transgene compared to the native *XMPP* gene (Supplementary Fig. 3c).
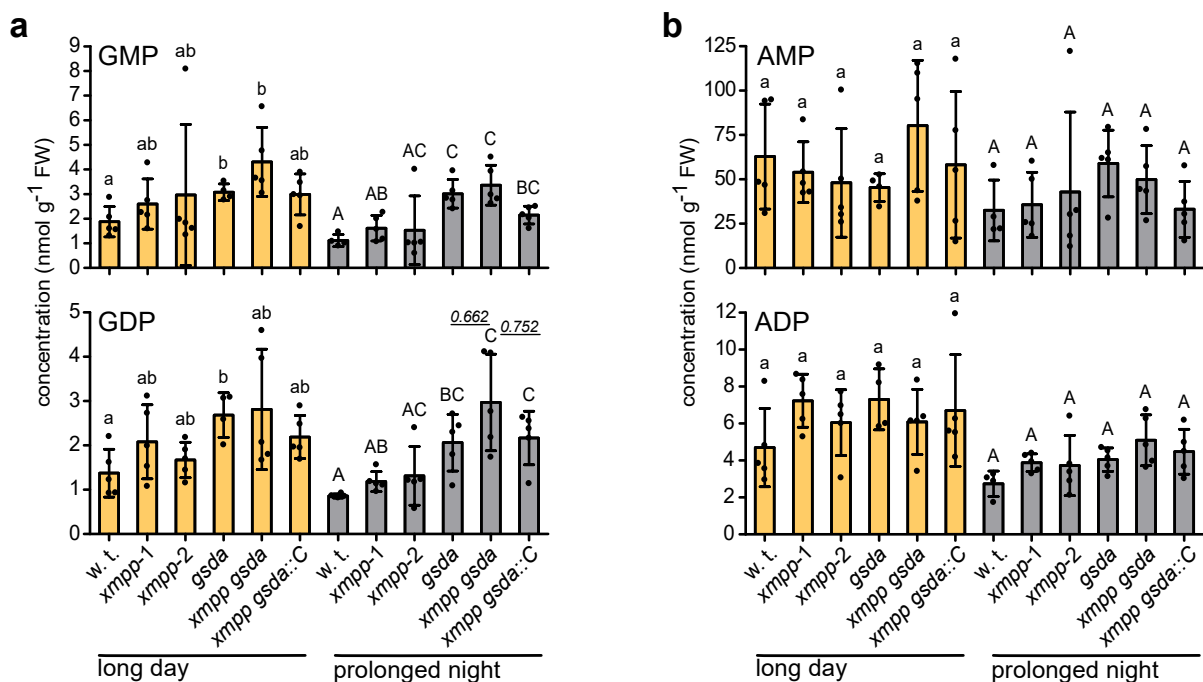
Supplementary Fig. 8. Nucleotide pool sizes in XMPP genetic variants in the context of other mutants of purine catabolism genes in seedlings under long-day or prolonged-night conditions. a GMP, GDP concentrations and **b** AMP, ADP concentrations in seedling extracts from the experiment shown in Fig. 3b. Error bars are SD, n = 5 but for *gsda* in the light and the wild type in the dark n = 4. Statistical analysis as in Fig. 3b.

The in tendency higher GTP concentration in *xmpp gsda* background (Fig. 3b) is weakly reflected at the level of GMP or GDP but not observed for AMP or ADP. Guanylates and also adenylates mostly accumulate as GTP and ATP and not in lower phosphorylation states.
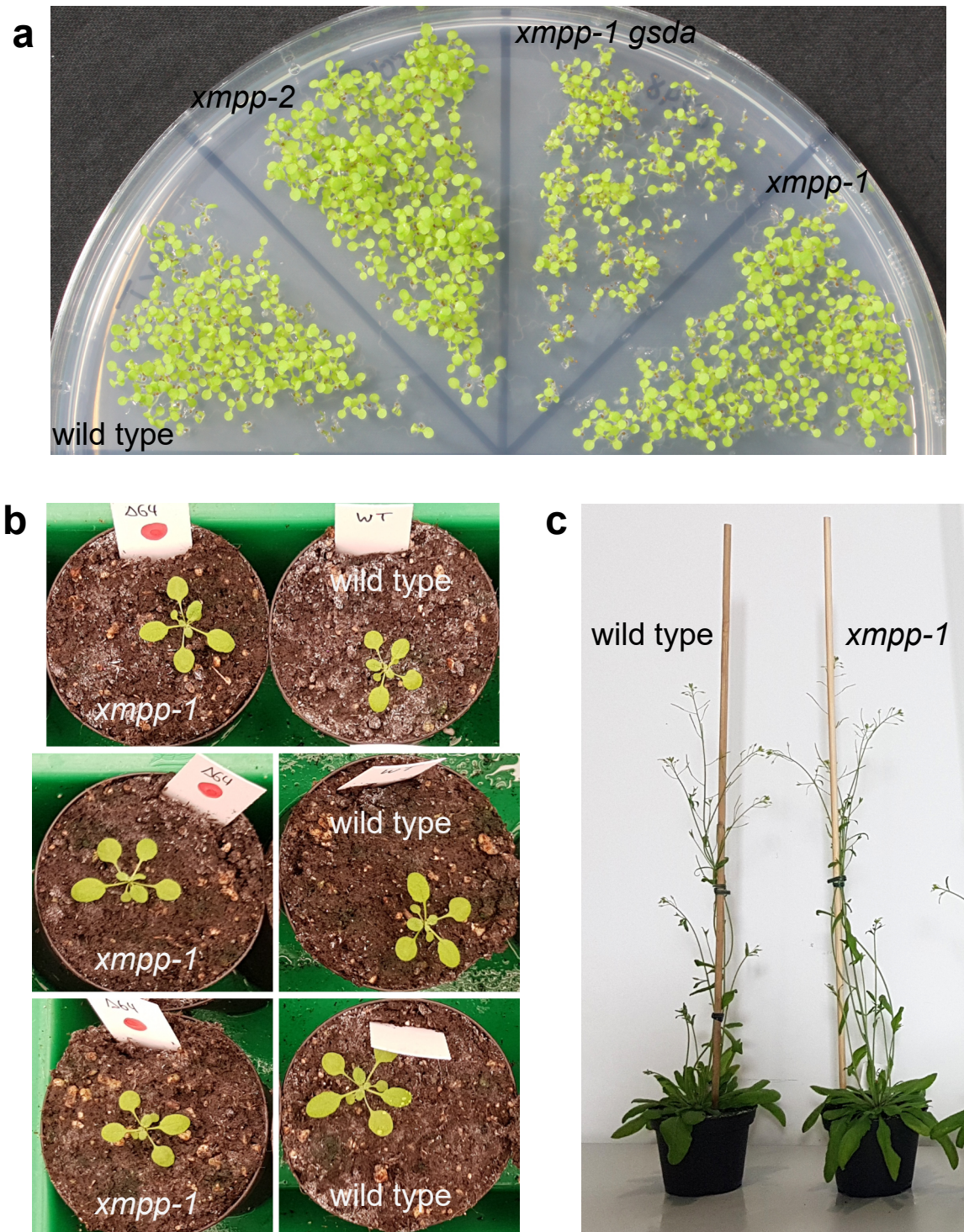
Supplementary Fig. 9. Growth of *XMPP* mutants in comparison to the wild type. All plants were grown under long day conditions with a 16 h light period. **a** 5-day-old seedlings on a half-strength Murashige and Skoog medium without sucrose and vitamins. The mutation of *GSDA* (here in *xmpp* background) is known to lead to a delay in germination[15]. **b** 17-day-old plants grown in a growth chamber. Six trays of 28 pots each containing plants with different genotypes were grown in a randomized fashion to obtain a homogenous seed batch. Wild-type and *xmpp-1* plants from one of those trays are shown. The whole tray is shown in the Source Data file. **c** Representative photograph of a wild type next to a *xmpp-1* plant in the flowering stage (40 days after germination).
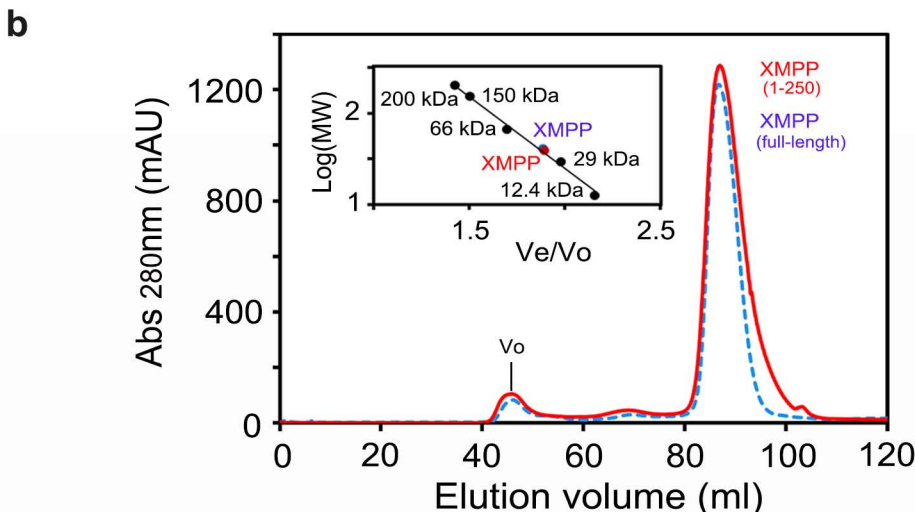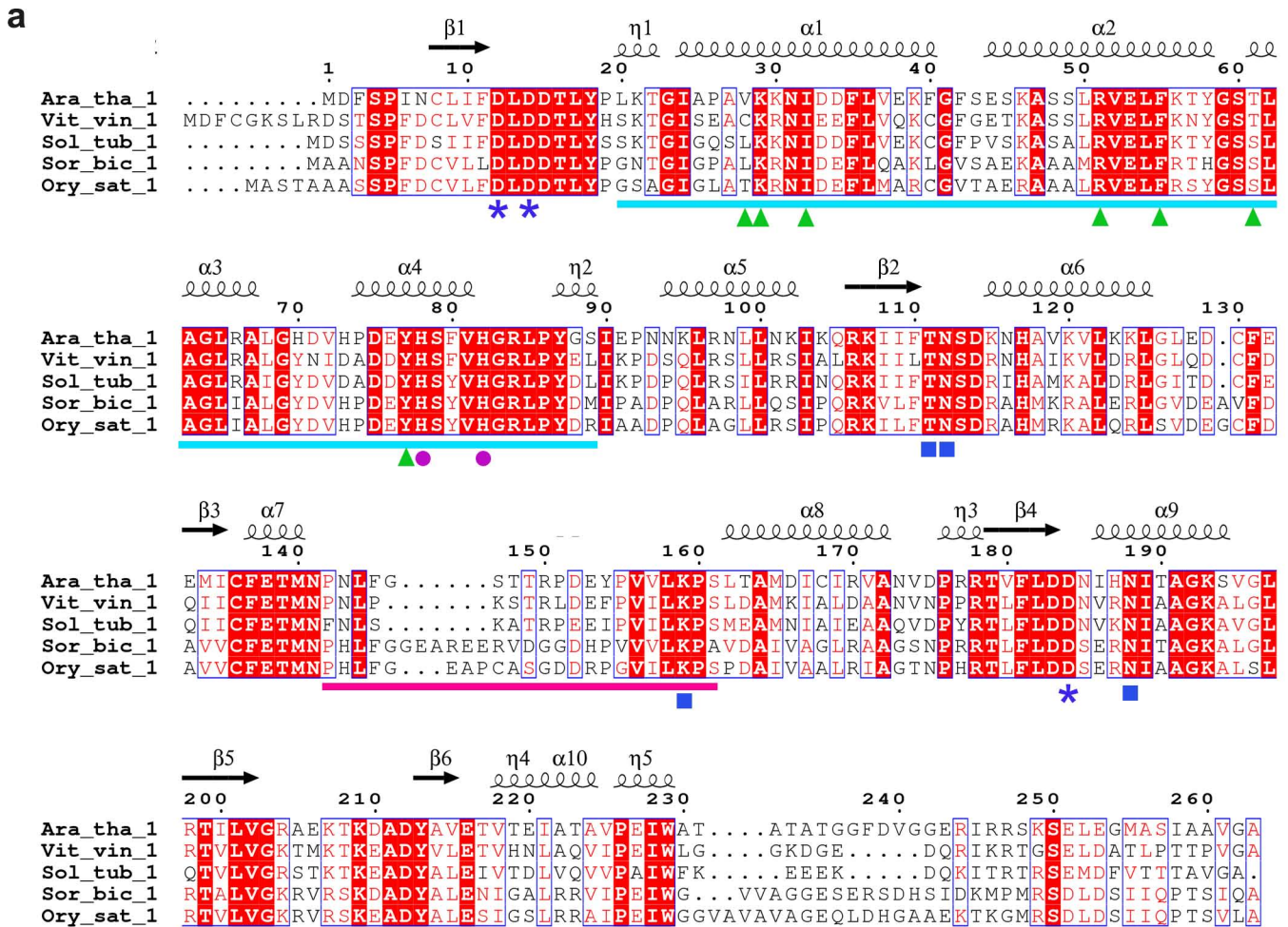
Supplementary Fig. 10. Sequence alignment and size exclusion chromatography of XMPP. **a** The amino acid sequence of XMPP from Arabidopsis (Ara_tha_1) aligned to the putative orthologs from grape (*Vitis vinifera*, Vit_vin_1), potato (*Solanum tuberosum*, Sol_tub_1), great millet (*Sorghum bicolor*, Sor_bic_1) and rice (*Oryza sativa*, Ory_sat_1). These sequences are also part of the analysis shown in Fig. 1b and are present in the XMPP/SDTL1 clade. Highly conserved residues are shown in red and boxed in blue, while strictly conserved residues are highlighted with a red background. Secondary structural elements defined in the unliganded XMPP are shown above the alignment. The cap domain and the protruding loop shown in Fig. 4a are indicated by a cyan and a red bar, respectively. Metal-binding residues are indicated by blue asterisks. Residues interacting with the phosphate, ribose and xanthine moieties of XMP are indicated by blue squares, magenta circles, and green triangles, respectively. This figure was prepared using the ESPript software[36] **b** Elution profiles from size exclusion chromatography of full-length XMPP (dashed blue line, calculated molecular weight of 29 kDa) and of the C-terminally truncated XMPP used for crystallization (solid red line, residues 1 to 250). The proteins were resolved on a Superdex 200 column (GE Healthcare) equilibrated in 50 mM Tris (pH 8.0), 100 mM NaCl, and 5% (v/v) glycerol. The void volume (V_o) was about 45 ml and the XMPP peaks corresponded to a size of ~33 kDa when calibrated with molecular weight standards of 12 to 200 kDa (insert).
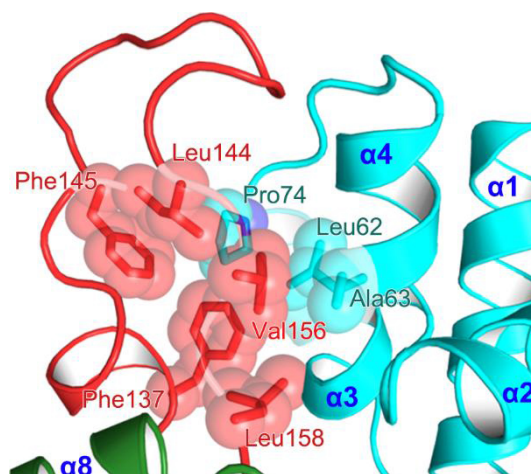
Supplementary Fig. 11. Interaction of the cap and protruding loop domains of XMPP. Hydrophobic residues with their surfaces are displayed in the unliganded XMPP. These are clustered at the interface between the cap domain and the protruding loop. The protruding loop (red) from the β3-α8 unit in the central core domain mediates extensive interactions with α3 and the N-terminal region of α4 (cyan) in the cap domain. Hydrophobic contacts of residues within 4.0 Å dominate the cap-loop interaction. Note, that the interacting hydrophobic residues are highly conserved in the XMPP homologs (Supplementary Fig. 10a).
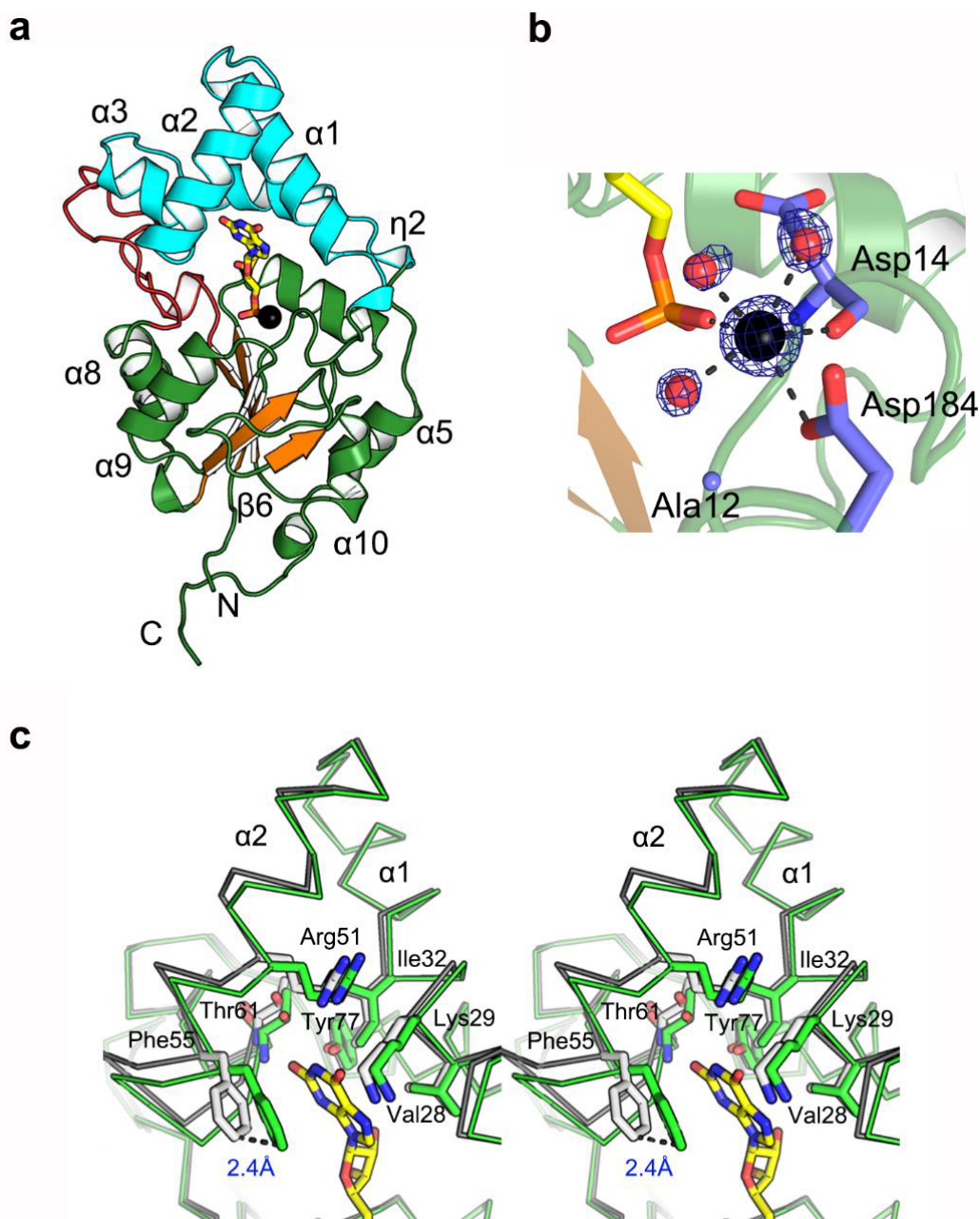
Supplementary Fig. 12. Structure of XMPP(D12A) in complex with XMP. **a** An overall structure of the XMPP(D12A)-XMP complex with a ball-and-stick presentation for XMP. Labels and color codes are identical to those in Fig. 4a. **b** Mg^{2+}-binding site in the XMPP(D12A)-XMP complex. Due to the mutation of D12A, the coordination of the metal ion (black sphere) is not in a square bi-pyramidal shell. Two water molecules above the metal ion are overlaid with a *2Fo-Fc* electron density contoured at 2.0 σ, with the other water molecules across Asp14 at 1.5 σ and the metal ion at 3.0 σ, respectively. **c** Comparison of the xanthine-binding environment of XMP for the unliganded XMPP (gray) and the complex (green). The two structures were superimposed using the central core domain, resulting in a root-mean square deviation (RMSD) of 0.66 Å for 145 Cα atoms. Note, there are no significant conformational changes in the binding environment, and the largest displacement of approximately 2.4 Å is observed for the CZ of Phe55. This region contains Phe55 and Thr61 that interact with the xanthine moiety (Fig. 4d). In the presence of XMP it shifts its position toward XMP without significantly perturbing any side chain conformations. Further analysis indicated that XMP is almost buried in the complex.
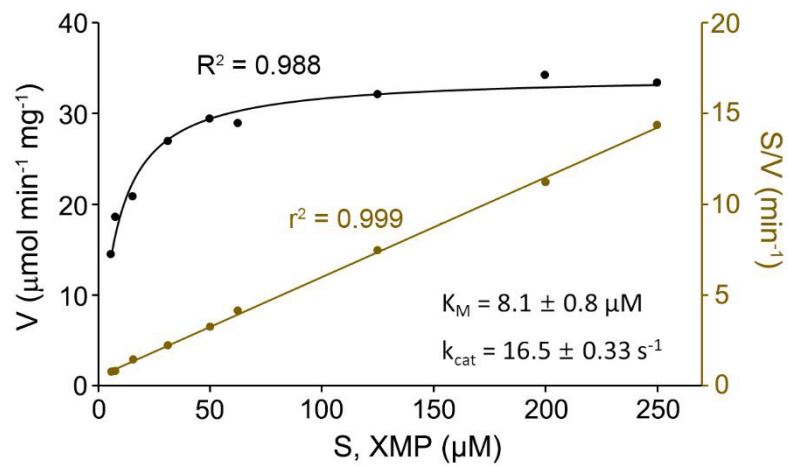
Supplementary Fig. 13. Steady-state kinetic analysis of C-terminal His-tagged full-length XMPP. The enzyme was expressed in *E. coli* and the enzyme assay performed at 30°C. Kinetic constants were determined by carrying out reactions containing 18 nM wild-type XMPP with 0 to 250 µM XMP.

Supplementary Table 1. Species names, protein abbreviations and corresponding locus identifiers (in Phytozome V12.1) for proteins used in the phylogenetic analysis

species	abbreviation	locus identifier	remark
Ananas comosus	Ana_com_1	Aco018447.1	Manually corrected according to consensus. Intron with GC acceptor site postulated.
	Ana_com_2a	Aco017201.1	
	Ana_com_2b	Aco013639.1	
	Ana_com_3A	Aco016820.1	
	Ana_com_3B	Aco001439.1	
Aquilegia coerulea	Aqu_coe_1	Aqcoe5G056900.1	
	Aqu_coe_2	Aqcoe2G238600.2	
	Aqu_coe_3	Aqcoe1G108100.2	
	Aqu_coe_3/4	Aqcoe2G223000.1	
Arabidopsis thaliana	Ara_tha_1	At2g32150.1	
	Ara_tha_2	At3g62040.1	
	Ara_tha_3	At5g02230.1	
	Ara_tha_4A	At5g59480.1	
	Ara_tha_4B	At5g59490.1	
Brachypodium distachyon	Bra_dis_1	Bradi1g11620.1	
	Bra_dis_2A	Bradi1g20280.2	
	Bra_dis_2B	Bradi1g02690.1	
	Bra_dis_3A	Bradi1g66710.1	
	Bra_dis_3B	Bradi2g62150.1	
Capsella rubella	Cap_rub_1	Carubv10025064m	
	Cap_rub_2	Carubv10019512m	Sequence manually corrected. Second intron has unusual GA-AG splice junctions and is not correctly annotated
	Cap_rub_3a	Carubv10001663m	
	Cap_rub_3b	Carubv10023829m	
	Cap_rub_4A	Carubv10026922m	
	Cap_rub_4Ba	Carubv10027751m	
	Cap_rub_4Bb	Carubv10028046m	
Carica papaya	Car_pap_1	evm.model.supercontig_58.124	
	Car_pap_2	evm.model.supercontig_57.45	

	Car_pap_3	evm.model.supercontig_169.20	
	Car_pap_4A	evm.model.supercontig_81.119	
	Car_pap_4B	evm.model.supercontig_81.118	
Citrus sinensis	Cit_sin_1	orange1.1g025190m	
	Cit_sin_2	orange1.1g024578m	
	Cit_sin_3	orange1.1g022360m	
	Cit_sin_4A	orange1.1g036723m	
	Cit_sin_4B	orange1.1g035566m	Genomic sequence has a gap. Translated mRNA sequence (accession XM_025102177.1) was used instead.
Medicago truncatula[1]	Med_tru_1	Medtr1g104610.1	
	Med_tru_2	Medtr1g085570.1	
	Med_tru_3a	Medtr4g021800.1	
	Med_tru_3b	Medtr7g062150.1	
	Med_tru_4a	Medtr2g097520.1	
	Med_tru_4b	Medtr4g064380.1	
Mimulus guttatus	Mim_gut_1	Migut.G00328.1	
	Mim_gut_2	Migut.L01000.1	
	Mim_gut_3a	Migut.L01953.1	
	Mim_gut_3b	Migut.J00320.1	
Oryza sativa	Ory_sat_1	LOC_Os03g49440.1	
	Ory_sat_2A	LOC_Os07g44060.1	
	Ory_sat_2B	LOC_Os03g61829.1	
	Ory_sat_3A	LOC_Os03g16670.1	
	Ory_sat_3B	LOC_Os01g74152.1	
Physcomitrium patens	Phy_pat_1	Pp3c27_8130V3.2	
Selaginella moellendorfii	Sel_moe_1	73861	Manually corrected according to consensus.
Setaria italica[1]	Set_ita_1	Seita.9G127000.1	
	Set_ita_2Aa	Seita.2G403800.1	
	Set_ita_2Ab	Seita.2G403900.1	
	Set_ita_2B	Seita.9G022100.1	
	Set_ita_3A	Seita.9G453700.1	
	Set_ita_3B	Seita.5G468600.1	
Solanum tuberosum	Sol_tub_1	PGSC0003DMT400081430	

	Sol_tub_2	PGSC0003DMT400002201	
	Sol_tub_3	PGSC0003DMT400022761	
	Sol_tub_4	PGSC0003DMT400078186	
Sorghum bicolor	Sor_bic_1	Sobic.001G125000.1	
	Sor_bic_2A	Sobic.002G388400.1	
	Sor_bic_2B	Sobic.001G027300.1	
	Sor_bic_3A	Sobic.001G422100.1	
	Sor_bic_3B	Sobic.003G444200.1	
Spagnum phallax	Spa_pha_1	Sphfalx0274s0003.1	
Vitis vinifera	Vit_vin_1	GSVIVT01017946001	
	Vit_vin_2	GSVIVT01028227001	Manually corrected according to consensus.
	Vit_vin_3	GSVIVT01034151001	Manually corrected according to consensus.
	Vit_vin_4	GSVIVT01024928001	
Zostera marina	Zos_mar1	Zosma185g00220.1	
	Zos_mar2	Zosma4g00550.1	
	Zos_mar3	Zosma241g00050.1	

[1] Sequences with strong deviations from consensus were not considered in the analysis: Medtr4g021825.1; Seita.2G403700.1.

Supplementary Table 2. Data collection and refinement statistics

Data set	XMPP	XMPP(D12A) complexed with XMP
PDB ID	7EF6	7EF7
Data collection		
Wavelength (Å)	0.97933	0.97933
Resolution (Å)	50.0 – 1.34 (1.39-1.34)[a]	50.0 – 1.50 (1.53-1.50)[a]
Unique reflections	51,330 (4,973)	37,816 (1,715)
Multiplicity	5.4 (4.9)	3.5 (2.9)
Completeness (%)	98.9 (95.4)	98.3 (91.4)
Mean I/sigma(I)	31.1 (3.0)	15.2 (1.0)
Wilson B-factors (Å^2)	16.4	20.5
R-merge	0.057 (0.391)	0.086 (0.875)
CC$_{1/2}$[b]	0.993 (0.969)	0.989 (0.622)
Space group	$P2_1$	$P2_1$
Unit cell a, b, c (Å)	37.1, 55.3, 57.2	37.7, 56.8, 57,0
α , β , γ (º)	90, 93.9, 90	90, 92.9, 90
Refinement		
R-work[c]	0.190	0.213
R-free[d]	0.226	0.240
No. of atoms		
Macromolecules	1828	1834
Ligands	1	25
Water	192	80
RMS (bonds) (Å)	0.012	0.009
RMS (angles) (º)	1.42	1.17
Ramachandran favored (%)	99.6	98.3
Ramachandran outliers (%)	0.0	0.0
Average B-factors (Å^2)		
Macromolecules	22.6	50.6
Ligands	16.3	43.2
Water	30.7	53.3

[a]Numbers in parentheses refer to data in the highest resolution shell.
[b]The CC$_{1/2}$ is the Pearson correlation coefficient (CC) calculated from each subset containing a random half of the measurements of unique reflection
[c]R_{work} = Σ ||F_{obs}|-|F_{cal}|| / Σ |F_{obs}|
[d]R_{free} is the same as R_{obs} for a selected subset (5%) of the reflections that was not included in prior refinement calculations.

Supplementary Table 3. Primer list and synthetic XMPP sequence

Primer	Sequence
N188	TCCATGGATTTCTCTCCGATCAACTGC
N189	ACCCGGGTCACGCACCGACGGCTGC
N359	TATCGATAAAATGGATTTCTCTCCGATCAACTGC
N360	TCCCGGGCGCACCGACGGCTGCTATT
1033	GTGAACGATTCCTGGACCTGCCTC
1034	GAGAGGTTACATGTTCACCACAAC
1709	TCGAGCTCAAAATGGGTTCTGCATGGTCTCATCCTCAATTTGAAAAAGGCGCCACCATGG
1710	AATTCCATGGTGGCGCCTTTTTCAAATTGAGGATGAGACCATGCAGAACCCATTTTGAGC
P243	TCCCGGGATGGTGAGCAAGGGCGAGGAGACCAC
P244	TGAGCTCTCACTTGTACAGCTCGTCCATGCCGTCGG
P1094	TGGCGCGCCTGCTGGTTAGATAGATTCATGGC
P1097	GCGCCGGTTCCGGTGGGGGCTCCGGGGGAAGTGCATGGTCCCATCCGCAGTTCGAGAAAGGTGGTTC
P1098	CATGGAACCACCTTTCTCGAACTGCGGATGGGACCATGCACTTCCCCCGGAGCCCCCACCGGAACCG
P1103	ACTCGAGTACAAGAGAACAAGTAAGTAAGAAAATATA
X0001	GGAGATATACATATGGATTTCAGCCCGATCACAA
X0002	GTGGTGGTGCTCGAGGCTTTTGCTACGACGGATA

Primers for the generation of point mutations

Primer	Name indicating mutation site	Sequence (mutation site in bold)
X0121	D12A_F	GCCCGATCAACTGCCTGATCTTT**GCG**CTGGATGACA
X0122	D12A_R	TGTCATCCAG**CGC**AAAGATCAGGCAGTTGATCGGGC
X0141	D14A_F	CTTTGACCTG**GCG**GACACCCTGTATCCGCTGAAGAC
X0142	D14A_R	GTCTTCAGCGGATACAGGGTGTC**CGC**CAGGTCAAAG
X0281	V28A_F	GTATCGCGCCGGCC**GCG**AAAAAAAAACATCGATGAC
X0282	V28A_R	GTCATCGATGTTTTTTTT**CGC**GGCCGGCGCGATAC
X0291	K29A_F	AAGACGGGTATCGCGCCGGCCGTG**GCA**AAAAACAT
X0292	K29A_R	ATGTTTTT**TGC**CACGGCCGGCGCGATACCCGTCTT
X0295	K29R_F	CCGGC**CGT**GCGTAAAAACATCGATGACTTCCTGG
X0266	K29R_R	CCAGGAAGTCATCGATGTTTTTACGC**ACG**GCCGG
X0321	I32A_F	CCGGCCGTGAAAAAAAAC**GCG**GATGACTTCCTGG
X0322	I32A_R	CCAGGAAGTCATC**CGC**GTTTTTTTTCACGGCCGG
X0511	R51A_F	CTGAATCCAAAGCGTCCAGCCTG**GCG**GTAGAACTC
X0512	R51A_R	GAGTTCTAC**CGC**CAGGCTGGACGCTTTGGATTCAG
X0551	F55A_F	AGCCTGCGTGTAGAACTC**GCG**AAAACGTACGGGT
X0552	F55A_R	ACCCGTACGTTTT**CGC**GAGTTCTACACGCAGGCT
X0611	T61A_F	AGAACTCTTCAAAACGTACGGGTCT**GCG**CTGGCA
X0612	T61A_R	TGCCAG**CGC**AGACCCGTACGTTTTGAAGAGTTCT
X0615	T61V_F	AGAACTCTTCAAAACGTACGGGTCT**GTG**CTGGCA
X0616	T61V_R	TGCCAG**CAC**AGACCCGTACGTTTTGAAGAGTTCT
X0771	Y77F_F	TCCTGATGAA**TTT**CATTCTTTCGTTCATGGTCGTC
X0772	Y77F_R	GACGACCATGAACGAAAGAATG**AAA**TTCATCAGGA
X0781	H78A_F	GCCACGACGTTCATCCTGATGAATAC**GCG**TCTTT
X0782	H78A_R	AAAGA**CGC**GTATTCATCAGGATGAACGTCGTGGC
X0821	H82A_F	CCTGATGAATACCATTCTTTCGTT**GCG**GGTCGTCT
X0822	H82A_R	AGACGACC**CGC**AACGAAAGAATGGTATTCATCAGG
X1111	T111A_F	ATCAAACAGCGCAAAATTATCTTT**GCG**AATTCTGA
X1112	T111A_R	TCAGAATT**CGC**AAAGATAATTTTGCGCTGTTTGAT
X1121	N112A_F	TATCTTTACG**GCG**TCTGATAAAAACCATGCCGTCAA
X1122	N112A_R	TTGACGGCATGGTTTTTATCAGA**CGC**CGTAAAGATA
X1591	K159A_F	GGTTGTTCTG**GCG**CCGAGTCTGACCGCCATGGATA
X1592	K159A_R	TATCCATGGCGGTCAGACTCGG**CGC**CAGAACAACC
X1831	D183A_F	GTGTTCCTG**GCG**GATAACATTCACAACATCACCGC
X1832	D183A_R	GCGGTGATGTTGTGAATGTTATC**CGC**CAGGAACAC
X1841	D184A_F	GTTCCTGGAT**GCG**AACATTCACAACATCACCGCCG
X1842	D184A_R	CGGCGGTGATGTTGTGAATGTT**CGC**ATCCAGGAAC
X1881	N188A_F	AACATTCAC**GCG**ATCACCGCCGGTAAATCTGTAGGC
X1882	N188A_R	GCCTACAGATTTACCGGCGGTGAT**CGC**GTGAATGTT

Sequence of the synthetic *XMPP* gene for expression in *E. coli*

```
ATGGATTTCAGCCCGATCAACTGCCTGATCTTTGACCTGGATGACACCCTGTATCCGCTG
AAGACGGGTATCGCGCCGGCCGTGAAAAAAAACATCGATGACTTCCTGGTAGAGAAATTC
GGTTTCTCTGAATCCAAAGCGTCCAGCCTGCGTGTAGAACTCTTCAAAACGTACGGGTCT
ACCCTGGCAGGCCTGCGCGCACTGGGCCACGACGTTCATCCTGATGAATACCATTCTTTC
GTTCATGGTCGTCTGCCGTACGGTTCCATCGAACCTAACAACAAACTGCGTAACCTGCTG
AATAAAATCAAACAGCGCAAAATTATCTTTACGAATTCTGATAAAAACCATGCCGTCAAA
GTTCTGAAGAAGCTGGGCCTGGAGGACTGTTTTGAGGAAATGATTTGTTTCGAAACCATG
AACCCGAACCTGTTCGGTAGCACTACCCGTCCGGACGAATACCCGGTTGTTCTGAAACCG
AGTCTGACCGCCATGGATATCTGTATTCGCGTAGCAAATGTTGATCCGCGTCGTACCGTG
TTCCTGGATGATAACATTCACAACATCACCGCCGGTAAATCTGTAGGCCTGCGTACTATC
CTGGTGGGTCGCGCGGAAAAAACTAAAGATGCTGACTACGCAGTTGAAACCGTGACCGAA
ATCGCGACCGCAGTTCCGGAAATCTGGGCGACTGCGACCGCTACCGGCGGTTTCGACGTG
GGTGGAGAACGTATCCGTCGTAGCAAAAGCGAGCTGGAGGGTATGGCTTCTATTGCAGCT
GTTGGCGCGTGA
```

Supplementary Table 4. MS parameter

MS source parameter	Positive ion mode	Negative ion mode
Ion source	AJS ESI	AJS ESI
Gas temperature	300 °C	350 °C
Gas flow	12 L min^{-1}	12 L min^{-1}
Nebulizer	30 psi	40 psi
Sheath gas heater	300 °C	350 °C
Sheath gas flow	11 L min^{-1}	12 L min^{-1}
Capillary	4,000 V	2,500 V

Analyte	Ion mode	Retention time (min)	Precursor Ion (m/z)	Product ion[1] (m/z)	Frag-mentor V)	Collision energy (V)	Qualifier ratio
allantoate	pos.	0.80	177.05	117.0	60	25	
				74.0	60	25	21.9 – 32.6
urate	neg.	1.05	167.03	124.0	91	14	
				69.2	91	20	11.1 – 16.5
xanthosine	pos.	1.50	285.08	153.0	60	5	
				135.9	60	35	15.6 – 23.3
guanine	pos.	2.20	152.05	135.0	118	18	
				110.0	118	22	43.9 – 65.8
guanosine	pos.	5.00	284.09	152.0	90	10	
				135.0	90	45	26.0 – 39.5

[1] the first listed product ion was used for quantification.

2.7 An inosine triphosphate pyrophosphatase safeguards nucleic acids from aberrant purine nucleotides

Henryk Straube[1], Jannis Straube[2], Jannis Rinne[1], Markus Niehaus[1], Claus-Peter Witte[1] and Marco Herde[1]

[1] Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany

[2] Department of Molecular Plant Breeding, Leibniz Universität Hannover, Hannover, 30419, Germany

An inosine triphosphate pyrophosphatase safeguards plant nucleic acids from aberrant purine nucleotides

Henryk Straube[1] (iD), Jannis Straube[2] (iD), Jannis Rinne[1] (iD), Lisa Fischer[1] (iD), Markus Niehaus[1] (iD), Claus-Peter Witte[1] (iD) and Marco Herde[1] (iD)

[1]Department of Molecular Nutrition and Biochemistry of Plants, Leibniz Universität Hannover, Hannover 30419, Germany; [2]Department of Molecular Plant Breeding, Leibniz Universität Hannover, Hannover 30419, Germany

Author for correspondence:
Marco Herde
Email: mherde@pflern.uni-hannover.de

Summary

• In plants, inosine is enzymatically introduced in some tRNAs, but not in other RNAs or DNA. Nonetheless, our data show that RNA and DNA from *Arabidopsis thaliana* contain (deoxy)inosine, probably derived from nonenzymatic adenosine deamination in nucleic acids and usage of (deoxy)inosine triphosphate (dITP and ITP) during nucleic acid synthesis.
• We combined biochemical approaches, LC–MS, as well as RNA-Seq to characterize a plant INOSINE TRIPHOSPHATE PYROPHOSPHATASE (ITPA) from *A. thaliana*, which is conserved in many organisms, and investigated the sources of deaminated purine nucleotides in plants.
• Inosine triphosphate pyrophosphatase dephosphorylates deaminated nucleoside di- and triphosphates to the respective monophosphates. *ITPA* loss-of-function causes inosine di- and triphosphate accumulation *in vivo* and an elevated inosine and deoxyinosine content in RNA and DNA, respectively, as well as salicylic acid (SA) accumulation, early senescence, and upregulation of transcripts associated with immunity and senescence. Cadmium-induced oxidative stress and biochemical inhibition of the INOSINE MONOPHOSPHATE DEHYDROGENASE leads to more IDP and ITP in the wild-type (WT), and this effect is enhanced in *itpa* mutants, suggesting that ITP originates from ATP deamination and IMP phosphorylation.
• Inosine triphosphate pyrophosphatase is part of a molecular protection system in plants, preventing the accumulation of (d)ITP and its usage for nucleic acid synthesis.

Introduction

Spontaneous or enzymatically catalyzed reactions can lead to damaged or aberrant metabolites (Linster *et al.*, 2013; Lerma-Ortiz *et al.*, 2016; Crécy-Lagard *et al.*, 2018). In recent years, several advances have been made in the field of damaged plant metabolites (Hanson *et al.*, 2016), like the discovery of enzymes repairing NAD(P)H hydrates (Niehaus *et al.*, 2014), removing damaged glutathione (Niehaus *et al.*, 2019) and oxidized deoxyguanosine triphosphate (8-oxo dGTP; Dobrzanska *et al.*, 2002; Yoshimura *et al.*, 2007). One of the most prominent metabolite classes that are prone to become damaged are nucleotides (Galperin *et al.*, 2006; Rampazzo *et al.*, 2010; Nagy *et al.*, 2014; Rudd *et al.*, 2016). In nonplant organisms, several enzymes have been identified that repair or remove aberrant and often toxic nucleotides, like DEOXYURIDINE TRIPHOSPHATE PYRO-PHOSPHATASE (DUT1), INOSINE TRIPHOSPHATE PYR-OPHOSPHATASE (ITPA), 8-OXO DEOXYGUANOSINE TRIPHOSPHATE PYROPHOSPHATASE (NUDX1), and MULTICOPY-ASSOCIATED FILAMENTATION (MAF) family proteins (Nagy *et al.*, 2014). However, little is known about such enzymes in plants. While plant homologues of DUT1 and NUDX1 (Yoshimura *et al.*, 2007; Dubois

et al., 2011) have been characterized, a plant ITPA has not been described.

Inosine triphosphate pyrophosphatase (ITPA) enzymes from *Escherichia coli* and *Homo sapiens* remove deaminated purine nucleotide triphosphates. They exclusively hydrolyze the phospho-anhydride bond between the alpha- and beta-phosphate of (deoxy)inosine triphosphate (ITP and dITP) probably derived from the deamination of (deoxy)adenosine triphosphate (ATP and dATP) or from spurious phosphorylation of inosine monophosphate (IMP). Xanthosine triphosphate (XTP), presumably derived from the deamination of guanosine triphosphate (GTP) or spurious phosphorylation of xanthosine monophosphate (XMP), is also a substrate (Lin *et al.*, 2001; Davies *et al.*, 2012). Inosine triphosphate pyrophosphatase generates the corresponding nucleotide monophosphate and pyrophosphate. Whereas (d)ITP and XTP are aberrant nucleotides, IMP and XMP are canonical intermediates of guanosine monophosphate (GMP) synthesis in most organisms and in plants XMP is additionally a starting point of purine nucleotide catabolism (Heinemann *et al.*, 2021).

Deamination of cytosine to uracil is well known to occur in nucleic acids (Lindahl & Nyberg, 1974; Frederico *et al.*, 1990), but purine bases can also be deaminated. *In vitro*, the deamination of adenine to hypoxanthine in DNA occurs at 2% to 3% of

the rate observed for cytosine to uridine conversion (Lindahl, 1993). The deamination of free nucleotides is less well documented. Purine deamination was shown to occur not only in the nucleic acids but also in the pool of free nucleotides under heat stress (Karren & Lindahl, 1980). In general abiotic stresses, like salinity, water stress, extreme temperature, wounding, UV radiation, or exposure to heavy metals (Valderrama *et al.*, 2007; Corpas *et al.*, 2011) may influence the rate of nucleotide deamination. These stresses generate reactive nitrogen or oxygen species (RNS, ROS; Valderrama *et al.*, 2007; Corpas *et al.*, 2011), which are mainly discussed in the context of protein modification and lipid peroxidation but are also likely damaging other metabolites although this has rarely been shown so far (Demidchik, 2015).

Inosine triphosphate and dITP are substrates of RNA and DNA polymerases, respectively. Random inosine or deoxyinosine in nucleic acids can thus not only be derived from the deamination of adenosine or deoxyadenosine in the polymer but may also be incorporated during polymerization. When DNA, containing deaminated sites, is replicated, mutations may be introduced. Deoxycytidine can be incorporated opposite of deoxyinosine and often thymidine is placed opposite of deoxyxanthosine (Kamiya, 2003). This misincorporation is not directly detrimental for *E. coli* (Budke & Kuzminov, 2006) because it is rapidly repaired. However, upon stimulation of deamination, the constant deployment of the repair machinery leads to strand breaks and chromosomal rearrangements (Budke & Kuzminov, 2010). By contrast, deoxyinosine in human DNA causes mutations and genome instability even when deamination is not stimulated (Yoneshima *et al.*, 2016; DeVito *et al.*, 2017). During RNA synthesis in nonplant organisms, incorporation of ITP has been shown to significantly reduce transcriptional velocity. Inosine in RNA alters RNA structure and stability and leads to mistranslation (Thomas *et al.*, 1998; Ji *et al.*, 2017) or stalling of translation (Schroader *et al.*, 2022). Apart from effects on nucleic acids, ITP and XTP can interfere with various ATP- and GTP-dependent cellular processes (Osheroff *et al.*, 1983; Carnelli *et al.*, 1992; Klinker & Seifert, 1997; Muraoka *et al.*, 1999; Weber & Senior, 2001).

Inosine triphosphate has been detected *in vivo* in only a few occasions. Human erythrocytes accumulated radioactive ITP after incubation with [^{14}C] adenine, [^{14}C] guanine, or [^{14}C] hypoxanthine, indicating that a biosynthetic route to deaminated purine nucleotide triphosphates might exist (Fraser *et al.*, 1975). Interestingly, a loss of ITPA function in erythrocytes of mice leads to a substantial accumulation of ITP of up to 10% to 25% of the ATP pool (Behmanesh *et al.*, 2009). However, erythrocytes are a special cell type, which lack nuclear DNA and seem to have a particular nucleotide metabolism. It is technically challenging to detect ITP and XTP at all. In human cancer cells, both were detected with a highly sensitive method in the fmol mg^{-1} protein range. Here, ITP was around 5–10 times more abundant than XTP (Jiang *et al.*, 2018). To date, the corresponding deoxyribonucleotide triphosphates, dITP and deoxyxanthosine triphosphate (dXTP), have never been detected in any organism.

Although these metabolites are low abundant, they can be incorporated into nucleic acids. The frequency of inosine in RNA is about one in 10^5 nucleosides in wild-type (WT) *E. coli*,

while deoxyinosine in DNA is 10-fold less abundant (Pang *et al.*, 2012). In mutants of rdgB, the bacterial ITPA homolog, the inosine content of RNA increased *c.* 10-fold and the deoxyinosine of DNA rose about fivefold compared with the WT (Pang *et al.*, 2012). In *ITPA* knockout mice, the inosine-to-adenosine ratio was about one per 100 nucleosides in RNA (Behmanesh *et al.*, 2009). In mammals, ITPA deficiency causes drastic phenotypes. In humans, it leads to a lethal Martsolf-like syndrome with neurological and developmental abnormalities (Handley *et al.*, 2019). Mice without a functional ITPA had RNA with an elevated inosine content and died shortly after birth (Behmanesh *et al.*, 2009). However, bacteria lacking ITPA are viable (Burgis *et al.*, 2003). Recently, a homolog of ITPA in *Cassava brown streak virus*, a plant-infecting RNA virus, was demonstrated to be an important factor during viral infection (Tomlinson *et al.*, 2019), interacting with the viral RNA-dependent RNA polymerase (Valli *et al.*, 2022).

In this study, we identified ITPA of *A. thaliana* and biochemically characterized the enzyme, showing that it catalyzes the dephosphorylation of deaminated purine nucleotides. We analyzed mutants of *ITPA* demonstrating that the enzyme safeguards nucleic acids from incorporation of (d)ITP. We were able to detect ITP and IDP accumulation in the mutants and investigated the metabolic source of ITP and IDP. In contrast to *itpa* mutants in mammals, the Arabidopsis mutants have only minor phenotypic alterations. They show slightly earlier senescence and have elevated concentrations of the defense phytohormone salicylic acid (SA), triggering a transcriptional response of genes associated with SA, systemic acquired resistance (SAR) and aging.

Materials and Methods

Plant cultivation

Arabidopsis thaliana (L.) Heynh and *Nicotiana benthamiana* plants were cultivated as described recently (Witte *et al.*, 2005; Straube *et al.*, 2021).

A segregating T-DNA insertion line from the SALK collection (*itpa*-1, SALK_052023; Alonso *et al.*, 2003) was acquired from the Nottingham Arabidopsis Stock Centre. A homozygous knockout line and a corresponding WT were selected from the segregating population (see Methods S1).

A second knockout line was obtained via genome editing (*itpa*-2) according to Rinne *et al.* (2021; see Methods S1). A line overexpressing ITPA (*itpa-2::C*) was generated (see Methods S1).

For root-length and germination assays, plants were grown in Petri dishes on previously described medium (Niehaus *et al.*, 2020; Straube *et al.*, 2021).

Plants for the cadmium sulfate feeding experiment were grown hydroponically according to Batista-Silva *et al.* (2019), with small modifications. Plants were grown in viscous liquid medium (Niehaus *et al.*, 2020; Straube *et al.*, 2021) containing 0.4% agar under long-day conditions. Cadmium treatment was performed for 24 h by adding a 500 mM cadmium sulfate solution to a final concentration of 10 µM to the medium, respectively. This concentration was chosen according to Di Sanità & Gabbrielli (1999).

DNA extraction, RNA extraction, genotyping, qRT-PCR, and cloning

DNA for genotyping was isolated according to the BOMB protocol using the TNA extraction from plants with the TNES/GITC lysis protocol (Oberacker *et al.*, 2019). Magnetic beads were silica-coated beads, synthesized using the BOMB protocol for coating ferrite magnetic nanoparticles with silica oxide (Oberacker *et al.*, 2019).

Total RNA was isolated using the Plant NucleoSpin II kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's specifications.

For information on genotyping, qRT-PCR, and cloning, see Methods S1. Primers used in this study are provided in Table S1.

Protein purification, quantification, and enzymatic assay

C-terminal strep-tagged ITPA, AMPK3, and AMPK4 (Chen *et al.*, 2018) were transiently overexpressed in *N. benthamiana* and affinity purified after 5 d as described previously (Werner *et al.*, 2008). The protein concentration was determined by an in-gel quantification assay using a Li-Cor Odyssey FC with Image Studio software. SDS gel electrophoresis and staining were performed as described previously (Witte *et al.*, 2005).

Enzymatic assays for ITPA were performed using the EnzChek™ Pyrophosphate Assay Kit (ThermoFisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions, with the exception of reducing the volume of all solution to one-fifth of the instructions using High Precision Cells (Hellma Analytics, Müllheim, Germany).

AMPK3 and 4 activity was determined using a coupled enzyme assay with pyruvate kinase and lactate dehydrogenase (LDH, P0294; Sigma-Aldrich, St Louis, MO, USA), measuring the production of NAD^+ that results in absorption decrease at 340 nm. For details, see Methods S1.

Subcellular localization

A construct encoding ITPA fused to C-terminal YELLOW FLUORESCENT PROTEIN (YFP) was transiently coexpressed in *N. benthamiana* with a construct encoding cytosolic β-ureidopropionase C-terminally fused to CYAN FLUORESCENT PROTEIN (CFP). Whole leaves were immersed in octadecafluordecahydronaphthalin (Sigma-Aldrich; Littlejohn *et al.*, 2010). The samples were analyzed using a Leica True Confocal Scanner SP8 microscope equipped with an HC PL APO CS2 40 × 1.10 water immersion objective (Leica, Wetzlar, Germany) as described in Dahncke & Witte (2013). The acquired images were processed using Leica Application Suite Advanced Fluorescence software (Leica Microsystems, Wetzlar, Germany). The immunoblot to analyze the construct was developed using a GREEN FLUORESCENT PROTEIN (GFP)-specific antibody. Co-localization was analyzed using the JACoP-plugin in IMAGEJ (Bolte & Cordelieres, 2006).

Metabolite analysis

The analysis of nucleotide concentration was performed as recently described with slight modifications (Straube *et al.*, 2021; Straube & Herde, 2022). For details, see Methods S1.

Salicylic acid (SA) was extracted and measured according to Cao *et al.* (2014) and Yang *et al.* (2017).

Quantitative determination of (deoxy)inosine per (deoxy) adenosine level in total RNA and DNA

DNA was extracted using a modified CTAB- (hexadecyltrimethylammonium bromide) based protocol (see Methods S1). Total RNA was isolated using the Plant NucleoSpin II kit (Macherey-Nagel) according to the manufacturer's specifications.

Nucleic acids were digested using the Nucleoside Digestion Mix (New England Biolabs, Ipswich, MA, USA) with the following modifications: 5 µl Nucleoside Digestion Mix Reaction Buffer (10×) and 0.5 µl Nucleoside Digestion Mix were used, adding RNA or DNA and water depending on the respective concentration of the RNA or DNA up to a total volume of 50 µl. The mixture was incubated for 4 h at 37°C. A total of 1 µg RNA or 3 µg DNA was digested.

Digested RNA and DNA was measured by HPLC–MS as recently published (Traube *et al.*, 2019), with minor modifications (Methods S1).

Analysis of chlorophyll content

Measurement of chlorophyll content was performed as described previously (Harris & Baulcombe, 2015).

Cell death staining and quantification

Trypan blue staining was performed as published recently (Fernández-Bautista *et al.*, 2016). To quantify the leaf area affected by cell death, RGB pictures of stained leaves were taken and stained cell areas were quantified using IMAGEJ.

Sequence analysis and phylogenetic analysis

Inosine triphosphate pyrophosphatases sequences were identified using the PHYTOZOME v.12.1 webserver (https://phytozome.jgi.doe.gov/pz/portal.html) and the National Center for Biotechnology (NCBI) for all nonplant organisms employing BLASTP searches with the *Arabidopsis* amino acid sequence. The alignment and phylogenetic tree were constructed as described recently (Heinemann *et al.*, 2021). The alignment used to construct the phylogenetic tree is shown in Fig. S1. For details, see Methods S1.

RNA-seq analysis

Total RNA from 35-d-old rosette leaves from *A. thaliana* Col-0 ('wild type') and *itpa*-1 plants was sent for Illumina mRNA

sequencing to Novogene (Cambridge, UK). For details, see Methods S1.

Ribavirin treatment

Arabidopsis thaliana and *N. benthamiana* plants were either mock-treated or with buffer containing ribavirin, an inhibitor of the IMP DEHYDROGENASE. For details, see Methods S1.

Statistical analysis

Statistical analysis was performed as stated in Heinemann *et al.* (2021). The number of replicates, test values, and multiplicity-adjusted *P*-values are reported in Table S4.

Results

Comprehensive phylogenetic analysis of inosine triphosphate pyrophosphatase in plants

We used the amino acid sequence of human ITPA to search with BLASTP for homologs in plants (Fig. S2). For *A. thaliana*, the protein encoded at the gene locus At4g13720 was the only candidate with a high sequence similarity to the human ITPA. It has been suggested previously that this enzyme may represent a plant ITPA (Lin *et al.*, 2001; James *et al.*, 2021). A phylogenetic analysis revealed that ITPA is present in all kingdoms of life including plants and algae (Figs 1, S1).

Several amino acid residues are highly conserved in ITPAs from different species. Many of these have been shown to be relevant for substrate binding, catalytic activity, and dimerization (Figs S1, S2; Savchenko *et al.*, 2007; Stenmark *et al.*, 2007; Gall *et al.*, 2013). Nonetheless, there are amino acid residues that are only conserved in plant ITPAs (Figs S1, S2). Among these, K/R 183 and K/R 189 can be acetylated in Arabidopsis (Liu *et al.*, 2018), but the function of these posttranslational modifications is unclear.

ITPA has atypical characteristics for an ITP pyrophosphatase

Inosine triphosphate pyrophosphatases are described in other organisms to hydrolyze the phosphoanhydride bond between the alpha- and beta-phosphate groups of inosine triphosphate resulting in IMP and pyrophosphate (Fig. 2a).

To assess the activity of the plant enzyme, a C-terminal Strep-tagged variant of ITPA was transiently expressed in *N. benthamiana* and affinity purified (Fig. 2b). Interestingly, two forms of the protein were purified represented by two bands in a Coomassie-stained SDS-PAGE, possibly indicating a posttranslational modification. The identity of the purified protein was confirmed by immunoblot with an antibody against the Strep-tag. The specific activity of the purified enzyme was determined with several substrates at concentrations of 200 μM, including (deoxy)ribonucleotide di-, and triphosphates, *para*-nitrophenyl-phosphate and nicotinamide adenine dinucleotide (Fig. 2c; Table S5). An enzymatic activity could only be observed with dITP, ITP, inosine
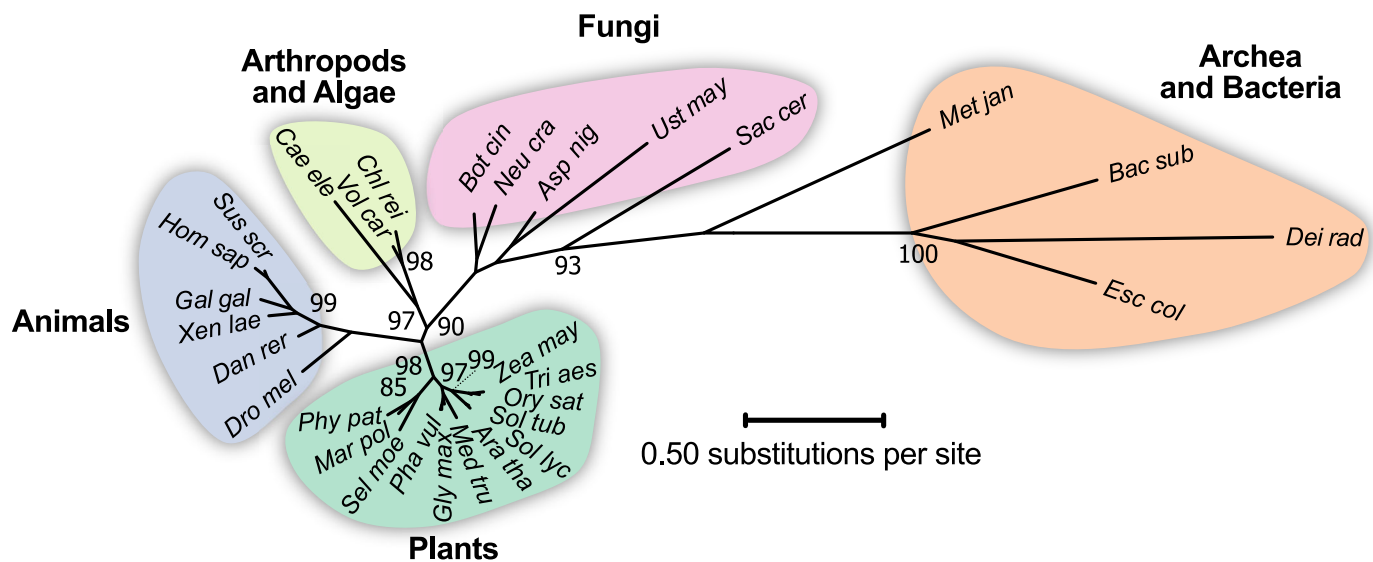
Fig. 1 Phylogenetic analysis of ITPA in several organisms. An unrooted maximum likelihood tree was built using MEGAX, comprising model species from a wide taxonomic range. The tree with the best log score is shown with branch length indicating the number of site substitutions (scale bar). Numbers at branches indicate the percentage of trees in which associated taxa clustered together in the bootstrap analysis (1000 bootstraps). Values below 80 are not shown. Species names are abbreviated as follows: Asp nig, *Aspergillus niger*; Ara tha, *Arabidopsis thaliana*; Bac sub, *Bacillus subtilis*; Bot cin, *Botrytis cinerea*; Cae ele, *Caenorhabditis elegans*; Chl rei, *Chlamydomonas reinhardtii*; Dan rer, *Danio rerio*; Dei rad, *Deinococcus radiodurans*; Dro mel, *Drosophila melanogaster*; Esc col, *Escherichia coli*; Gal gal, *Gallus gallus*; Gly max, *Glycine max*; Hom sap, *Homo sapiens*; Mar pol, *Marchantia polymorpha*; Met jan, *Methanocaldococcus jannaschii*; Med tru, *Medicago truncatula*; Neu cra, *Neurospora crassa*; Ory sat, *Oryza sativa*; Phy pat, *Physcomitrium patens*; Pha vul, *Phaseolus vulgaris*; Sac cer, *Saccharomyces cerevisiae*; Sel moe, *Selaginella moellendorffii*; Sol lyc, *Solanum lycopersicum*; Sol tub, *Solanum tuberosum*; Sus scr, *Sus scrofa*; Tri aes, *Triticum aestivum*; Ust may, *Ustilago maydis*; Vol car, *Volvox carteri*; Xen lae, *Xenopus laevis*; Zea may, *Zea mays*.

Fig. 2 Characterization of the enzymatic properties of inosine triphosphate pyrophosphatases (ITPA). (a) Reaction scheme of ITPA with inosine triphosphate (ITP). (b) Affinity purified, C-terminal Strep-tagged ITPA on a Coomassie-stained gel (left) and detection of Arabidopsis ITPA by immunoblotting using an antibody against the Strep-tag (right). Numbers indicate the respective size in kDa (kDa). Crude, centrifuged extract; unbound, centrifuged extract after incubation with Strep-Tactin matrix; 5th wash, supernatant after the fifth wash of the affinity matrix; elution, eluate of the affinity matrix. (c) Specific activity of ITPA with different substrate at a concentration of 200 μM. nd, not detected. (d) Determination of kinetic constants for the ITPA-catalyzed reaction of dITP to dIMP. Kinetic data were fitted using the Michaelis–Menten equation (orange) or by linear regression for the Hanes plot (S/V, purple). Error bars are standard deviation (SD), $n = 3$ reactions. S, substrate concentration; V, enzymatic activity. (e) The same as in (d), but using ITP as substrate. All P-values are provided in Table S4.
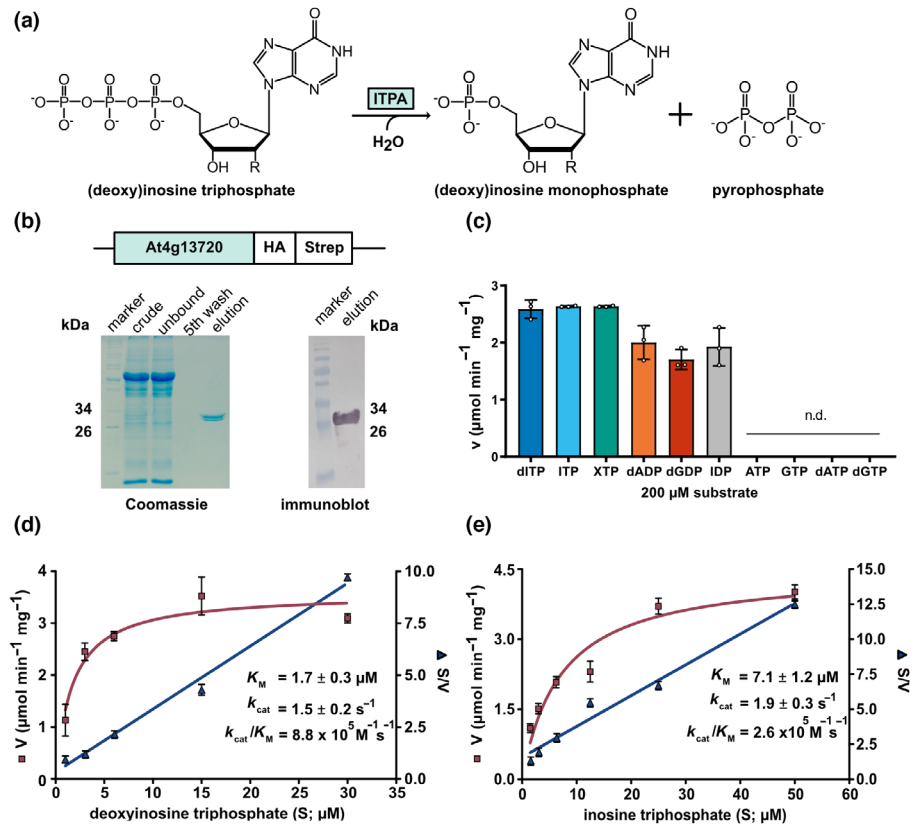
diphosphate (IDP), and XTP, as well as with deoxyadenosine diphosphate (dADP) and deoxyguanosine diphosphate (dGDP). In contrast to the ITPA from *H. sapiens*, the ITPA of *A. thaliana* efficiently dephosphorylates IDP (Davies *et al.*, 2012). For dITP and ITP, the kinetic constants were determined (Fig. 2d,e). With ITP as substrate, a K_M of 7.1 ± 1.2 μM and a turnover number of 1.9 ± 0.25 s^{-1} were measured. With dITP, the K_M was about fourfold lower at 1.7 ± 0.3 μM while the turnover number was similar at 1.5 ± 0.17 s^{-1}. Thus, Arabidopsis ITPA has catalytic efficiencies of 2.6×10^5 M^{-1} s^{-1} for ITP and 8.8×10^5 M^{-1} s^{-1} for dITP (Table 1).

ITPA mutants show an altered phenotype during senescence

To investigate the role of ITPA *in vivo*, two mutant lines and an overexpression line were created (Fig. S4). One T-DNA line with

an insertion in an exon of *ITPA* was available from the SALK collection (SALK_053023; Fig. S4a).

We identified plants with a homozygous insertion of the T-DNA and observed that they lacked an intact transcript (Fig. S4b). This mutant was named *itpa-1*. A second line (*itpa-2*) was generated using CRISPR, resulting in a plant with two homozygous single base pair insertions causing frameshifts, one in exon 2 and one in exon 5. From the segregating population harboring the editing construct, *itpa-2* plants without the transgene were selected and used for further characterization. Both mutant lines had no obvious developmental phenotypes (Fig. S3a–d) but showed an early onset of senescence at 40 d after germination (Fig. 3a) and a slightly increased senescence at an age of 8 wk with *c.* 50% less chlorophyll in the oldest leaves, as well as an increased percentage of dead cells per leaf in *itpa* plants when compared to the WT (Fig. 3b–d). The plant fresh weight, the number of seeds per silique, and the length of the

Table 1 Comparison of kinetic parameters of plant, fungi, and mammal ITPA enzymes.

	ITPA (*Arabidopsis thaliana*; this study)			HAM1 (*Saccharomyces cerevisiae*; Davies *et al.*, 2012)			ITPA (*Homo sapiens*; Lin *et al.*, 2001)		
	K_M	k_{cat}	k_{cat}/K_M	K_M	k_{cat}	k_{cat}/K_M	K_M	k_{cat}	k_{cat}/K_M
Substrate	μM	s^{-1}	M^{-1} s^{-1}	μM	s^{-1}	M^{-1} s^{-1}	μM	s^{-1}	M^{-1} s^{-1}
dITP	1.7 ± 0.3	1.5 ± 0.17	8.8×10^5	3.06 ± 0.6	1.3	4.21×10^5	450 ± 100	360	1.16×10^6
ITP	7.1 ± 1.2	1.9 ± 0.25	2.6×10^5	2.38 ± 0.4	1.0	4.17×10^5	510 ± 100	580	1.14×10^6

±, standard deviation (SD). dITP, (deoxy)inosine triphosphate; ITP, inosine triphosphate; ITPA, inosine triphosphate pyrophosphatase.
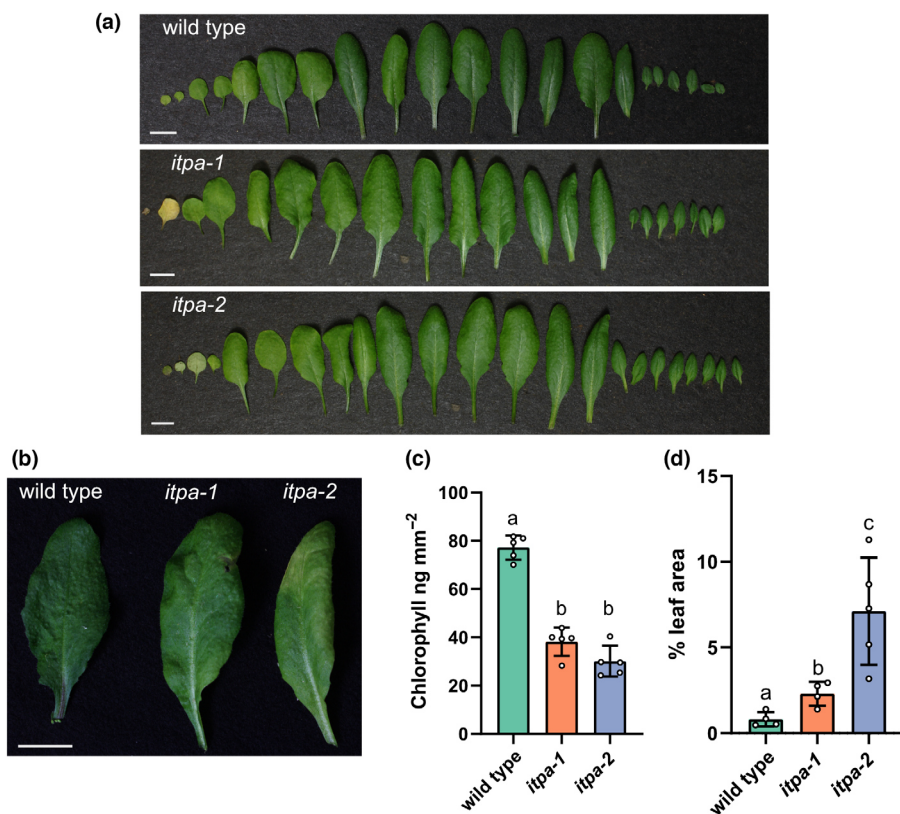
Fig. 3 Phenotypic comparison of wild-type (WT), *itpa-1*, and *itpa-2* plants. (a) Arabidopsis rosette leaves of 40-d-old WT, *itpa-1*, and *itpa-2* plants grown under long-day conditions. Bar, 10 mm. (b) Representative picture of the oldest leaves from 56-d-old plants of different genotypes used in (c). (c) Chlorophyll content of the respective oldest leaves from 56-d-old plants of different genotypes, $n = 5$. Every replicate represents a different plant of the same genotype. (d) Leaves of 56-d-old WT, *itpa-1*, and *itpa-2* plants were stained by Trypan blue, and dark pixels per leaf area were measured using IMAGEJ. Each open circle represents the result of measuring an independent biological replicate. Error bars are standard deviation (SD), $n = 4$–5. Statistical analysis was performed by a two-sided Tukey's pairwise comparison using the sandwich variance estimator. Different letters indicate $P < 0.05$. All P-values are provided in Table S4.

siliques were not different between the mutants and the WT (Fig. S3b–d).

ITPA is localized in the cytosol, nucleus, and putatively in plastids

To determine the subcellular localization of ITPA, cells of *N. benthamiana* co-expressing a C-terminally YFP-tagged variant of ITPA and cytosolic β-ureidopropionase C-terminally fused to CFP. Inosine triphosphate pyrophosphatase (ITPA) was clearly localized in the cytosol and nucleus, co-localizing with the cytosolic marker as the signals correlate with an *R*-value of 0.98; additionally, a good fit in the cross-correlation analysis was observed (Fig. 4a,b).

Both fusion proteins are stable as shown by an immunoblot (Fig. 4c). Interestingly, while the plastids are completely dark in the channel corresponding to the CFP signal of the cytosolic marker, there is signal in the YFP channel at these locations, illustrated by the arrows. Furthermore, the YFP signal and the plastidic autofluorescence overlap, indicated by a cyan coloration in the overlay of micrographs (Fig. 4a). Although the signals do not coincide completely, the data indicate that ITPA might also be localized in the plastids. This is in line with the predictions of several bioinformatic analysis tools summarized in the SUBA4 database, predicting nuclear, cytosolic, and plastidic localization for ITPA (Hooper *et al.*, 2017). Additionally, ITPA was found in a proteomic analysis of the plastid proteome (Kong *et al.*, 2011). A localization in mitochondria, also predicted by SUBA4, was not confirmed in our experiments.

ITPA is essential for the catabolism of ITP and IDP *in vivo*

In 35-d-old plants, IDP and ITP were only detectable in *itpa-1* and *itpa-2* background, but not in the WT (Fig. 5a).

Inosine triphosphate and IDP were reliably measured because they were baseline separated from the isobaric ATP and ADP isotopes (Fig. S5). The *itpa-1* and *itpa-2* lines contained 1.7 and 1.8 pmol g^{-1} IDP and 1.8 and 1.3 pmol g^{-1} ITP, respectively. This accounts for *c.* 30 molecules of ITP per 10^6 molecules of adenosine triphosphate (ATP) and between 300 and 400 molecules of IDP per 10^6 molecules of adenosine diphosphate (ADP, Table 2).

The ITP concentration per cell was *c.* 9.6 nM and that of IDP *c.* 11.2 nM estimated according to Straube *et al.* (2021). The IMP concentration was not altered significantly in the knockout lines (Fig. 5a), suggesting that ITP and IDP dephosphorylation is not a main contributor to the IMP pool. This is not surprising because IMP is a comparatively abundant intermediate of purine nucleotide biosynthesis and purine nucleotide catabolism, compared with the low abundant IDP and ITP (Baccolini & Witte, 2019; Witte & Herde, 2020). Although ITPA is active with dADP and dGDP *in vitro*, mutation of *ITPA* had no effect on the dADP and dGDP concentrations *in vivo* (Fig. S6a). The concentrations of other purine and pyrimidine nucleotides were also not altered in the mutant lines (Figs 5b, S6a–c).

What is the metabolic source of ITP and IDP?

Hitherto, the metabolic source of ITP and IDP is unknown. One possibility is that these metabolites arise from promiscuous
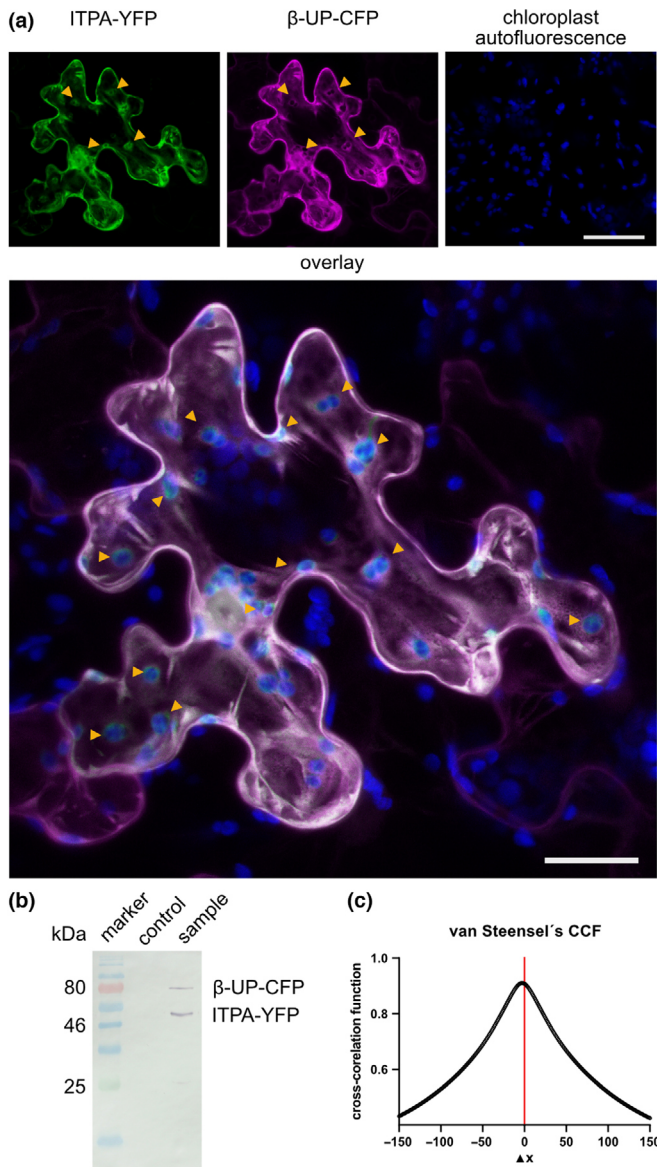
Fig. 4 Subcellular localization of inosine triphosphate pyrophosphatases (ITPA). (a) Confocal fluorescence microscopy images of leaf epidermal cells of *N. benthamiana* transiently co-expressing Arabidopsis ITPA C-terminally fused to YFP (ITPA-YFP) and cytosolic β-ureidopropionase C-terminally fused to CFP (β-UP-CFP). Left panel, YFP fluorescence; middle panel, CFP fluorescence; right panel, chloroplast autofluorescence. Bar, 50 μm. The large picture is an overlay of all three signals. Bar, 25 μm. Orange arrows indicate YFP signal not overlapping with the CFP signal. Analysis of the Pearson's correlation coefficient gave an *R*-value of 0.98 for the YFP and CFP signals. The images are representative for at least 10 micrographs of different cells. (b) Van Steensel's cross-correlation analysis for the YFP and CFP signals. (c) Analyzing the stability of the ITPA-YFP and β-UP-CFP by an immunoblot developed with a GFP-specific antibody. The control was generated from infiltrated leaves with *Rhizobium radiobacter* carrying the RNA silencing suppressor P19.

enzyme activity (Hanson *et al.*, 2016) of AMP kinases, which in addition to AMP may also be able to phosphorylate IMP. Consistent with this hypothesis, it has been shown in *E. coli* and *Saccharomyces cerevisiae* that an increase in the IMP content is concomitant with an accumulation of (deoxy)inosine in RNA

and DNA (Pang *et al.*, 2012). In plants, more IMP could also lead to more ITP and IDP, which would indicate that the phosphorylation of IMP by kinases is a source of ITP and IDP.

A biochemical approach was used to investigate the question of possible aberrant IMP phosphorylation. We determined the enzymatic activities of the two cytosolic AMP kinases, AMPK3 (At5g50370) and AMPK4 (At5G63400) of *A. thaliana* with either 1 mM AMP or IMP as substrates (Lange *et al.*, 2008; Chen *et al.*, 2018; Fig. 6a,b). Although IMP is an *in vitro* substrate for both enzymes, the activity with IMP was only 3.9% of that with AMP for AMPK3 (Fig. 6a) and 5.2% for IMP vs AMP for AMPK4 (Fig. 6b).

In vivo the relative activity with IMP will be lower because the cellular IMP pool is several orders of magnitude smaller than that of AMP. Nonetheless, the data show that AMPKs can phosphorylate IMP in principle resulting in the formation of ITP and IDP.

To further test the hypothesis of spurious IMP phosphorylation as a source for IDP and ITP, we fed ribavirin to *A. thaliana* WT, *itpa-2*, and *itpa-2::C* plants (Fig. 7), an inhibitor of IMP DEHYDROGENASE (IMPDH, Keya *et al.*, 2003). This enzyme represents a major sink of IMP in plant metabolism (Witte & Herde, 2020), and its inhibition likely increases the IMP concentration *in vivo* and thus the availability of substrate for spurious phosphorylation by AMPKs (Fig. 7c). Ribavirin treatment for 24 h caused a significant increase in the IMP content in ribavirin-treated *A. thaliana* seedlings compared with the mock-treated samples and a decreased concentration of downstream metabolites like XMP (Fig. 7). We noticed a higher abundance of IMP in ribavirin-treated *itpa-2::C* plants and suspect that this is a result of a higher abundance of ITPA, as ITPA from other organisms is described dephosphorylate ribavirin triphosphate (RTP) to its respective monophosphate (Vidal *et al.*, 2022). In the overexpression line, this leads to an increased concentration of ribavirin monophosphate, which is the inhibitory metabolite of IMPDH. Instead of RTP being incorporated into RNA, it is dephosphorylated, leading to an even stronger block of IMPDH and thus a greater accumulation of IMP.

Concomitant with higher IMP concentrations, the abundance of IDP and ITP was increased in all genotypes treated with ribavirin (Fig. 7b), suggesting that spurious phosphorylation of IMP cannot be excluded as a source of IDP and ITP despite the comparably low activity of AMP kinases with IMP as a substrate. The concentration of IDP and ITP was significantly decreased in *itpa-2::C* plants compared with the WT and *itpa-2* plants demonstrating a gain-of-function concerning IDP/ITP dephosphorylation caused by ectopic expression of *ITPA*.

We used the ribavirin treatment as a tool to extend the concept of IMP phosphorylation to other plant species and developmental stages. Thus, we infiltrated *N. benthamiana* plants with ribavirin and the mock control (Fig. S7), similarly to Arabidopsis, the treatment with ribavirin led to increased concentrations of IMP, whereas XMP was not detected (Fig. S7). However, in *N. benthamiana*, the treatment did not result in increased IDP and ITP concentrations, suggesting that the contribution of spurious IMP phosphorylation to IDP and ITP pools varies between
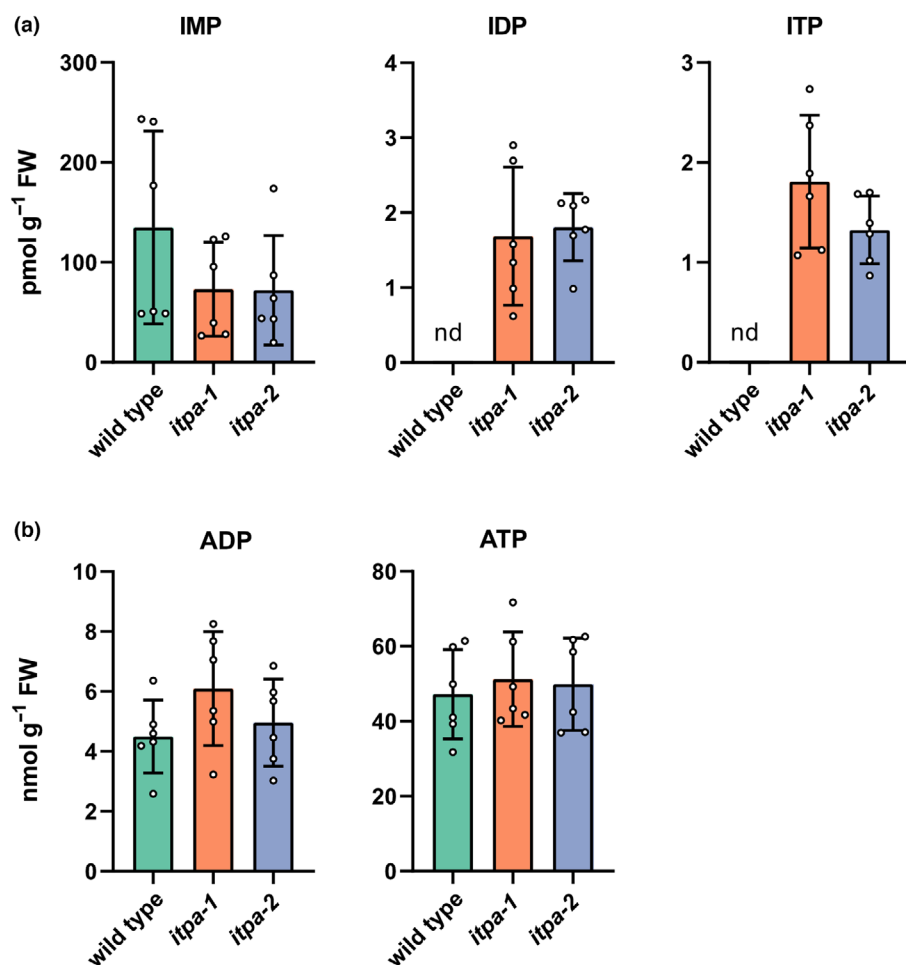
Fig. 5 Loss-of-function of *itpa* causes an accumulation of inosine diphosphate (IDP) and inosine triphosphate (ITP), while adenylate concentrations remain unchanged. (a) Concentrations of inosylates in rosette leaves of 35-d-old Arabidopsis wild-type (WT), *itpa-1*, and *itpa-2* plants. (b) As in (a) but for adenosine diphosphate (ADP) and adenosine triphosphate (ATP). Each open circle represents the result of measuring an independent biological replicate. Error bars are standard deviation (SD), $n = 6$. Every replicate represents a different plant. Statistical analysis was performed by a two-sided Tukey's pairwise comparison using the sandwich variance estimator. Different letters indicate $P < 0.05$. All *P*-values are provided in Table S4. nd, not detected.

Table 2 Ratios of ITP and IDP per million molecules of ATP and ADP, respectively.

	itpa-1	*itpa-2*
ITP/1×10^6 ATP	36.6 ± 12	28.7 ± 11
IDP/1×10^6 ADP	300.8 ± 162	407.5 ± 178

$\pm$, standard deviation (SD). ADP, adenosine diphosphate; ATP, adenosine triphosphate; IDP, inosine diphosphate; ITP, inosine triphosphate.

plant species (Arabidopsis and Nicotiana) or developmental stages (seedlings and fully developed leaves).

Aberrant metabolites can also be derived from nonenzymatic processes – e.g. from contact with ROS (Hanson *et al.*, 2016; Lerma-Ortiz *et al.*, 2016). Heavy metals, such as cadmium, are known to cause ROS production in plants and comprise important anthropogenic pollutants (Di Sanità & Gabbrielli, 1999; Cho & Seo, 2005; Lin *et al.*, 2007). It has been well established that cadmium-induced oxidative stress not only damages macromolecules such as the DNA directly but also causes an increased rate of lipid peroxidation (Lin *et al.*, 2007). To test whether this oxidative stress also has an impact on nucleotides, we grew WT and *itpa-1* plants hydroponically for 14 d and applied cadmium sulfate (0 and 10 μM). Without Cd, WT plants only

accumulated small amounts of ITP, but with 10 μM Cd the ITP content rose significantly (Fig. 8). Inosine diphosphate was not detectable in the WT. As expected, the mutant generally contained more IDP and ITP than the WT and the application of Cd resulted in a further increase in these pools. The concentrations of ITP and IDP measured in this experiment in 14-d-old plants grown in a hydroponic system were comparable to those in 35-d-old plants grown on soil (Figs 5a, 8). These data show that ITP and IDP can be formed under stress conditions that involve ROS probably by the deamination of ATP and ADP. Overall, our data suggest that spurious IMP phosphorylation and ATP deamination caused by oxidative stress both can contribute to IDP and ITP pools but the impact of both processes depends on environmental conditions and putatively on the developmental context.

ITPA prevents (d)ITP incorporation into DNA and RNA

Measuring deoxyinosine (dI) in digested DNA of 7-d-old plants by liquid chromatography-mass spectrometry (LC–MS) revealed that both mutant lines accumulate dI in the DNA (Fig. 9a). Wild-type (WT) DNA contained 290 ± 96 dI per million molecules deoxyadenosine (dA). This amount was more than doubled in the mutants (Table S6). In DNA from older plants (56 d), dI
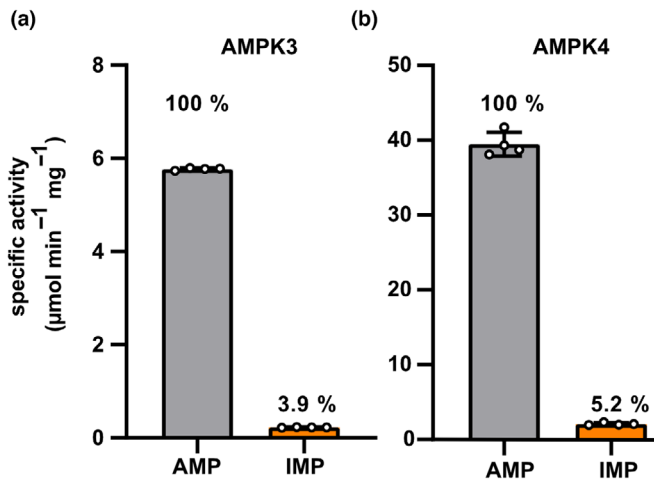
Fig. 6 Enzymatic activity of AMPK 3 and 4 with adenosine monophosphate (AMP) and inosine monophosphate (IMP). Arabidopsis enzymes were transiently expressed in *Nicotiana benthamiana* as C-terminal Strep-tagged variants and affinity purified. (a) Specific activity of AMPK3 with 1 mM AMP or IMP in the presence of 1 mM ATP. Linear rates were observed in all cases. Error bars are standard deviation (SD). Each open circle represents the result of measuring a replicate, each replicate represents an independent enzyme assay from the same enzyme preparation ($n = 4$). (b) Analogous data for AMPK4.

could not be detected in the WT and accumulated in DNA of the mutants to approximately the same level as in seedlings. Therefore, a marked increase in dI accumulation in DNA with age could not be observed. Interestingly, the higher dI concentration in DNA of *itpa* plants demonstrates that a dITP pool likely exists in plants. Deoxyxanthosine was not detected in DNA. To test whether the concentration of dI is really dependent on *ITPA* expression, we analyzed the concentration of dI in DNA of 7-d-old WT, *itpa-2*, and *itpa-2::C* plants grown hydroponically (Fig. 9b).

Consistently, itpa-2 plants accumulated dI and the overexpression of ITPA (itpa-2::C) led to a concentration of dI comparable to what was measured in WT plants.

In digested total RNA of 35-d-old plants, 405 ± 117 molecules of inosine per million molecules of adenosine were detected (Fig. 9c). The *itpa* lines contained about twofold more inosine in the RNA (Table S6), showing that also the RNA is protected by ITPA. In a pilot experiment, the RNA of younger, 21-d-old plants was analyzed and contained only 200 molecules of inosine per million molecules of adenosine (Fig. S8), which is about half the amount found in 35-d-old plants. A similar tendency was observed for the *itpa-1* line. In contrast to deoxyinosine in DNA, it appears that the inosine content of RNA increases with age. Interestingly, inosine-containing RNA can be specifically recognized and cleaved by an endonuclease in Arabidopsis, which may limit the amount of inosine that can accumulate in RNA (Endo *et al.*, 2021). Xanthosine was not detected in RNA. In summary, the data demonstrate that ITPA is an enzyme required for the removal of dITP and ITP to protect DNA and RNA from random incorporation of (deoxy)inosine (Fig. 10).

Loss of ITPA leads to an upregulation of transcripts involved in biotic stress

We further investigated whether the loss of *ITPA* leading to the accumulation of (deoxy)inosine in nucleic acids and inosylates in nucleotide pools has an influence on the transcriptome. An RNA-seq analysis was performed with RNA isolated from rosette leaves of 35-d-old WT and *itpa-1* plants that had been grown side-by-side in a randomized fashion and for which we had already confirmed differential levels of inosine in RNA (Fig. 9c). Plants lacking ITPA showed an upregulation of 256 transcripts and a downregulation of only 28 transcripts at a $\log_2$ fold change ($\log_2$ FC) of ≥ 2 and ≤ -2, respectively, in addition to a false discovery rate (FDR) of ≤ 0.05 (Table S2). Many regulated transcripts were involved in biological processes associated with stress responses and often encode proteins with kinase activity or transcription factors as reported by GO (gene ontology) enrichment analysis (Fig. S9; Table S3). Upon closer inspection, we found that especially transcripts related to the SA response (Fig. 11a), SAR (Fig. 11b), and aging (Fig. 11c) were upregulated.

Since the transcriptional profiles suggested that the concentration of SA is altered, we quantified it. Significantly higher SA concentrations were found in 35-d-old plants lacking ITPA compared with the corresponding WT (Fig. 11d).

Consistent with public data sets (Winter *et al.*, 2007), the expression of *ITPA* itself was not influenced by a treatment with SA (Fig. 11e), but treating WT plants with cadmium sulfate resulted in a minor, but significant, decrease in ITPA transcript abundance (Fig. 11e). Although counter-intuitive, a reduced expression of ITPA in oxidatively stressed plants has been reported recently (Arruebarrena Di Palma *et al.*, 2022). During development, the abundance of the ITPA transcript is decreasing (Fig. 11f), which is also in accordance with publicly available data (Winter *et al.*, 2007).

Discussion

Although the existence of ITPA in plants has been suggested based on sequence comparisons (Lin *et al.*, 2001; James *et al.*, 2021), these *in silico* predictions have never been experimentally tested nor has an *itpa* mutant been characterized in any plant species. Here, we demonstrate that plants have ITPA and that it is involved in the metabolite damage repair system in Arabidopsis. The kinetic properties of the enzyme from Arabidopsis are similar to those of the yeast ortholog HAM1p. Interestingly, the human ITPA has a 250-fold higher K_M and a 300- to 500-fold greater catalytic activity than the plant and yeast enzymes (Table 1), which might be an adaptation to high ITP concentrations in some human cells, for example, erythrocytes (Behmanesh *et al.*, 2009). However, in other human cells, the ITP concentration must be relatively low because it was generally not detectable (Sakumi *et al.*, 2010).

The K_M of Arabidopsis ITPA for ITP of 7.1 µM is *c.* 1000-fold higher than the cellular concentration of ITP (9.6 nM) determined in *itpa* background, which will be even lower in the WT. We were unable to detect dITP but could show that
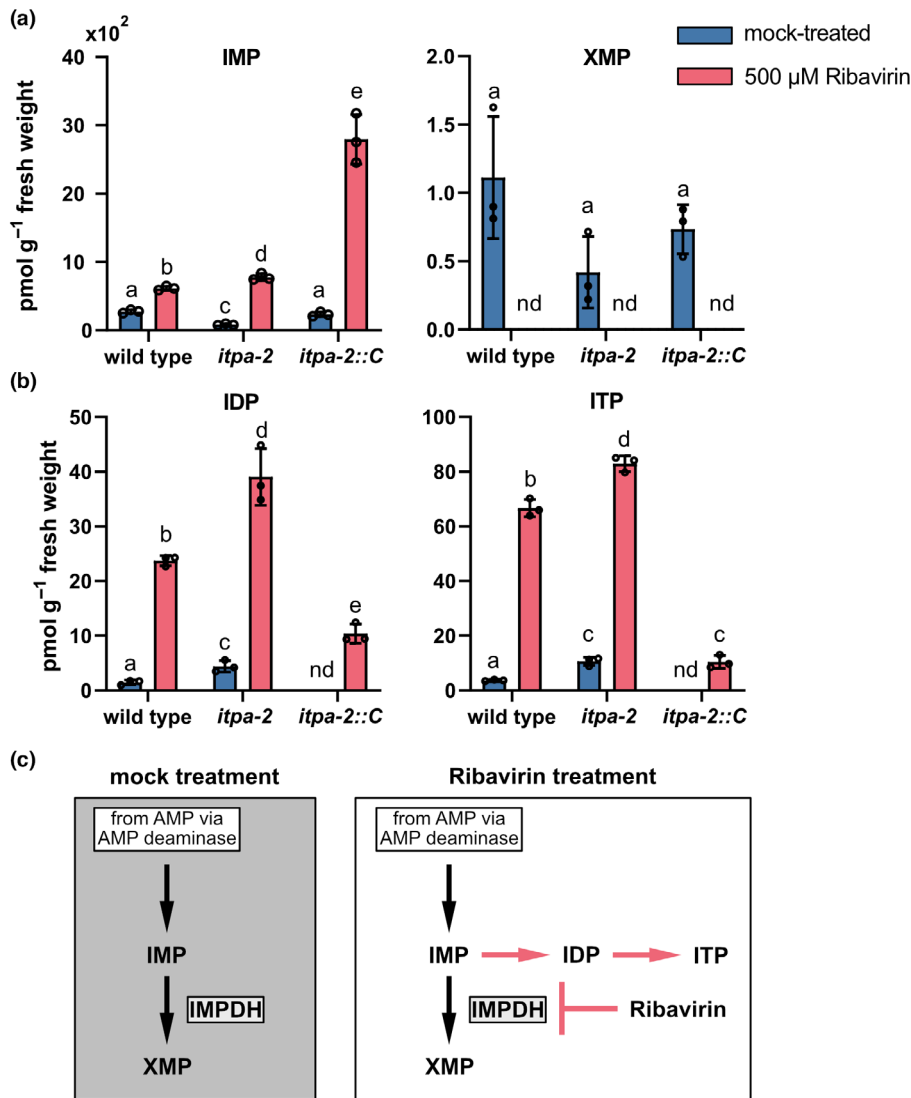
Fig. 7 Manipulation of inosine monophosphate (IMP) content in 7-d-old *Arabidopsis thaliana* plants alters the concentration of xanthosine monophosphate (XMP), inosine diphosphate (IDP), and inosine triphosphate (ITP). (a) Concentrations of IMP and XMP in 7-d-old Arabidopsis wild-type (WT), *itpa-2*, and *itpa-2::C* plants, grown hydroponically. Plants were either mock-treated or treated with 500 μM ribavirin for 24 h. (b) as in (a) but for IDP and ITP. Each open circle represents the result of measuring an independent biological replicate. Error bars are standard deviation (SD), $n = 3$. Every replicate represents a pool of seedlings from a different hydroponic vessel. Statistical analysis was performed by a two-sided Tukey's pairwise comparison using the sandwich variance estimator. Different letters indicate $P < 0.05$. All P-values are provided in Table S4. nd, not detected. (c) model of metabolic flux of parts of the purine metabolism in mock-treated and ribavirin-treated plants. Black arrows indicate the flux of purine metabolites in mock-treated plants, whereas red arrows indicate the proposed flux of purine metabolites in ribavirin-treated plants. The blunt-ended red arrow displays the inhibition of IMP DEHYDROGENASE (IMPDH) by ribavirin.

mutation of *ITPA* led to more deoxyinosine in DNA. Thus, dITP is likely an *in vivo* substrate of ITPA ($K_M = 1.7\,\mu M$ for dITP). We estimate that the dITP concentration is in the low pM range, assuming that the ratio of ITP to dITP is comparable to the ATP-to-dATP ratio of *c.* 500 to 1000 (Straube *et al.*, 2021). A difference in several orders of magnitude between K_M and cellular substrate concentration is also known for other housekeeping enzymes (Galperin *et al.*, 2006; Yoshimura *et al.*, 2007; Sakumi *et al.*, 2010; Heinemann *et al.*, 2021). Furthermore, it has been shown that a K_M above cellular concentrations of a metabolite seems to be typical for enzymes in *E. coli* involved in the degradation of nucleotides (Bennett *et al.*, 2009).

Xanthosine triphosphate, dXTP, and (deoxy)xanthosine in nucleic acids could not be detected, indicating that the deamination of (deoxy)guanosine in DNA and RNA and the phosphorylation of (d)XMP are very rare. Therefore, ITPA seems mainly involved with ITP and IDP and somewhat with dITP dephosphorylation *in vivo* (Fig. 10). The activity of Arabidopsis ITPA with IDP is interesting, since the yeast and human orthologs have

a rather poor activity with IDP (Holmes *et al.*, 1979; Davies *et al.*, 2012). Human cells possess a nudix hydrolase (NUDT16), which hydrolyzes preferably IDP and dIDP, acting synergistically with ITPA to protect cells from high levels of (d)ITP (Abolhassani *et al.*, 2010). This enzyme is missing in Arabidopsis, which may explain why the plant enzyme is also an efficient diphosphate hydrolase. IDP is likely the product of ADP deamination (Fig. 8), and ADP is also a substrate of the ribonucleotide reductase (RNR; Wang & Liu, 2006), an enzyme complex catalyzing the first committed step in the synthesis of deoxyribonucleotide triphosphates. Assuming that RNR accepts IDP as a substrate, the reaction would lead to the formation of deoxyinosine diphosphate, which could be phosphorylated and then be incorporated into DNA.

According to our data, an important source of ITP is probably the deamination of ATP. The concentration of the adenylates is by far the highest of all nucleotides (Straube *et al.*, 2021), and even a small rate of deamination will generate significant amounts of ITP. Additionally, this rate can be enhanced by
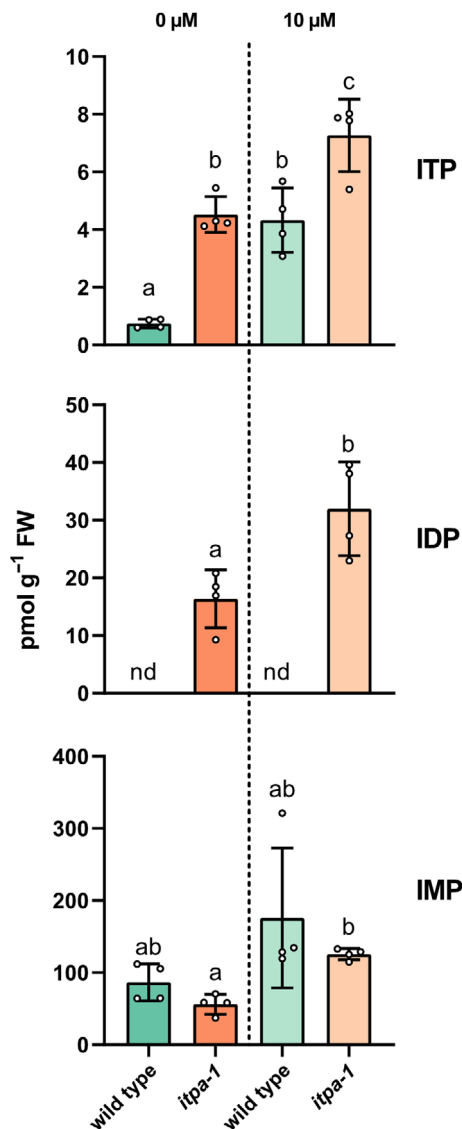
Fig. 8 Effect of cadmium causing oxidative stress on inosylate concentrations. Arabidopsis mutant and wild-type (WT) plants grown in a hydroponic system for 14 d were treated with 10 μM cadmium sulfate for 24 h or were left untreated. inosine triphosphate (ITP), inosine diphosphate (IDP), and inosine monophosphate (IMP) were quantified. Each open circle represents the result of measuring an independent biological replicate. Error bars are standard deviation (SD), $n = 4$. Every replicate represents a pool of plants from an independent hydroponic vessel. Statistical analysis was performed by a two-sided Tukey's pairwise comparison using the sandwich variance estimator. Different letters indicate $P < 0.05$. All P-values are provided in Table S4. nd, not detected.

Fig. 9 Concentration of deoxyinosine and inosine in nucleic acids of different genotypes. (a) Molecules of deoxyinosine per million molecules of deoxyadenosine in DNA isolated from pools of 7-d-old Arabidopsis seedlings grown on agar plates and 56-d-old Arabidopsis rosettes. Error bars are standard deviation (SD), $n = 5$ and 3. (b) Molecules of deoxyinosine per million molecules of deoxyadenosine in DNA isolated from pools of 7-d-old Arabidopsis seedlings grown hydroponically. Each open circle represents the result of measuring an independent biological replicate. Error bars are SD, $n = 3$. (c) Molecules of inosine per million molecules adenosine in total RNA isolated from 35-d-old Arabidopsis rosettes. Error bars are SD, $n = 5$. Every replicate represents a different plant. Statistical analysis was performed by a two-sided Tukey's pairwise comparison using the sandwich variance estimator. Different letters indicate $P < 0.05$. All P-values are provided in Table S4. nd, not detected.

oxidative stress, which can occur in plant cells because of abiotic and biotic stress. AMP kinases may also contribute to the IDP and ITP pool because they can use IMP as substrate albeit with lower activity than for AMP. The possibility of spurious phosphorylation by kinases involved in nucleotide metabolism has been discussed recently (Chen *et al.*, 2018; Chen & Witte, 2020). However, *in vivo* there is far more AMP than IMP and our manipulation of IMP in *N. benthamiana* leaves showed no correlation between IMP and ITP concentrations (Fig. S7). However,

in 7-d-old *A. thaliana* plants, the concentrations of IDP and ITP rose significantly when plants were treated with ribavirin. We hypothesize that this discrepancy is due to a different developmental context, as there is likely more overall nucleotide kinase activity in developing tissues than in fully developed leaves (Niehaus *et al.*, 2022), which would also increase the rate of spurious IMP phosphorylation.

It is surprising that ITPA is only located in the cytosol, nucleus, and plastids but not in the mitochondria because ATP occurs in each of these compartments and mitochondria also perform nucleic acid biosynthesis. Additionally, the respiratory chain and photosystems are the main sources for ROS in plant cells (Janků *et al.*, 2019), potentially leading to nucleotide and nucleic acid damage, such as nucleotide deamination (Cadet &

Fig. 10 Overview of deaminated purine nucleotide metabolism and its interplay with purine *de novo* synthesis. Enzymatic and nonenzymatic steps that lead to the generation of deaminated purine nucleotides are shown, as well as their removal by the inosine triphosphate pyrophosphatase (ITPA) and reintegration into canonical purine nucleotide metabolism or potential incorporation into nucleic acids. AMP, adenosine monophos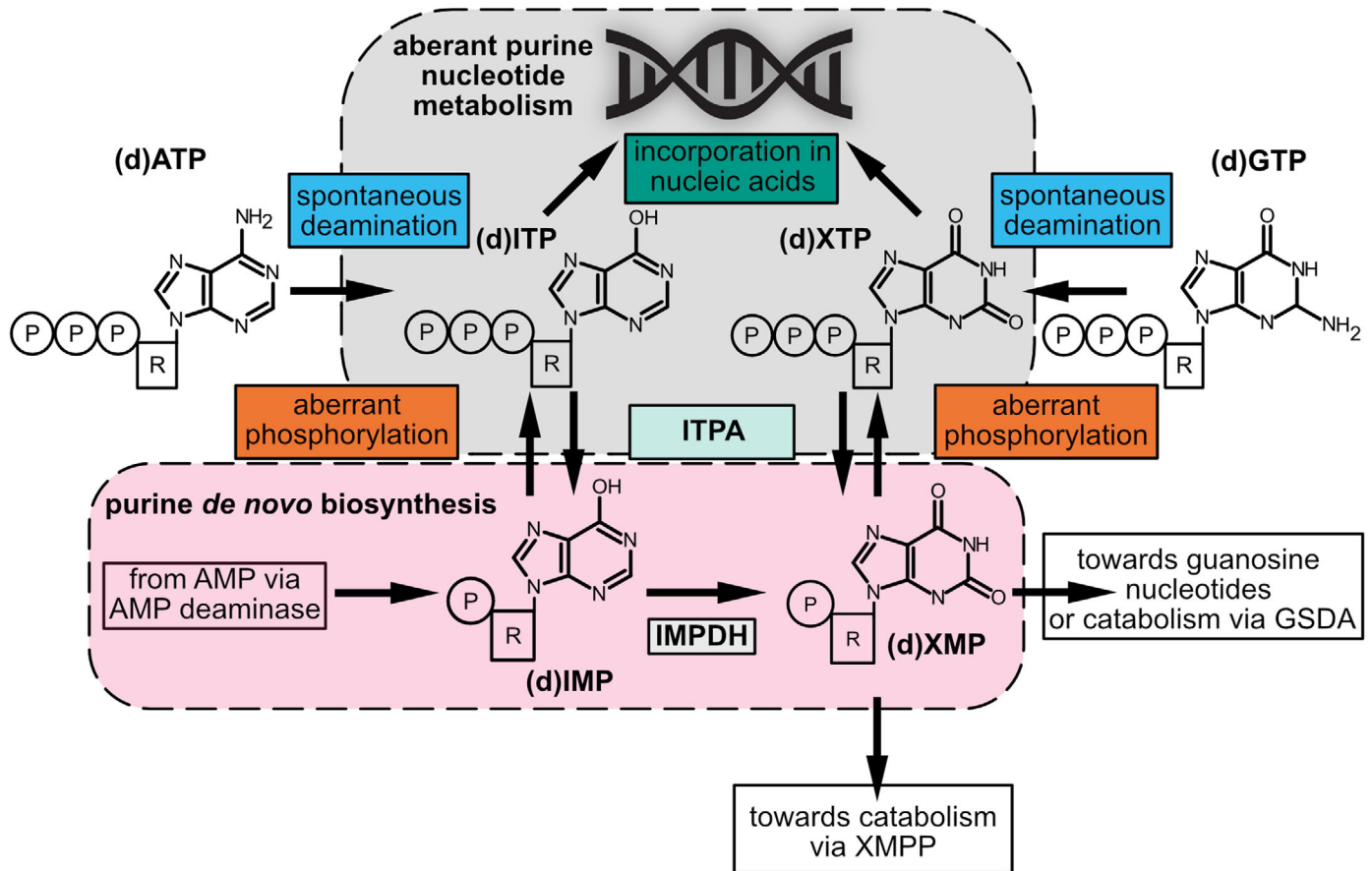phate; (d)ATP, (deoxy) adenosine triphosphate; (d)ITP, (deoxy)inosine triphosphate; (d)XTP, (deoxy)xanthosine triphosphate; (d)GTP, (deoxy)guanosine triphosphate; (d)IMP, (deoxy)inosine monophosphate; (d)XMP, (deoxy)xanthosine monophosphate; IMPDH, inosine monophosphate dehydrogenase; XMPP, xanthosine monophosphate phosphatase.

Wagner, 2013), in mitochondria and plastids, respectively (Tripathi *et al.*, 2020). It is known that ATP and ADP are rapidly exchanged over the mitochondrial membrane because the mitochondria supply the cell with ATP produced by oxidative phosphorylation (Braun, 2020). Therefore, cytosolic ITPA may be able to also purify the mitochondrial adenylate pool – this however implies that ITP and IDP are also rapidly exchanged over the mitochondrial membrane (Haferkamp *et al.*, 2011). By contrast, plastids exchange adenylates with the cytosol mainly at night for ATP import whereas during the day they internally generate ATP from ADP driven by photosynthesis for triosephosphate production. Plastids may require their own ITPA because in the light ROS production is high and adenylates are not exchanged with the cytosol.

Mutant plants lacking *ITPA* show a slightly earlier senescence and increased cell death (Fig. 3a–d). This is a relatively moderate phenotype compared to what occurs in mammals. Mice without functional ITPA die shortly after birth (Behmanesh *et al.*, 2009). A null mutation of *ITPA* in humans leads to severe developmental abnormalities and is also lethal (Handley *et al.*, 2019). Humans with missense mutations in *ITPA* develop severe side

effects when treated with purine analog like ribavirin or thiopurines (Simone *et al.*, 2013). By contrast, the single-cell eukaryote *S. cerevisiae* and also the model bacterium *E. coli* show no particular physiological phenotype when they lack ITPA (Pavlov, 1986; Clyman & Cunningham, 1987; Noskov *et al.*, 1996). Nonetheless, the *E. coli* mutant has an induced SOS response, a pathway that is turned on by DNA damage (Clyman & Cunningham, 1987).

In contrast to other housekeeping genes, a decrease in *ITPA* transcript abundance is observed during senescence (Fig. 11f). The housekeeping function of ITPA is centered on the proofreading of damaged metabolites, which is likely not as important in senescent plants as it is in developing tissues. Consistently, other metabolite damage repair enzymes, like NUDX1 (AT1G68760), show a similar expression profile during development (Winter *et al.*, 2007).

So why do plants lacking ITPA activity show symptoms of early senescence? This could be related to the accumulation of (deoxy)inosine in DNA and RNA in the mutant (Figs 9, S8). Deaminated purines in RNA were shown to lead to reduced polymerase fidelity, RNA structure changes, altered stability and
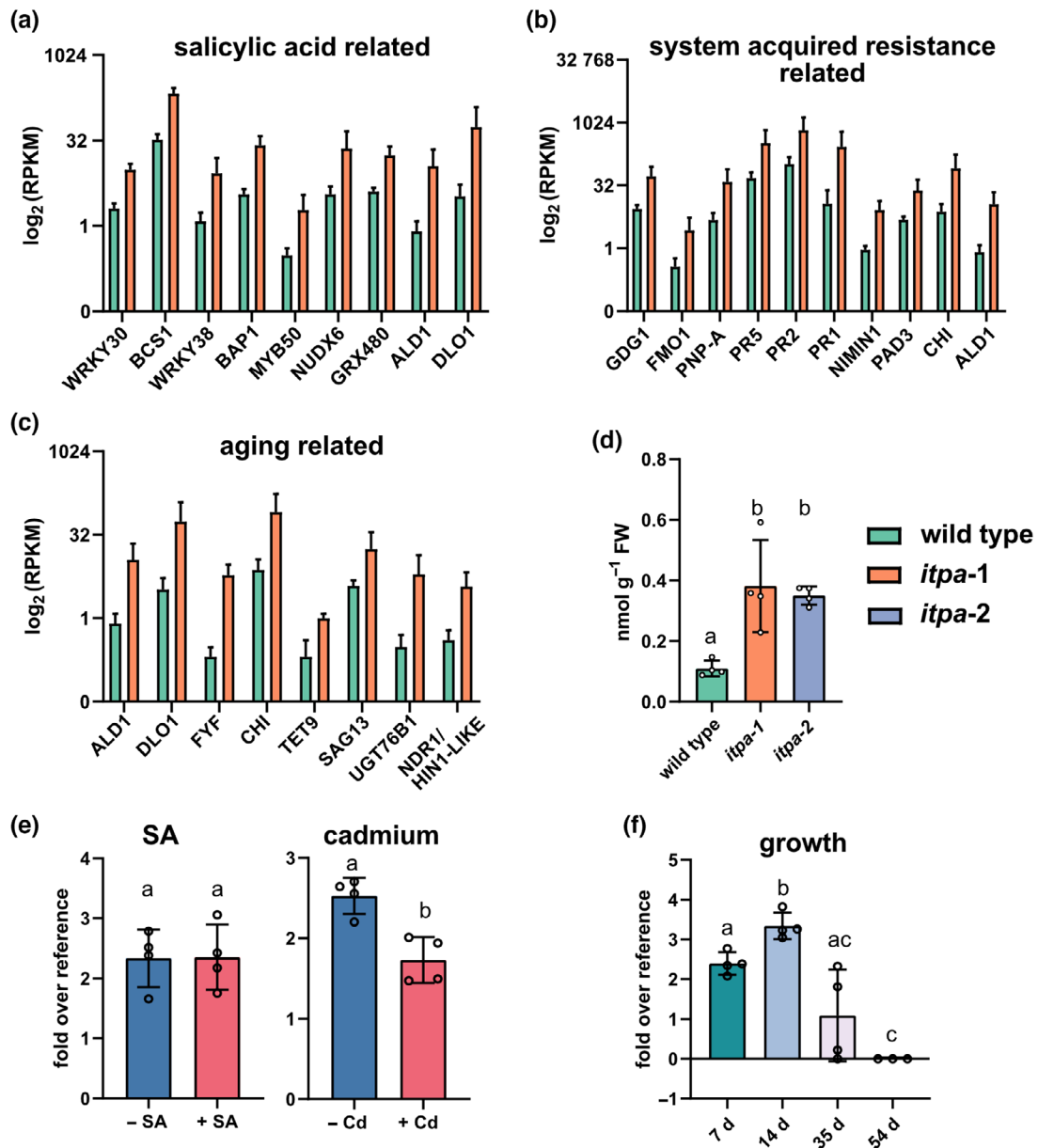
Fig. 11 Loss-of-function of inosine triphosphate pyrophosphatase (ITPA) leads to differential upregulation of transcripts associated with salicylic acid (SA), systemic acquired resistance (SAR) as well as aging and increased endogenous concentrations of SA. (a) RNA was isolated from 35-d-old Arabidopsis rosettes and analyzed by RNA-seq. Log$_2$ (RPKM) values are shown for genes associated with response to SA (GO:0009751) and differentially regulated between wild-type (WT) and *itpa-1* samples. Error bars are standard deviation (SD), $n = 3$. Every replicate represents a pool of two different plants. (b) As in (a) but for genes associated with SAR (GO:0009627). (c) As in (a) but for genes associated with aging (GO:0007568). (d) Concentrations of SA in 35-d-old Arabidopsis rosettes analyzed by LC–MS. Error bars are SD, $n = 4$. Every replicate represents a different plant. (e) Transcript abundance of the ITPA transcript in 14-d-old *A. thaliana* WT plants grown hydroponically. Plants were either mock-treated, treated with 1 mM SA or 500 μM cadmium sulfate for 24 h. Every replicate represents a pool of seedlings from a different hydroponic vessel. (f) ITPA transcript abundance in *A. thaliana* WT plants grown on soil for 7, 14, 35, or 54 d. Each open circle represents the result of measuring an independent biological replicate. Error bars are SD, $n = 3–4$. Every replicate represents a different plant, only in case of 7-d-old plants seedlings were pooled for RNA isolation. Statistical analysis was performed by a two-sided Tukey's pairwise comparison using the sandwich variance estimator. Different letters indicate $P < 0.05$. All P-values are provided in Table S4.

mistranslation (Thomas *et al.*, 1998; Ji *et al.*, 2017; Schroader *et al.*, 2022), and in case of DNA to mutations (Spee *et al.*, 1993; DeVito *et al.*, 2017) and genome instability (Yoshimura *et al.*, 2007). We observed a higher amount of inosine in RNA of older *itpa* plants (Figs 9, S8), indicating that the problems may become more severe with age. DNA damage can lead to programmed cell death (PCD) in Arabidopsis (Wang & Liu, 2006),

and we saw a higher incidence of PCD in the *itpa* background during senescence (Fig. 3d). An early senescence phenotype similar to that in *itpa* plants was also observed in plants with a higher frequency of unrepaired DNA double-strand breaks (Li *et al.*, 2020). In summary, damage to DNA and RNA by aberrant nucleotide incorporation may be the cause for the earlier senescence and the increased PCD rate. Alternatively, the detrimental

effects of *ITPA* mutation may also be unrelated to the alteration of nucleic acids but caused by ITP accumulation. ITP can directly influence cellular processes that either require or are inhibited by ATP or GTP (Vanderheiden, 1979; Osheroff *et al.*, 1983; Klinker & Seifert, 1997; Burton *et al.*, 2005). However, compared with the sizes of the ATP and GTP pools, the ITP pool is very small even in the *itpa* background, which argues against a direct metabolic interference of ITP.

The abundance of transcripts associated with SA, senescence, and SAR increases in *itpa* plants, and they accumulate SA (Figs 11, S9). Mutants with constitutive high levels of SA are known to senesce prematurely. The higher SA levels in *itpa* plants may thus present an explanation for the early senescence phenotype observed (Fig. 3a–d) and may affect plant–pathogen interactions. How the mutation of ITPA and the induction of SA are connected is unclear. We hypothesize that an accumulation of deoxyinosine in DNA causes more DNA repair and turnover associated with occasional double-strand breaks. It is known that in plants with more unrepaired DNA double-strand breaks the concentration of SA is increased (Li *et al.*, 2020). It also has been shown that SA can trigger DNA damage (Hadwiger & Tanaka, 2017) and that proteins involved in DNA damage response, like RAD51, interact synergistically to trigger increased transcription of genes involved in plant immunity (Yan *et al.*, 2013). Taken together, it is possible that the transcripts involved in SA response, SAR, and aging are upregulated as a consequence of DNA damage.

Deamination of adenylates is an unavoidable chemical reaction, and most organisms including plants have ITPA to protect their nucleic acids from the reaction products. Whether (d)ITP and (d)IDP are more than just an undesirable metabolic byproduct remains to be investigated.

Competing interests

None declared.

Author contributions

HS and MH designed the research; HS, JS, LF, JR and MN performed the research; HS, JS, LF, JR, MN and MH analyzed the data, and HS, CPW and MH interpreted the data; HS, CPW and MH wrote the manuscript. All authors read and revised the manuscript and agreed on the final version.

ORCID

Lisa Fischer https://orcid.org/0000-0001-5033-2490
Marco Herde https://orcid.org/0000-0003-2804-0613
Markus Niehaus https://orcid.org/0000-0003-3057-7823
Jannis Rinne https://orcid.org/0000-0003-0361-5257
Henryk Straube https://orcid.org/0000-0001-9286-7784
Jannis Straube https://orcid.org/0000-0002-6813-3809
Claus-Peter Witte https://orcid.org/0000-0002-3617-7807

Data availability

The data that support the findings of this study are available in the Supporting Information of this article. Information about the mutant lines in this article can be obtained in the GenBank/EMBL data libraries under the following accession number: At4g13720 (ITPA). Raw read data from Illumina sequencing are deposited at the Sequence Read Archive (SRA) on NCBI under the following accession number: PRJNA807491.

References

Abolhassani N, Iyama T, Tsuchimoto D, Sakumi K, Ohno M, Behmanesh M, Nakabeppu Y. 2010. NUDT16 and ITPA play a dual protective role in maintaining chromosome stability and cell growth by eliminating dIDP/IDP and dITP/ITP from nucleotide pools in mammals. *Nucleic Acids Research* 38: 2891–2903.

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R *et al.* 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653–657.

Arruebarrena Di Palma A, Perk EA, Carboni ME, Garcia-Mata C, Budak H, Tör M, Laxalt AM. 2022. The isothiocyanate sulforaphane induces respiratory burst oxidase homologue D-dependent reactive oxygen species production and regulates expression of stress response genes. *Plant Direct* 6: e437.

Baccolini C, Witte C-P. 2019. AMP and GMP catabolism in Arabidopsis converge on xanthosine, which is degraded by a nucleoside hydrolase heterocomplex. *Plant Cell* 31: 734–751.

Batista-Silva W, Heinemann B, Rugen N, Nunes-Nesi A, Araújo WL, Braun H-P, Hildebrandt TM. 2019. The role of amino acid metabolism during abiotic stress release. *Plant, Cell & Environment* 42: 1630–1644.

Behmanesh M, Sakumi K, Abolhassani N, Toyokuni S, Oka S, Ohnishi YN, Tsuchimoto D, Nakabeppu Y. 2009. ITPase-deficient mice show growth retardation and die before weaning. *Cell Death and Differentiation* 16: 1315–1322.

Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical Biology* 5: 593–599.

Bolte S, Cordelieres FP. 2006. A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy* 224: 213–232.

Braun H-P. 2020. The oxidative phosphorylation system of the mitochondria in plants. *Mitochondrion* 53: 66–75.

Budke B, Kuzminov A. 2006. Hypoxanthine incorporation is nonmutagenic in *Escherichia coli*. *Journal of Bacteriology* 188: 6553–6560.

Budke B, Kuzminov A. 2010. Production of clastogenic DNA precursors by the nucleotide metabolism in *Escherichia coli*. *Molecular Microbiology* 75: 230–245.

Burgis NE, Brucker JJ, Cunningham RP. 2003. Repair system for noncanonical purines in *Escherichia coli*. *Journal of Bacteriology* 185: 3101–3110.

Burton K, White H, Sleep J. 2005. Kinetics of muscle contraction and actomyosin NTP hydrolysis from rabbit using a series of metal-nucleotide substrates. *The Journal of Physiology* **563**: 689–711.

Cadet J, Wagner JR. 2013. DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Perspectives in Biology* **5**: a012559.

Cao Y, Tanaka K, Nguyen CT, Stacey G. 2014. Extracellular ATP is a central signaling molecule in plant stress responses. *Current Opinion in Plant Biology* **20**: 82–87.

Carnelli A, Michelis MI, Rasi-Caldogno F. 1992. Plasma membrane Ca-ATPase of radish seedlings I. Biochemical characterization using ITP as a substrate. *Plant Physiology* **98**: 1196–1201.

Chen M, Urs MJ, Sánchez-González I, Olayioye MA, Herde M, Witte C-P. 2018. m^{6}A RNA degradation products are catabolized by an evolutionarily conserved N^6-methyl-AMP deaminase in plant and mammalian cells. *Plant Cell* **30**: 1511–1522.

Chen M, Witte C-P. 2020. A kinase and a glycosylase catabolize pseudouridine in the peroxisome to prevent toxic pseudouridine monophosphate accumulation. *Plant Cell* **32**: 722–739.

Cho U-H, Seo N-H. 2005. Oxidative stress in *Arabidopsis thaliana* exposed to cadmium is due to hydrogen peroxide accumulation. *Plant Science* **168**: 113–120.

Clyman J, Cunningham RP. 1987. *Escherichia coli* K-12 mutants in which viability is dependent on recA function. *Journal of Bacteriology* **169**: 4203–4210.

Corpas FJ, Leterrier M, Valderrama R, Airaki M, Chaki M, Palma JM, Barroso JB. 2011. Nitric oxide imbalance provokes a nitrosative response in plants under abiotic stress. *Plant Science* **181**: 604–611.

Crécy-Lagard V, Haas D, Hanson AD. 2018. Newly-discovered enzymes that function in metabolite damage-control. *Current Opinion in Chemical Biology* **47**: 101–108.

Dahncke K, Witte C-P. 2013. Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in Arabidopsis. *Plant Cell* **25**: 4101–4109.

Davies O, Mendes P, Smallbone K, Malys N. 2012. Characterisation of multiple substrate-specific (d)ITP/(d)XTPase and modelling of deaminated purine nucleotide metabolism. *BMB Reports* **45**: 259–264.

Demidchik V. 2015. Mechanisms of oxidative stress in plants: from classical chemistry to cell biology. *Environmental and Experimental Botany* **109**: 212–228.

DeVito S, Woodrick J, Song L, Roy R. 2017. Mutagenic potential of hypoxanthine in live human cells. *Mutation Research* **803–805**: 9–16.

Di Sanità TL, Gabbrielli R. 1999. Response to cadmium in higher plants. *Environmental and Experimental Botany* **41**: 105–130.

Dobrzanska M, Szurmak B, Wyslouch-Cieszynska A, Kraszewska E. 2002. Cloning and characterization of the first member of the Nudix family from *Arabidopsis thaliana*. *Journal of Biological Chemistry* **277**: 50482–50486.

Dubois E, Córdoba-Cañero D, Massot S, Siaud N, Gakière B, Domenichini S, Guérard F, Roldan-Arjona T, Doutriaux M-P. 2011. Homologous recombination is stimulated by a decrease in dUTPase in Arabidopsis. *PLoS ONE* **6**: e18658.

Endo M, Kim JI, Shioi NA, Iwai S, Kuraoka I. 2021. *Arabidopsis thaliana* endonuclease V is a ribonuclease specific for inosine-containing single-stranded RNA. *Open Biology* **11**: 210148.

Fernández-Bautista N, Domínguez-Núñez JA, Moreno MMC, Berrocal-Lobo M. 2016. Plant tissue trypan blue staining during phytopathogen infection. *Bio-Protocol* **6**: e2078.

Fraser JH, Meyers H, Henderson JF, Brox LW, McCoy EE. 1975. Individual variation in inosine triphosphate accumulation in human erythrocytes. *Clinical Biochemistry* **8**: 353–364.

Frederico LA, Kunkel TA, Shaw BR. 1990. A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry* **29**: 2532–2537.

Gall AD, Gall A, Moore AC, Aune MK, Heid S, Mori A, Burgis NE. 2013. Analysis of human ITPase nucleobase specificity by site-directed mutagenesis. *Biochimie* **95**: 1711–1721.

Galperin MY, Moroz OV, Wilson KS, Murzin AG. 2006. House cleaning, a part of good housekeeping. *Molecular Microbiology* **59**: 5–19.

Hadwiger LA, Tanaka K. 2017. Non-host resistance: DNA damage is associated with SA signaling for induction of PR genes and contributes to the growth suppression of a pea pathogen on pea endocarp tissue. *Frontiers in Plant Science* **8**: 446.

Haferkamp I, Fernie AR, Neuhaus HE. 2011. Adenine nucleotide transport in plants: much more than a mitochondrial issue. *Trends in Plant Science* **16**: 507–515.

Handley MT, Reddy K, Wills J, Rosser E, Kamath A, Halachev M, Falkous G, Williams D, Cox P, Meynert A et al. 2019. ITPase deficiency causes a Martsolf-like syndrome with a lethal infantile dilated cardiomyopathy. *PLoS Genetics* **15**: e1007605.

Hanson AD, Henry CS, Fiehn O, Crécy-Lagard V. 2016. Metabolite damage and metabolite damage control in plants. *Annual Review of Plant Biology* **67**: 131–152.

Harris C, Baulcombe D. 2015. Chlorophyll content assay to quantify the level of necrosis induced by different R gene/elicitor combinations after transient expression. *Bio-Protocol* **5**: e1670.

Heinemann KJ, Yang S-Y, Straube H, Medina-Escobar N, Varbanova-Herde M, Herde M, Rhee S, Witte C-P. 2021. Initiation of cytosolic plant purine nucleotide catabolism involves a monospecific xanthosine monophosphate phosphatase. *Nature Communications* **12**: 6846.

Holmes SL, Turner BM, Hirschhorn K. 1979. Human inosine triphosphatase: catalytic properties and population studies. *Clinica Chimica Acta* **97**: 143–153.

Hooper CM, Castleden IR, Tanz SK, Aryamanesh N, Millar AH. 2017. SUBA4: the interactive data analysis centre for Arabidopsis subcellular protein locations. *Nucleic Acids Research* **45**: D1064–D1074.

James AM, Seal SE, Bailey AM, Foster GD. 2021. Viral inosine triphosphate: a mysterious enzyme with typical activity, but an atypical function. *Molecular Plant Pathology* **22**: 382–389.

Janků M, Luhová L, Petřivalský M. 2019. On the origin and fate of reactive oxygen species in plant cell compartments. *Antioxidants* **8**: 105.

Ji D, Stepchenkova EI, Cui J, Menezes MR, Pavlov YI, Kool ET. 2017. Measuring deaminated nucleotide surveillance enzyme ITPA activity with an ATP-releasing nucleotide chimera. *Nucleic Acids Research* **45**: 11515–11524.

Jiang H-P, Xiong J, Liu F-L, Ma C-J, Tang X-L, Yuan B-F, Feng Y-Q. 2018. Modified nucleoside triphosphates exist in mammals. *Chemical Science* **9**: 4160–4167.

Kamiya H. 2003. Mutagenic potentials of damaged nucleic acids produced by reactive oxygen/nitrogen species: approaches using synthetic oligonucleotides and nucleotides: survey and summary. *Nucleic Acids Research* **31**: 517–531.

Karren P, Lindahl T. 1980. Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus. *Biochemistry* **19**: 6005–6011.

Keya CA, Crozier A, Ashihara H. 2003. Inhibition of caffeine biosynthesis in tea (*Camellia sinensis*) and coffee (*Coffea arabica*) plants by ribavirin. *FEBS Letters* **554**: 473–477.

Klinker JF, Seifert R. 1997. Functionally nonequivalent interactions of guanosine 5′-triphosphate, inosine 5′-triphosphate, and xanthosine 5′-triphosphate with the retinal G-protein, transducin, and with Gi-proteins in HL-60 leukemia cell membranes. *Biochemical Pharmacology* **54**: 551–562.

Kong RPW, Siu SO, Lee SSM, Lo C, Chu IK. 2011. Development of online high-/low-pH reversed-phase-reversed-phase two-dimensional liquid chromatography for shotgun proteomics: a reversed-phase-strong cation exchange-reversed-phase approach. *Journal of Chromatography A* **1218**: 3681–3688.

Lange PR, Geserick C, Tischendorf G, Zrenner R. 2008. Functions of chloroplastic adenylate kinases in Arabidopsis. *Plant Physiology* **146**: 492–504.

Lerma-Ortiz C, Jeffryes JG, Cooper AJL, Niehaus TD, Thamm AMK, Frelin O, Aunins T, Fiehn O, Crécy-Lagard V, Henry CS et al. 2016. 'Nothing of chemistry disappears in biology': the top 30 damage-prone endogenous metabolites. *Biochemical Society Transactions* **44**: 961–971.

Li Z, Kim JH, Kim J, Lyu JI, Zhang Y, Guo H, Nam HG, Woo HR. 2020. ATM suppresses leaf senescence triggered by DNA double-strand break through epigenetic control of senescence-associated genes in Arabidopsis. *New Phytologist* **227**: 473–484.

Lin A-J, Zhang X-H, Chen M-M, Cao Q. 2007. Oxidative stress and DNA damages induced by cadmium accumulation. *Journal of Environmental Sciences* 19: 596–602.

Lin S, McLennan AG, Ying K, Wang Z, Gu S, Jin H, Wu C, Liu W, Yuan Y, Tang R et al. 2001. Cloning, expression, and characterization of a human INOSINE TRIPHOSPHATE PYROPHOSPHATASE encoded by the *itpa* gene. *Journal of Biological Chemistry* 276: 18695–18701.

Lindahl T. 1993. Instability and decay of the primary structure of DNA. *Nature* 362: 709–715.

Lindahl T, Nyberg B. 1974. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry* 13: 3405–3410.

Linster CL, Van Schaftingen E, Hanson AD. 2013. Metabolite damage and its repair or pre-emption. *Nature Chemical Biology* 9: 72–80.

Littlejohn GR, Gouveia JD, Edner C, Smirnoff N, Love J. 2010. Perfluorodecalin enhances in vivo confocal microscopy resolution of *Arabidopsis thaliana* mesophyll. *New Phytologist* 186: 1018–1025.

Liu S, Yu F, Yang Z, Wang T, Xiong H, Chang C, Yu W, Li N. 2018. Establishment of dimethyl labeling-based quantitative acetylproteomics in Arabidopsis. *Molecular & Cellular Proteomics* 17: 1010–1027.

Muraoka M, Fukuzawa H, Nishida A, Okano K, Tsuchihara T, Shimoda A, Suzuki Y, Sato M, Osumi M, Sakai H. 1999. The effects of various GTP analogues on microtubule assembly. *Cell Structure and Function* 24: 101–109.

Nagy GN, Leveles I, Vértessy BG. 2014. Preventive DNA repair by sanitizing the cellular (deoxy)nucleoside triphosphate pool. *The FEBS Journal* 281: 4207–4223.

Niehaus M, Straube H, Künzler P, Rugen N, Hegermann J, Giavalisco P, Eubel H, Witte C-P, Herde M. 2020. Rapid affinity purification of tagged plant mitochondria (Mito-AP) for metabolome and proteome analyses. *Plant Physiology* 182: 1194–1210.

Niehaus M, Straube H, Specht A, Baccolini C, Witte C-P, Herde M. 2022. The nucleotide metabolome of germinating *Arabidopsis thaliana* seeds reveals a central role for thymidine phosphorylation in chloroplast development. *Plant Cell* 34: 3790–3813.

Niehaus TD, Patterson JA, Alexander DC, Folz JS, Pyc M, MacTavish BS, Bruner SD, Mullen RT, Fiehn O, Hanson AD. 2019. The metabolite repair enzyme NIT1 is a dual-targeted amidase that disposes of damaged glutathione in Arabidopsis. *The Biochemical Journal* 476: 683–697.

Niehaus TD, Richardson LGL, Gidda SK, ElBadawi-Sidhu M, Meissen JK, Mullen RT, Fiehn O, Hanson AD. 2014. Plants utilize a highly conserved system for repair of NADH and NADPH hydrates. *Plant Physiology* 165: 52–61.

Noskov VN, Staak K, Shcherbakova PV, Kozmin SG, Negishi K, Ono B-C, Hayatsu H, Pavlov YI. 1996. *HAM1*, the gene controlling 6-N-hydroxylaminopurine sensitivity and mutagenesis in the yeast *Saccharomyces cerevisiae*. *Yeast* 12: 17–29.

Oberacker P, Stepper P, Bond DM, Höhn S, Focken J, Meyer V, Schelle L, Sugrue VJ, Jeunen G-J, Moser T et al. 2019. Bio-On-Magnetic-Beads (BOMB): open platform for high-throughput nucleic acid extraction and manipulation. *PLoS Biology* 17: e3000107.

Osheroff N, Shelton ER, Brutlag DL. 1983. DNA TOPOISOMERASE II from *Drosophila melanogaster*. Relaxation of supercoiled DNA. *Journal of Biological Chemistry* 258: 9536–9543.

Pang B, McFaline JL, Burgis NE, Dong M, Taghizadeh K, Sullivan MR, Elmquist CE, Cunningham RP, Dedon PC. 2012. Defects in purine nucleotide metabolism lead to substantial incorporation of xanthine and hypoxanthine into DNA and RNA. *Proceedings of the National Academy of Sciences, USA* 109: 2319–2324.

Pavlov II. 1986. Mutants of *Saccharomyces cerevisiae* supersensitive to the mutagenic effect of 6-N-hydroxylaminopurine. *Genetika* 22: 2235–2243.

Rampazzo C, Miazzi C, Franzolin E, Pontarin G, Ferraro P, Frangini M, Reichard P, Bianchi V. 2010. Regulation by degradation, a cellular defense against deoxyribonucleotide pool imbalances. *Mutation Research* 703: 2–10.

Rinne J, Witte C-P, Herde M. 2021. Loss of MAR1 Function is a marker for co-selection of CRISPR-induced mutations in plants. *Frontiers in Genome Editing* 3: 723384.

Rudd SG, Valerie NCK, Helleday T. 2016. Pathways controlling dNTP pools to maintain genome stability. *DNA Repair* 44: 193–204.

Sakumi K, Abolhassani N, Behmanesh M, Iyama T, Tsuchimoto D, Nakabeppu Y. 2010. ITPA protein, an enzyme that eliminates deaminated purine nucleoside triphosphates in cells. *Mutation Research* 703: 43–50.

Savchenko A, Proudfoot M, Skarina T, Singer A, Litvinova O, Sanishvili R, Brown G, Chirgadze N, Yakunin AF. 2007. Molecular basis of the antimutagenic activity of the house-cleaning inosine triphosphate pyrophosphatase RdgB from *Escherichia coli*. *Journal of Molecular Biology* 374: 1091–1103.

Schroader JH, Jones LA, Meng R, Shorrock HK, Richardson JI, Shaughnessy SM, Lin Q, Begley TJ, Berglund JA, Fuchs G et al. 2022. Disease-associated inosine misincorporation into RNA hinders translation. *Nucleic Acids Research* 50: 9306–9318.

Simone PD, Pavlov YI, Borgstahl GEO. 2013. ITPA (inosine triphosphate pyrophosphatase): from surveillance of nucleotide pools to human disease and pharmacogenetics. *Mutation Research* 753: 131–146.

Spee JH, Vos WM, Kuipers OP. 1993. Efficient random mutagenesis method with adjustable mutation frequency by use of PCR and dITP. *Nucleic Acids Research* 21: 777–778.

Stenmark P, Kursula P, Flodin S, Gräslund S, Landry R, Nordlund P, Schüler H. 2007. Crystal structure of human inosine triphosphatase. Substrate binding and implication of the inosine triphosphatase deficiency mutation P32T. *Journal of Biological Chemistry* 282: 3182–3187.

Straube H, Herde M. 2022. Purification and analysis of nucleotides and nucleosides from plants. In: Ayyar BV, Arora S, eds. *Affinity chromatography. Methods in molecular biology, vol. 2466*. New York, NY: Humana, 145–155.

Straube H, Niehaus M, Zwittian S, Witte C-P, Herde M. 2021. Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. *Plant Cell* 33: 270–289.

Thomas MJ, Platas AA, Hawley DK. 1998. Transcriptional fidelity and proofreading by RNA POLYMERASE II. *Cell* 93: 627–637.

Tomlinson KR, Pablo-Rodriguez JL, Bunawan H, Nanyiti S, Green P, Miller J, Alicai T, Seal SE, Bailey AM, Foster GD. 2019. *Cassava brown streak virus* HAM1 protein hydrolyses mutagenic nucleotides and is a necrosis determinant. *Molecular Plant Pathology* 20: 1080–1092.

Traube FR, Schiffers S, Iwan K, Kellner S, Spada F, Müller M, Carell T. 2019. Isotope-dilution mass spectrometry for exact quantification of noncanonical DNA nucleosides. *Nature Protocols* 14: 283–312.

Tripathi D, Nam A, Oldenburg DJ, Bendich AJ. 2020. Reactive oxygen species, antioxidant agents, and DNA damage in developing maize mitochondria and plastids. *Frontiers in Plant Science* 11: 596.

Valderrama R, Corpas FJ, Carreras A, Fernández-Ocaña A, Chaki M, Luque F, Gómez-Rodríguez MV, Colmenero-Varea P, Del Río LA, Barroso JB. 2007. Nitrosative stress in plants. *FEBS Letters* 581: 453–461.

Valli AA, López RG, Ribaya M, Martínez FJ, Gómez DG, Garcia B, Gonzalo I, de Prádena AG, Pasin F, Montanuy I et al. 2022. Maf/ham1-like pyrophosphatases of non-canonical nucleotides are host-specific partners of viral RNA-dependent RNA polymerases. *PLoS Pathogens* 18: e1010332.

Vanderheiden BS. 1979. Possible implication of an inosinetriphosphate metabolic error and glutamic acid decarboxylase in paranoid schizophrenia. *Biochemical Medicine* 21: 22–32.

Vidal AE, Yagüe-Capilla M, Martínez-Arribas B, García-Caballero D, Ruiz-Pérez LM, González-Pacanowska D. 2022. Inosine triphosphate pyrophosphatase from *Trypanosoma brucei* cleanses cytosolic pools from deaminated nucleotides. *Scientific Reports* 12: 6408.

Wang C, Liu Z. 2006. Arabidopsis ribonucleotide reductases are critical for cell cycle progression, DNA damage repair, and plant development. *Plant Cell* 18: 350–365.

Weber J, Senior AE. 2001. Bi-site catalysis in F1-ATPase: does it exist? *Journal of Biological Chemistry* 276: 35422–35428.

Werner AK, Sparkes IA, Romeis T, Witte C-P. 2008. Identification, biochemical characterization, and subcellular localization of Allantoate amidohydrolases from Arabidopsis and soybean. *Plant Physiology* 146: 418–430.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. 2007. An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2: e718.

Witte C-P, Herde M. 2020. Nucleotide metabolism in plants. *Plant Physiology* **182**: 63–78.

Witte C-P, Rosso MG, Romeis T. 2005. Identification of three urease accessory proteins that are required for urease activation in Arabidopsis. *Plant Physiology* **139**: 1155–1162.

Yan S, Wang W, Marqués J, Mohan R, Saleh A, Durrant WE, Song J, Dong X. 2013. Salicylic acid activates DNA damage responses to potentiate plant immunity. *Molecular Cell* **52**: 602–610.

Yang Z-B, He C, Ma Y, Herde M, Ding Z. 2017. Jasmonic acid enhances Al-induced root growth inhibition. *Plant Physiology* **173**: 1420–1433.

Yoneshima Y, Abolhassani N, Iyama T, Sakumi K, Shiomi N, Mori M, Shiomi T, Noda T, Tsuchimoto D, Nakabeppu Y. 2016. Deoxyinosine triphosphate induces MLH1/PMS2- and p53-dependent cell growth arrest and DNA instability in mammalian cells. *Scientific Reports* **6**: 32849.

Yoshimura K, Ogawa T, Ueda Y, Shigeoka S. 2007. *ATNUDX1*, an 8-oxo-7,8-dihydro-2′-deoxyguanosine 5′-triphosphate pyrophosphohydrolase, is responsible for eliminating oxidized nucleotides in *Arabidopsis. Plant & Cell Physiology* **48**: 1438–1449.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Comparative sequence analysis of inosine triphosphate pyrophosphatases from several phylogenetically distant model organisms of all kingdoms of life.

Fig. S2 Comparative sequence analysis of inosine triphosphate pyrophosphatases.

Fig. S3 Phenotyping of wild-type and *itpa* mutant plants.

Fig. S4 Characterization of *itpa* mutant and overexpressor lines.

Fig. S5 Chromatograms of inosine triphosphate and inosine diphosphate analyzed by hypercarb chromatography.

Fig. S6 Relative concentration of different nucleotides in rosette leaves of 35-d-old wild-type and itpa mutant plants.

Fig. S7 Manipulation of inosine monophosphate content in leaves of *Nicotiana benthamiana* alters the concentration of guanosine monophosphate and xanthosine monophosphate, but not inosine triphosphate and guanosine triphosphate.

Fig. S8 Concentration of inosine in isolated RNA of wild-type and *itpa-1* plants.

Fig. S9 Loss-of-function of inosine triphosphate pyrophosphatase leads to differential upregulation of transcripts associated with GO terms of biological processes, molecular functions, and cellular components.

Methods S1 Supporting Methods.

Table S1 List of primers used in this study.

Table S2 List of differentially expressed genes.

Table S3 List of the GO term enrichment analysis.

Table S4 Raw and statistical data.

Table S5 Substrates tested for biochemical activity with inosine triphosphate pyrophosphatase.

Table S6 Ratios (deoxy)inosine per million molecules (deoxy) adenosine in DNA and RNA, respectively.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

```
Gly max   1 --------------------------MAAARGLVLSRPVTFVTANAKKLEEVRAIL--G
Pha vul   1 MFINIDEYGLILRRSALLQKNQTEIWEMAAARGLLLSRPVTFVTGNAKKLEEVRAIL--G
Med tru   1 ----------------------MAAIKSAASGLVLPRPVTFVTGNAKKLEEVRAIL--G
Ara tha   1 ---------------------MAAAAAKAAVVLPRPVTFVTGNAKKLEEVKAII--G
Sol tub   1 ------------------MAAAARTVGSGVVLPRSVTFVTGNAKKLEEVRAIL--G
Sol lyc   1 ------------------MAAAAARTVGSGVVLPRSVTFVTGNAKKLEEVRTIL--G
Tri aes   1 ----------------------MSGAAAAVARVLPKAVTFVTGNAKKLEEVRAIL--G
Zea may   1 -----------------------MSAAARVLPKAVTFVTGNAKKLEEVRAIL--G
Ory sat   1 ----------------------MSGAAARALPKAVTFVTGNAKKLEEVRAIL--G
Mar pol   1 ----------------------MSAMGVRMQLSSPVTFVTGNANKLEEVRMIL--G
Sel moe   1 ----------------------AMALVRAEVVLKKPVTFVTGNAKKLEEVRMIL--G
Phy pat   1 ---------------------MAAMKVQMTLTKPVTFVTGNAKKLEEVRMIL--G
Dro mel   1 -----------------------------MSKPITFVTGNAKKLEEIVAIL--G
Hom sap   1 --------------------------MAASLVGKKIVFVTGNAKKLEEVVQIL--G
Sus scr   1 --------------------------MANSLVGKKIVFVTGNAKKLEEVIQIL--G
Xen lae   1 ----------------------MAAAAGRSIVFVTGNAKKLEEVVQIL--G
Dan rer   1 -----------------------MAVPTGRALVFVTGNAKKLEEVVQIL--G
Gal gal   1 ----------------------MAAPVRRSVVFVTGNAKKLEEVTQIL--G
Cae ele   1 ---------------------------MSLRKINFVTGNVKKLEEVRAIL---
Vol car   1 ---------------------------MATPKKIYFATGNKKKLEEVTAILQSG
Chl rei   1 ---------------------------MTTPSKIHFATGNKKKLEEVNAILAAG
Bot cin   1 ---------------------------MAPPKTLNFITGNKNKLIEVKAIL--G
Neu cra   1 ------------------------XMSAPSQARHIVNFITGNANKIGEVKAIL--E
Asp nig   1 ---------------------------MASLEKLNFITGNKNKLAEVKAIL--G
Ust may   1 ---------------------------XMSNPTLTFVTGNANKLREVQQIFSLT
Sac cer   1 ----------------------------MSNNEIVFVTGNANKLKEVQSII--T
Met jan   1 --------------------------------MKIYFATGNPNKIKEANIIL--K
Bac sub   1 --------------------------------XMKEAIIATHNPGKVKEFREIL---
Dei rad   1 ---------------------XMQQQTGGRRRQIRRVVVATSNAGKVRETQGAL---
Esc col   1 --------------------------------MQKVVLATGNAGKVRELASIL---

Gly max  32 NSIPF--QSLK------LDLPELQG-EPEDISKEKARIAA--LQVNGPVLVEDTCLCFNA
Pha vul  59 NSIPF--QSLK------LDLPELQG-EPEDISKEKARMAA--LQVNGPVLVEDTCLCFNA
Med tru  36 HSIPF--QSLK------LDLPELQG-EPEDISKEKARLAA--IQVKGPVLVEDTCLCFNA
Ara tha  35 NSIPF--KSLK------LDLPELQG-EPEDISKEKARLAA--LQVDGPVLVEDTCLCFNA
Sol tub  37 QSIPF--QSLK------LDLPELQG-EPEDISKEKARIAA--KEVNGPVLVEDTCLCFNA
Sol lyc  38 QSIPF--QSLK------LDLPELQG-EPEDISKEKARIAA--KEVNGPVLVEDTCLCFNA
Tri aes  35 SSIPF--QSLK------LDLPELQG-EPEDISKEKARMAA--SQVNGPVLVEDTCLCFNA
Zea may  31 SSVPF--QSLK------LDLPELQG-EPEYISKEKARIAA--SQVNGPVLVEDTCLCFNA
Ory sat  32 SSIPF--QSLK------LDLPELQG-EPEDISKEKARMAA--SQVNGPVLVEDTCLCFNA
Mar pol  33 QTIPF--RSVK------LDLPELQG-EPEDISKEKCRLAA--KEVNGPVLVEDTCLCFNA
Sel moe  34 NSIPF--STLR------VDLPELQG-EPEEISKEKARIAA--KQIDGAVLVEDTCLCFNA
Phy pat  33 QSIPF--QSLK------LDLPELQG-EPEDISKEKARLAA--KEIGGPVLVEDTCLCFNA
Dro mel  24 PSFPRTIVSKK------IDLPELQG-DIDEIAIKKCKEAA--RQVNGPVLVEDTSLCFNA
Hom sap  29 DKFPCTLVAQK------IDLPEYQG-EPDEISIQKCQEAV--RQVQGPVLVEDTCLCFNA
Sus scr  29 DKFPCTLVAQK------IDLPEYQG-EPDEISIQKCKEAA--RQVQGPVLVEDTCLCFNA
Xen lae  28 DKFPCKLVAKK------IDLPEYQG-EPDEISIQKCREAA--KQIQGPVIVEDTCLCFNA
Dan rer  28 DKFPYKLISKK------IDLPEYQG-EPDDISIQKCKEAA--RQVDGPVLVEDTCLCFRA
Gal gal  28 DSSPYTLVARK------IDLPEYQG-EPDEISVQKCREAA--RQIRGPVIVEDTCLCFNA
Cae ele  24 --KNFEVSNVD------VDLDEFQG-EPEFIAERKCREAV--EAVKGPVLVEDTSLCFNA
Vol car  28 APLPFVMEAVK------LDLPELQG-EPEEISKEKCRIAA--KLVGGAVMVEDTSLCFNA
Chl rei  28 AELPFEVVAAK------LDLPELQG-EPEEISKEKCRIAA--KLVGGAVMVEDTSLCFNA
Bot cin  26 DTIDL--QSQS------LDIVEIQG-TIEEISSDKCRRAA--EIIQGPVLVEDTCLCFNA
Neu cra  31 PATQV--ENQA------LDILEIQG-TLEEVTLDKCRRAAXXDLVQGPVLVEDTCLCFNA
Asp nig  26 TVIDV--ENQA------VDLPEIQG-TIEETAREKCRRAA--EVVGGPVLTEDTALEFHA
Ust may  28 PNFPYELTNKD-----LDLPEIQG-TTRDVAQAKCAAAAKXXALGGACITEDTALGFHA
Sac cer  25 QEVDNNNKTIHLINEA-LDIEELQDTDLNAIALAKGKQAVAALGKGKPVFVEDTAIRFDE
Met jan  22 DLKDVEIEQIK------ISYPEIQG-TLEEVAEFGAKWVY--NILKKPVIVEDSGFFVEA
Bac sub  23 EPRGYDVKSLA--EIGFTEIEETGHTFEENAILKAEAVAKXXAVNKMVIAFDSGLSIDN
Dei rad  34 --APLGWQCEG---LGAVTLPEETGSTYEENAAXXLKACAAAMATGLPAIAFDSGIEVLA
Esc col  22 --SDFGLDIVAQTDLG-VDSAFETGLTFIENAILKARHAA--KVTGLPAIAFDSGLAVDV
```

210

```
Gly max   81  LKGLPGPYIKWF--LQKLGHEGLNNLLMAY-----DDKSAYA---LCVFSFAA---GPDS
Pha vul  108  LKGLPGPYIKWF--LQKLGHEGLNNLLMAY-----DDKSAFA---LCVFSFAA---GPDS
Med tru   85  LKGLPGPYIKWF--LQKLGHEGLNNLLMAY-----DDKSAFA---LCVFSFAI---GPDS
Ara tha   84  LKGLPGPYIKWF--LEKLGHEGLNNLLMAY-----EDKSAYA---LCAFSFSR---GPGA
Sol tub   86  LKGLPGPYIKWF--LQKISHEGLNNLLMAY-----EDKTAYA---MCIFSLAL---GPNT
Sol lyc   87  LKGLPGPYIKWF--LQKLGHEGLNNLLMAY-----EDKTAYA---MCIFSLAL---GPNA
Tri aes   84  LKGLPGPYIKWF--LEKLGHEGLNNLLKAY-----EDKSAFA---MCIFSLAL---GPEE
Zea may   80  LKGLPGPYIKWF--LEKLGHEGLNNLLKAY-----EDKSAFA---MCIFSLAL---GPGE
Ory sat   81  LKGLPGPYIKWF--LEKTGHEGLNNLLLAY-----EDKSAFA---MCIFSL-----GPGE
Mar pol   82  LKGLPGPYVKWF--LQKTGHEGLNNLLAAY-----EDKSAYA---LCVFSLAL---GPDA
Sel moe   83  LHGLPG-FEKWF--LQKLGHEGLNNVLAAY-----KDKSAYA---LCVFSLAL---GPGF
Phy pat   82  LNGLPGPVKWF---IMKTGHEGLNNLLAAY-----EDKTAYA---LCVFSLAL---GPDF
Dro mel   75  LEGLPGPYIKWF--LEKIQPEGLHRLLHGW-----ENKSAQA---ICTFGYCD---GVDA
Hom sap   80  LGGLPGPYIKWF--LEKLKPEGLHQLLAGF-----EDKSAYA---LCTFALST---GDPS
Sus scr   80  FGGLPGPYIKWF--LEKLKPEGLHQLLAGF-----QDKSAYA---LCTFALST---GDPN
Xen lae   79  LGGLPGPYIKWF--LEKIKPEGLHRVLEGF-----EDKSAIA---LCTFAYCN---GNPD
Dan rer   79  LEGLPGPYIKWF--LDKLKPEGLYKVLAGF-----EDKSAWA---LCTFAFCA---GKE-
Gal gal   79  LGGLPGPYIKWF--LEKLKPEGLYKLLAGF-----EDKSAYA---LCTFAFST---GNPE
Cae ele   73  MGGLPGPYIKWF--LKNLKPEGLHNVLAGF-----SDKTAYA---QCIFAYTE---GLG-
Vol car   79  LKGLPGPYIKWF--LEKLGHDGLNKVLAGF-----DDKTAYA---QCIFAYTT---GPEV
Chl rei   79  MHGLPGPYIKWF--LEKLGHDGLNRVLAGF-----EDKSAYA---QCIFAYTP---GPDT
Bot cin   75  LKELPGPYIKWF--MDALGHDGLNNVLAGF-----PDKSAQA---VCTFAYSE---GPGH
Neu cra   82  LKGLPGPYIKWF--MNSLGHEGLNNLLAAY-----EDKSAKA---VCTFGYSXXXAGPGH
Asp nig   75  LKGLPGPYIKSF--LDALGHEGLNKLLDSF-----ETRAAEA---VCTFAFSS---GPGS
Ust may   81  LGGLPGPYIKDF--MKTLGHDGLNKVLDGF-----EDRTASA---ICTFAYCAXXXGPDE
Sac cer   84  FNGLPGAYIKWF--LKSMGLEKLVKMLEPF-----ENKNAEA---VTTICFAD---SRGE
Met jan   73  LNGFPGTYSKFV--QETLGNEGILKLLECK-----DNRNAYF---KTVIGYCD------E
Bac sub   81  LGGRPGVYSARYAGEQKDDQANLNKVLSELKGIEKEQRTARFRXXXCALAVSI----PGE
Dei rad   89  LGGQPGVYSARF---GNVNSDVERNVLLLEKMRRHTDRXXXRAKFVSVLVLAY----PDG
Esc col   77  LGGAPGIYSARYSGEDATDLKNLQKLLETLKDVPDGQRQARF---HCVLVYLR--HAEDP

Gly max  128  EPI-TFSGKTPGKIVP-PRGPNDFG--WDPIFEP-DGYDQTYAQMPKEE---KNKIS---
Pha vul  155  EPI-TFSGKTPGKIVP-PRGPNDFG--WDPIFEP-EGYDQTYAQMPKEE---KNKIS---
Med tru  132  EPL-TFSGKTPGKIVP-PRGPNDFG--WDPVFQP-DGYDQTYAEMSKEE---KNKIS---
Ara tha  131  EPL-TFLGKTPGKIVP-ARGPTDFG--WDPVFQP-DGYDQTYAEMAKEE---KNKIS---
Sol tub  133  EPM-TFVGKTLGRIVP-ARGPNDFG--WDPIFQP-HGYDQTYAEMPKEE---KNKIS---
Sol lyc  134  EPM-TFVGKTQGRIVP-ARGPNDFG--WDPIFQP-HGYDQTYAEMPNEE---KHKIS---
Tri aes  131  EPI-TFVGKTAGKIVP-ARGPADFG--WDPVFQP-DGFEQTYAEMPKPE---KNQIS---
Zea may  127  EPI-TFVGKTAGKIVP-ARGPNYFG--WDPVFQP-DGFEQTYAEMPKSV---KNNIS---
Ory sat  126  EPM-TFVGKTAGKIVP-ARGPADFG--WDPVFQP-DGFDQTYAEMPKSV---KNQIS---
Mar pol  129  EPL-TFSGRTEGKIVP-ARGPGDFG--WDPVFQP-DGYDLTFAEMKKED---KNKIS---
Sel moe  129  EPI-TFVGRTEGKIVP-ARGPADFG--WDPVFQP-DGSDFTYAEMPKDE---KNKIS---
Phy pat  129  EPI-TFSGRTEGKIVP-ARGSGNFG--WDPIFQP-VGSDFTFAEMLKDE---KNKIS---
Dro mel  122  EPL-IFKGITEGVIVE-PRGPRDFG--WDPVFQP-SGYDKTYAELPKSE---KNTIS---
Hom sap  127  QPVRLFRGRTSGRIVA-PRGCQDFG--WDPCFQP-DGYEQTYAEMPKAE---KNAVS---
Sus scr  127  EPVRLFKGRTSGQIVV-PRGSRDFG--WDPCFQP-DGYEQTYAEMPKAE---KNTIS---
Xen lae  126  DTVLLFRGKTLGQIVL-PRGPRDFG--WDPCFQP-DGFQQTYAELPKEV---KNTIS---
Dan rer  125  EPVQLFRGITEGHIVE-PRGPRDFG--WDPCFQP-EGYDKTYAELPKEV---KNSIS---
Gal gal  126  EPVKLFKGQTHGVIVE-PRGPRDFG--WDPCFQP-DGYDQTYAELPKAV---KNSIS---
Cae ele  119  KPIHVFAGKCPGQIVA-PRGDTAFG--WDPCFQP-DGFKETFGEMDKDV---KNEIS---
Vol car  126  EPI-VFVGRTPGRIVA-ARGPPDFG--WDPIFEP-EGFETTYAEMDKET---KNKIS---
Chl rei  126  EPI-VFVGRTEGRIVQ-ARGPTDFG--WDPVFLP-DGFTDTYAEMDKTT---KNTIS---
Bot cin  122  EPI-IFQGRTDGKIVP-ARGPTAFG--WDPIFEY-EG--QTYAEMEKVE---KNKIS---
Neu cra  132  EPI-LFQGITDGKIVP-PRGPPNFG--WDAIFEY-EG--QTYAEMDKAE---KNKIS---
Asp nig  122  DPI-LFQGRTEGAIVR-PRGPANFG--WDPIFEH-QG--KTYAEMDKEE---KNRIS---
Ust may  131  QVH-LFEGRTEGVIVP-PRGPTHFG--WDPILEI-KGTGLTYAEMDPKQ---KNTIS---
Sac cer  131  YHF--FQGITRGKIVP-SRGPTTFG--WDSIFEPFDGHGLTYAEMSKDA---KNAIS---
Met jan  117  NGVRLFKGIVKGRVSE-EIRSKGYGFAYDSIFIP-EEEERTFAEMTTEE---KSQIS---
Bac sub  137  ETK-TVEGHVEGYIAEEPRGEYGFG--YDPIFIV-KDKDKTMAELTSDE---KNKISXXX
Dei rad  142  KLE-EYRGEVTGQLLEGPRGESGFG--YDPLFLP-DGSELSMGEMTLEQXXXKCAIS---
Esc col  132  TPL-VCHGSWPGVITREPAGTGGFG--YDPIFFV-PSEGKTAAELTREE---KSAIS---
```

```
Gly max  177  HRSKSLAL---VKSHFAEARF----------------TFDVKNF----
Pha vul  204  HRSKSLAL---VKSHFAEAGY----------------AFEINNV----
Med tru  181  HRSKSLAL---VKSHFAEAGY----------------TFQI-------
Ara tha  180  HRYKSLAL---VKSHFKEAGY----------------VFQTDDGTI--
Sol tub  182  HRGKALEL---VKLHFAEARY----------------TFQTDSTTQN-
Sol lyc  183  HRGKALEL---VKLHFAEARY----------------TFQTDSTT---
Tri aes  180  HRGRALAL---VKEHFASANY----------------EVQGDGLA---
Zea may  176  HRGKALAL---VKEHFASASY----------------TVQSNDST---
Ory sat  175  HRGKALAL---VKEHFAAANY----------------KVQNDGSA---
Mar pol  178  HRRRALDK---VREHFYENDY----------------APPV-------
Sel moe  178  HRRRALDK---VRDHFREYDF----------------VVRNDDS----
Phy pat  178  HRRRALDK---VKEYFYDFNY----------------APRSGNV----
Dro mel  171  HRYRALAL---LRQHFEKQDK----------------LIN--------
Hom sap  177  HRFRALLE---LQEYFGSLAA-----------------------
Sus scr  177  HRFRALLE---LQEYFGSLTPPGAEAAHVISPLRPWSAVHSLSPNPVQ-
Xen lae  176  HRYRALKE---MSDYFIQNGT----------------KV--------
Dan rer  175  HRYRALAA---LSEHFCQDNG----------------APETKRSKHQD
Gal gal  176  HRYRALSE---LSAFFLQSNP----------------TEAPSSPS---
Cae ele  169  HRAKALEL---LKEYFQNN-------------------------
Vol car  175  HRYRSLDL---LRSHLQHHSE----------------KSS-------
Chl rei  175  HRYRSLDK---LRTYLLSHAA----------------SK--------
Bot cin  169  HRFRALEK---LKTWLQEEA------------------------
Neu cra  179  HRAKALAKXXXLQEWFAKEMTA------------------------
Asp nig  169  HRYKALVK---LQDWLADRRP------------------------
Ust may  180  HRYKALTXXXLLQDYLVGLSKQN----------------------
Sac cer  180  HRGKAFAQ---FKEYLYQNDF------------------------
Met jan  169  HRKKAFEE---FKKFLLDRI-------------------------
Bac sub  190  HRADALKK---LSKLLEA--------------------------
Dei rad  195  HRGQALAA---LLAAHGA--------------------------
Esc col  182  HRGQALKL---LLDALRNG-------------------------
```

Figure S1. Comparative sequence analysis of inosine triphosphate pyrophosphatases from several phylogenetically distant model organisms of all kingdoms of life.

Multiple alignment of sequences of inosine triphosphate pyrophosphatases (ITPA) from *Arabidopsis thaliana* and phylogenetically distant model organisms mentioned in the materials and methods. The alignment was generated using the MUSCLE algorithm in MEGAX. Shading indicating conserved residues was generated using BOXSHADE. Abbreviations and loci used for the phylogenetic analysis are as follows (Phytozome V12.1 locus identifiers for plant and algae sequences and NCBI locus identifiers for non-plant sequences are given): *Asp nig*, *Aspegillus niger* (XP_001398459.1); *Ara tha*, *Arabidopsis thaliana* (At4g13720); *Bac sub*, *Baccillus subitilis* (OIS57636.1); *Bot cin*, *Botrytis cinerea* (XP_024547817.1); *Cae ele*, *Caenorhabditis elegans* (NP_498121.1); *Chl rei*, *Chlamydomonas reinhardtii* (Cre02g095089.t1.1); *Dan rer*, *Danio rerio* (NP_001093456.1); *Dei rad*, *Deinococcus radiodurans* (WP_010886825.1); *Dro mel*, *Drosophila melanogaster* (NP_608890.1); *Esc col*, *Escherichia coli* (NAA10980.1); *Gal gal*, *Gallus gallus* (NP_001258859.1); *Gly max*, *Glycine max* (Glyma10G125100.1); *Hom sap*, *Homo sapiens* (AAK21848.1); *Mar pol*, *Marchantia polymopha* (Mapoly0026s0018.1); *Met jan*, *Methanocaldococcus jannaschii* (WP_064496428.1); *Med tru*, *Medicago truncatula* (Medtr1g047390.1); *Neu cra*, *Neurospora crassa* (XP_955963.1); *Ory sat*, *Oryza sativa* (Os10g31940.1); *Phy pat*, *Physcomitrium patens* (Pp3c16_19640V3.1); *Pha vul*, *Phaseolus vulgaris* (Phvul.007G236950.1); *Sac cer*, *Saccharomyces cerevisiae* (AJR66682.1); *Sel moe*, *Selangiella moellendorffii* (e_gw1.0.2702.1); *Sol lyc*, *Solanum lycopersicum* (Solyc09g091420.2.1); *Sol tub*, *Solanum tuberosum* (PGSC0003DMT400076168); *Sus scr*, *Sus scrofa* (XP_020933236.1); *Tri aes*, *Triticum aestivum* (KAF6994464); *Ust may*, *Ustilago maydis* (XP_011387918.1); *Vol car*, *Volvox carteri* (Vocar0027s0182.1); *Xen lae*, *Xenopus laevis* (XP_018114098.1); *Zea may*, *Zea mays* (GRMZM2G020295_T02).
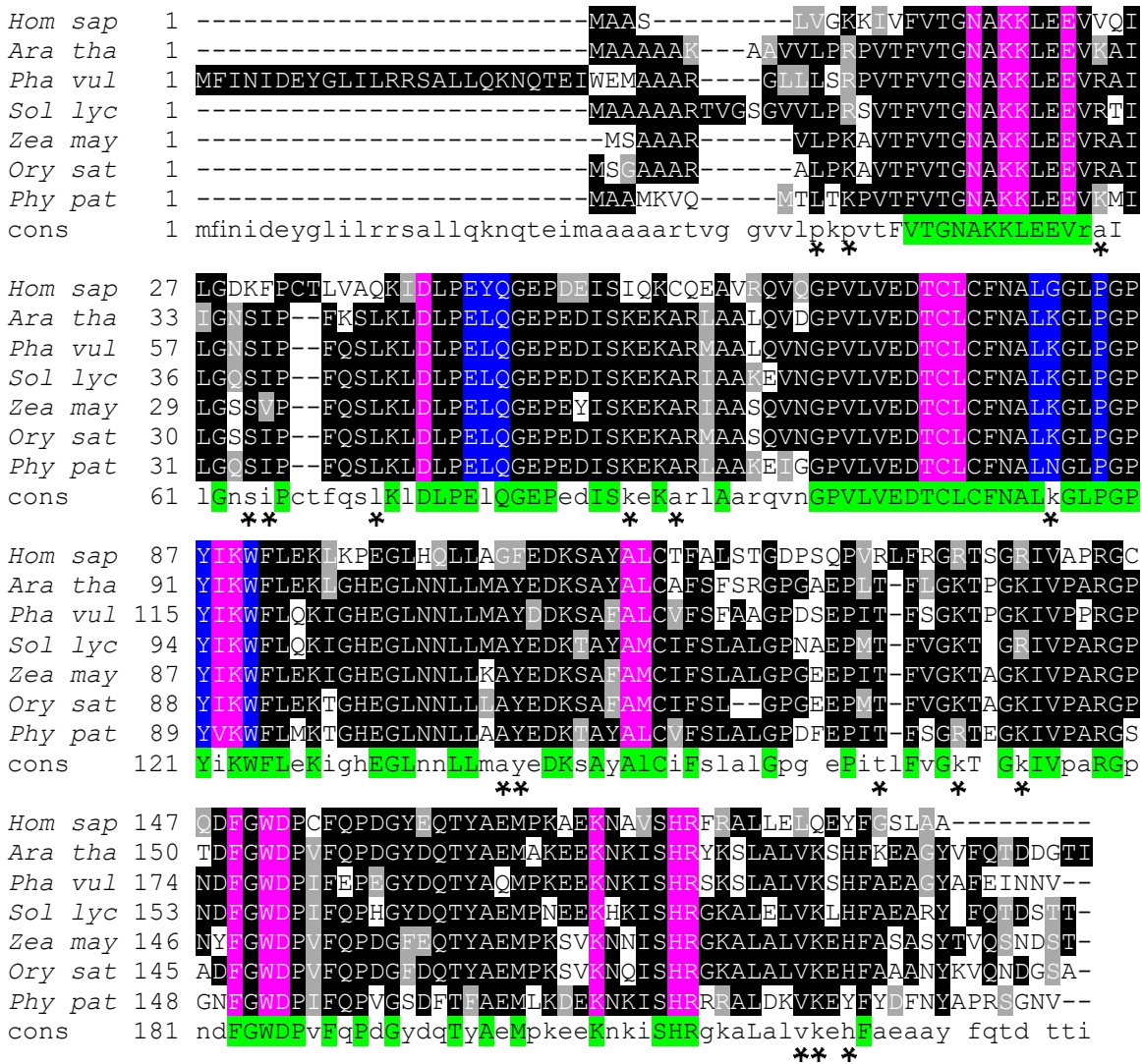
Figure S2. Comparative sequence analysis of inosine triphosphate pyrophosphatases.

Multiple alignment of sequences of inosine triphosphate pyrophosphatases (ITPA) from Homo sapiens and phylogenetically distant model plant organisms mentioned in the materials and methods. The alignment was generated using the MUSCLE algorithm in MEGAX. Shading indicating conserved residues was generated using BOXSHADE. Amino acids participating in substrate specificity and catalytic activity are indicated in pink, residues that are participating in enzyme dimerization are indicated in blue. Information about conserved amino acids is derived from Stenmark et al. (2006) and the Conserved Domain Database (Lu et al., 2020). Highly conserved amino acids occurring in all analyzed sequences are indicated in green in a consensus sequence below the analyzed sequences. Abbreviations and loci used for the phylogenetic analysis are as follows (Phytozome V12.1 locus identifiers for plant and algae sequences and NCBI locus identifiers for non-plant sequences are given): *Ara tha*, *Arabidopsis thaliana* (At4g13720); *Hom sap*, *Homo sapiens* (AAK21848.1); *Ory sat*, *Oryza sativa* (Os10g31940.1); *Phy pat*, *Physcomitrium patens* (Pp3c16_19640V3.1); *Pha vul*, *Phaseolus vulgaris* (Phvul.007G236950.1); *Sol lyc*, *Solanum lycopersicum* (Solyc09g091420.2.1); *Zea may*, *Zea mays* (GRMZM2G020295_T02). Some amino acid residues are specifically conserved in plant ITPAs, but not in the enzymes from other organisms and are indicated by an asterisk under the alignment (the given position refers to the ITPA sequence of A. thaliana): L12, R/K14, R/K30, S37, I/V38, L/V42, K58, A61, K85, A110, Y111, T134, K/R138, K142, V188, K/R189, H191.
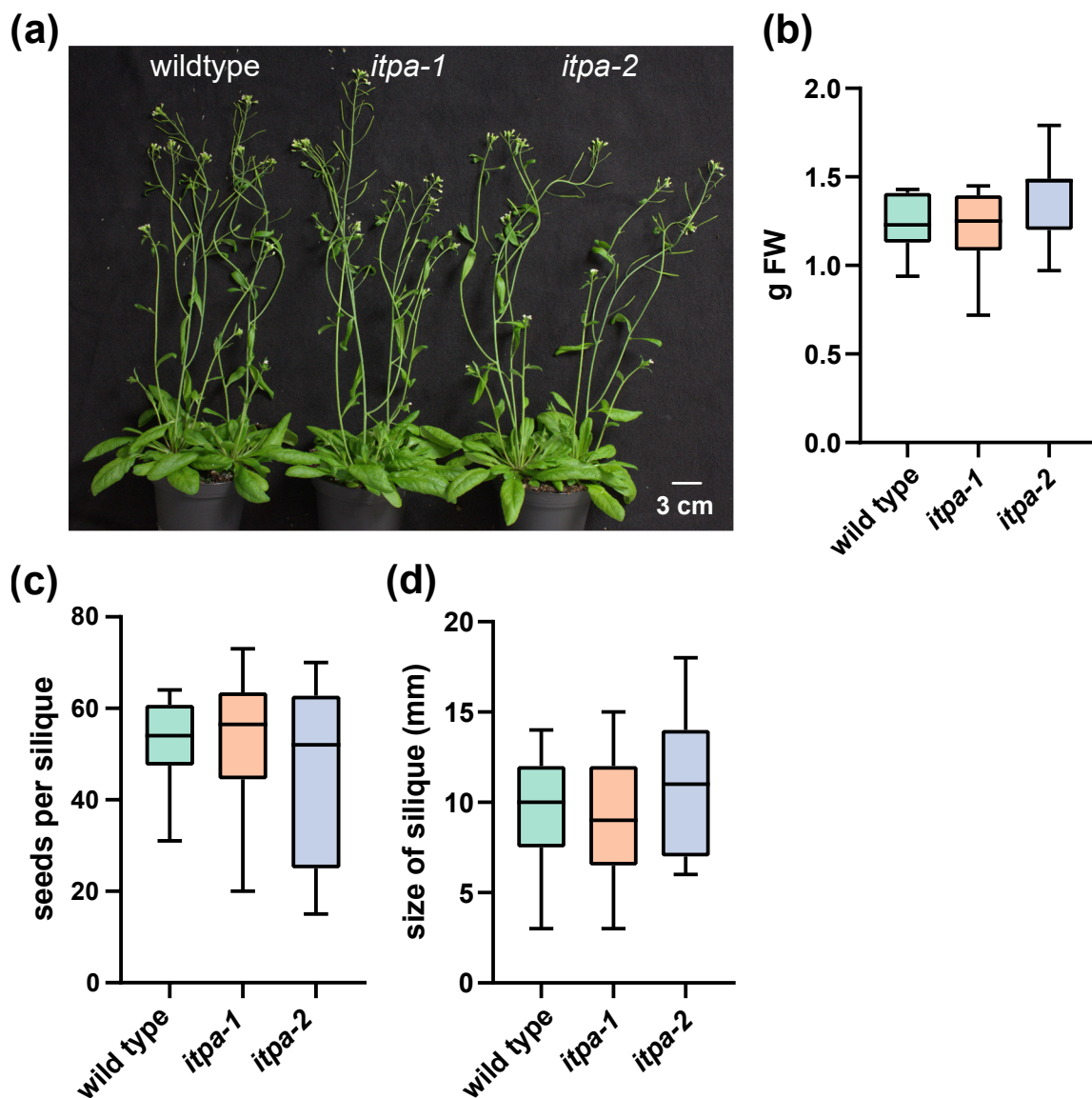
Figure S3. Phenotyping of wild type and *itpa* mutant plants

(a) Image of 35-day-old wild type, *itpa-1* and *itpa-2* plants
(b) Freshweight of 40-day-old wild type-, *itpa-1* and *itpa-2* plants. Rosettes and inflorescences were weighed without roots. Error bars are SD, *n* = 9.
(c) Seeds per silique of wild type, *itpa-1* and *itpa-2* plants. Error bars are SD, *n* = 12.
(d) Size of siliques of wild type, *itpa-1* and *itpa-2* plants. Error bars are SD, *n* = 21.

Statistical analysis was performed by a two-sided Tukey's pairwise comparision using the sandwich variance estimator. Different letters indicate *p* values < 0.05. All *p* values can be found in Table S4.
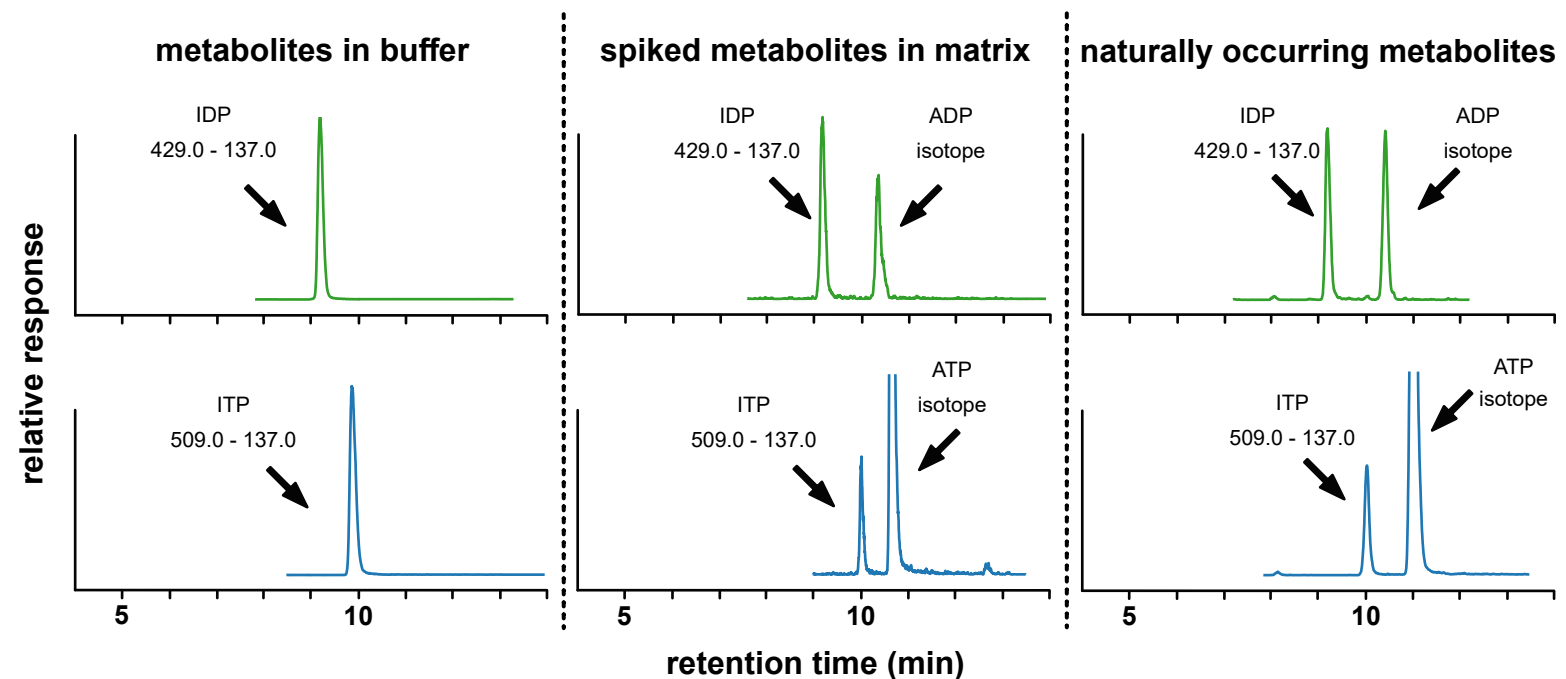
Figure S4. Chromatograms of ITP and IDP analyzed by hypercarb chromatography.

Representative depiction of Hypercarb multiple reaction monitoring (MRM) chromatograms of IDP (upper panel) and ITP (lower panel) in buffer (left column), wild type plant matrix isolated by SPE (middle column) or naturally occuring in *itpa* mutant plant matrix (right column). Chromatograms shown in green represent the IDP signal, while chromatograms shown in blue represent the ITP signal. MRM transitions used for quantification are given for the shown analytes. The signals of baseline separated ATP and ADP isotopes that are isobaric with ITP and IDP respectively are indicated by arrows.
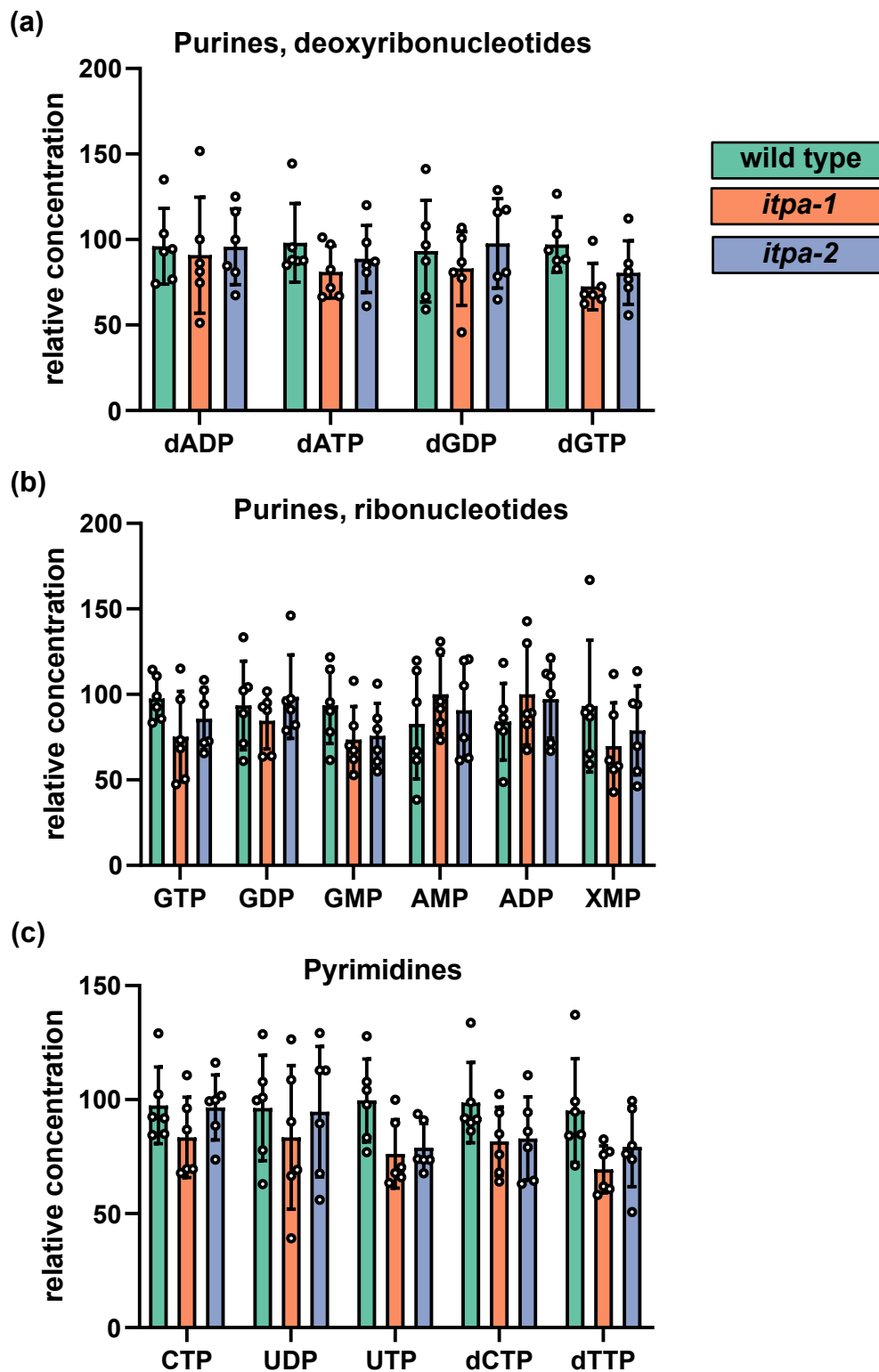
Figure S5. Relative concentration of different nucleotides in rosette leaves of 35-day-old wild type and itpa mutant plants.

Bars indicate the relative concentration of the respective metabolite in relation to the highest measured signal of the respective metabolite.

(a) Purine deoxynucleotides

(b) Purine ribonucleotides

(c) Pyridimiden nucleotides

Error bars are SD, n = 6, each replicate representing a different plant.
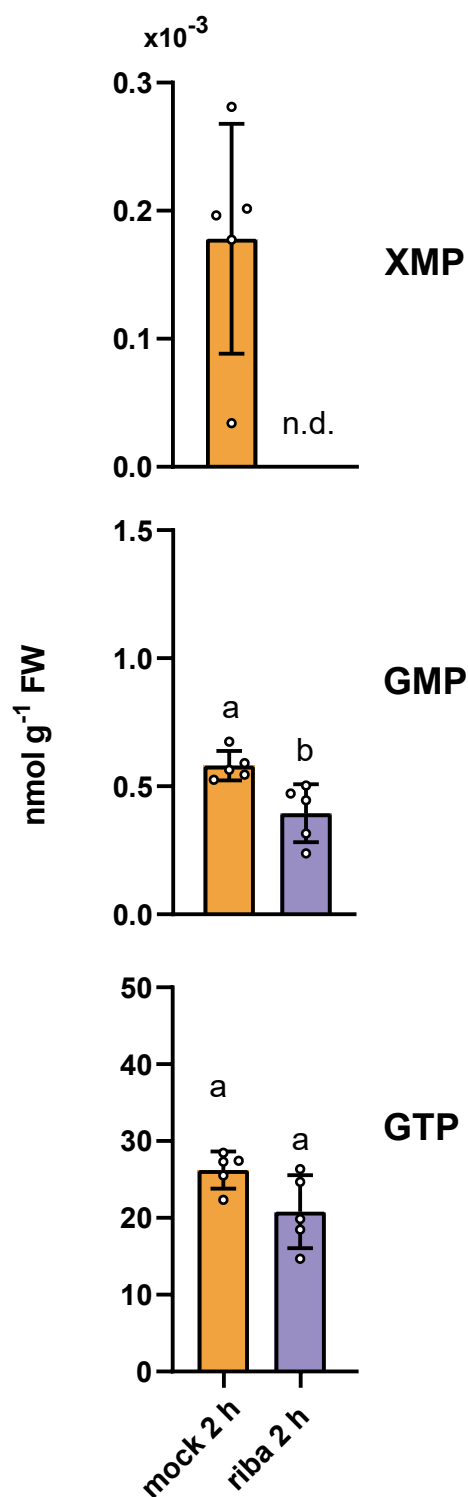
Figure S6. Manipulation of the IMP content in leaves of *N. benthamiana* alters the concentration of GMP and XMP, and in tendency also GTP.

XMP (upper panel), GMP (middel panel) and GTP (lower panel) in leaves of 21-day-old *N. benthamiana* plants.
The leaves were either mock infitrated with buffer on one leaf or infiltrated with buffer containing ribavirin (500 µM) on the opposite leaf. Error bars are SD, $n = 5$, every replicate representing a leaf from an independent plant. Statistical analysis was performed by a two-sided Tukey's pairwise comparision using the sandwich variance estimator. Different letters indicate *p* values < 0.05. All *p* values can be found in the Supplemental Table 1. n.d. = not detected.
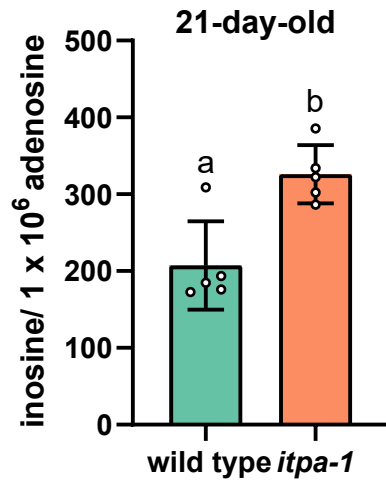
21-day-old

Figure S7. Concentration of inosine in isolated RNA of wild type and *itpa-1* plants.

Molecules of inosine per million adenosine molecules in RNA isolated from 21-day-old arabidopsis

rosette leaves. Error bars are SD, *n* = 5, every replicate representing a different plants. Statistical analysis was performed by a two-sided Tukey's pairwise comparision using the sandwich variance estimator. Different letters indicate *p* values < 0.05. All *p v*alues can be found in the Table S4.
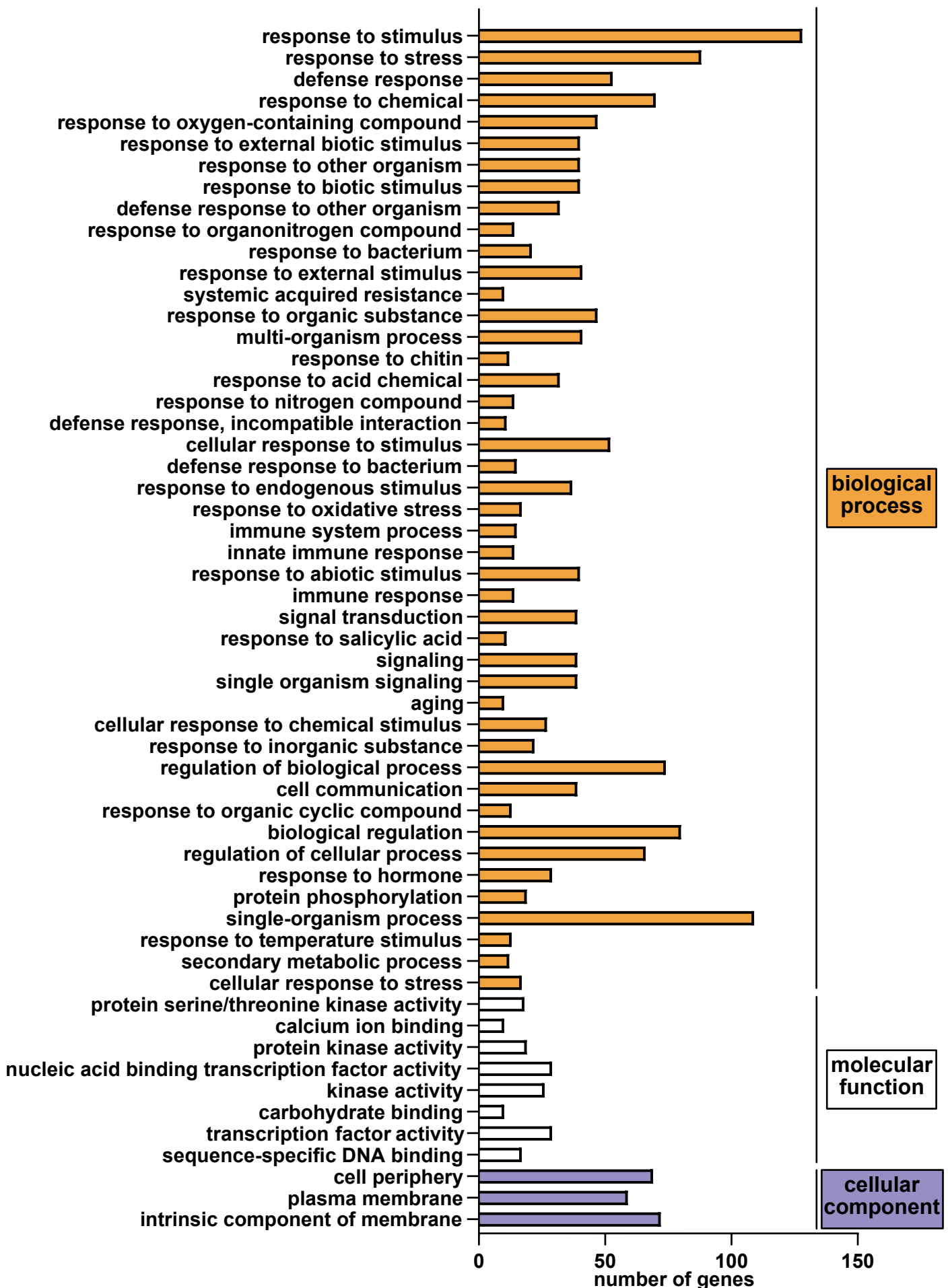
Figure S8. Loss-of-function of ITPA leads to differential upregulation of transcripts associated with GO:terms of biological processes, molecular functions and cellular components.

RNA was isolated from 35-day-old Arabidopsis rosettes. GO:term analysis of differentially upregulated genes ($\log_2$ FC > 2; FDR < 0.05) in *itpa-1* plants compared to wildtype plants. The analysis was performed with AgriGO v2.0 based on singular enrichment analysis. Statistics were performed using hypergeometric test after Hochberg with significant level adjusted P < 0.01. Only GO:terms with a minimum of 10 mapping entries are displayed.

SUPPORTING METHODS

Genotyping and cloning

Arabidopsis T-DNA insertion mutants (*itpa-1*) and respective wild type plants were identified by PCR using the following primer combinations: P1628 and P1629 (wild type reaction); P1628 and P1316 (T-DNA reaction). The PCR product of the T-DNA reaction was sent for sequencing. The T-DNA insertion was found to be positioned as described on the Signal website (http://signal.salk.edu/cgi-bin/tdnaexpress). Gene-specific mRNA level were determined by extracting total RNAs from seedlings of wild type and knockout plants and subsequent reverse transcription (Revert Aid H Minus first strand cDNA synthesis kit, ThermoFisher). PCRs were facilitated using T-DNA insertion flanking primers P2487 and P2488. An amplification of *actin2* was used as control with the primers 1033 and 1034. The PCR was carried out with 35 cycles using an annealing temperature of 55°C. Vectors for CRISPR-induced editing of the ITPA encoding locus At4g13720 were constructed as previously described (Rinne et al., 2021). We used two gRNAs (target sequences: TGAGCTCCAAGGTGAGCCTG; CTAGAGAAGCTTGGTCACGA), each expressed by an AtU6-26-promoter on the same vector. The final vector contained (I) Selectable phosphinotricin resistance cassette encoding a phosphinotricin acetyltransferase; (II) Cas9 under the controle of the EC1.2-promoter for egg cell-specific Cas9-expression (Wang et al., 2015); (III) the gRNA-expression cassette and (IV) green fluorescent protein (GFP) expressed by seed specific At2S3-promoter (Aliaga-Franco et al., 2019) for selection of non-glowing T-DNA-free seeds. The vector was transformed into *Agrobacterium tumefaciens* strain AGL-1. *Arabidopsis thaliana* Col-0 plants were grown as described above until flowers developed and transformed by floral dip (Clough and Bent, 1998). T1 seeds were harvested and screened for transgenicity by spraying 1:1000 diluted phosphinotricin (BASTA®, Bayer). Resistant plants were grown until seeds developed and T2-seedlings were genotyped for mutations in the ITPA locus. Genotyping was performed via amplified fragment length polymorphism (AFLP) analysis of the genomic area around the targeted sites as described previously (Rinne et al., 2021). The first target site located in the second exon was amplified with P1774 and P1775 and the second target site located in the fifth exon was amplified with P1776 and P1777. Mutations were confirmed by Sanger sequencing (Microsynth-Seqlab). Plants harboring mutations in the ITPA locus were grown further until seeds developed. To obtain T-DNA-free T3 seeds, non-glowing seeds were selected and germinated. Absence of the

T-DNA was confirmed by PCR-amplification with P1164 and P1165 and inheritance of mutations in the ITPA locus were confirmed by Sanger sequencing. Plants that showed no DNA-amplification with T-DNA-specific oligonucleotides were used for further experiments.

The ITPA coding region was amplified from cDNA using the primers P1075 and P1076. The fragment was cloned into pXCScpmv-HA-Strep (V69, Claus-Peter Witte et al., 2004) for protein production and into pXCS-YFP (V36, Myrach et al., 2017; Dahncke and Witte, 2013) for localization studies.

Enzymatic assays

The enzymatic activity of ITPA was measured at 360 nm using an UV-2700 UV-Vis-spectrophotometer (Shimadzu). Resulting data were analyzed with Graph Pad Prism 4, determining kinetic constants fitting the data to the Michaelis-Menten equation. Specific activities were screened with 200 µM of the respective substrate, enzymatic constants for ITP were determined using concentrations of 1.5, 3.0, 6.25, 12.5, 25.0 and 50.0 µM ITP; for dITP concentrations of 1, 3, 6, 15, and 30 µM dITP were used.

AMPK3 and 4 activity was determined using a coupled enzyme assay with pyruvate kinase and lactate dehydrogenase (LDH, Sigma, P0294), measuring the production of NAD^+ that results in absorption-decrease at 340 nm (AMPK3) or 1 µl purified enzyme (AMPK4) were used in a total volume of 300 µl containing 40 mM Tris-SO_4, 20 mM KCL, 4 mM $MgCl_2$, 2 mM phosphoenolpyruvate, 0.17 mM NADH, 2 µl 1:50 diluted coupling enzyme solution (600-1000 U/ml PK and 900-1400 U/ml LDH). The reaction was started by adding ATP to a final concentration of 1 mM and measured by a microplate spectrophotometer (Multiskan GO, Thermo Scientific) at 22°C for 20 min. The absorption change was calculated from the linear phase and subtracted by the change in a blank-sample, containing everything but the substrate. Absolute amounts were calculated utilizing the Beer-Lambert law.

Quantitative determination of (deoxy)inosine per (deoxy)adenosine level in total RNA and DNA

Briefly, 1 g plant material was frozen in liquid nitrogen and ground in a mortar with 50 mg polyvinylpolypyrrolidone. 1.5 ml CTAB Buffer (2% CTAB; 20 mM EDTA; 1.4 M NaCl; 150 mM Tris-HCL pH 7.5) were added and the material was thawed with occasional gentle stirring, transferred to 2 x 2 ml reaction tubes and placed in a water bath for 10 min at 60°C. 0.5 ml chloroform:isoamylacohol (24:1) was then added and

the tubes were inverted 10 times. The tubes were centrifuged at room temperature for 5 min at 21000 x *g* and the supernatant was transferred to a new 2 ml tube containing 0.5 ml phenol:chloroform:isoamylacohol (25:24:1). This solution was inverted thoroughly and centrifuged as stated before. The phenol-step was repeated and the resulting supernatant was transferred to a tube containing 0.5 ml chloroform, inverted and centrifuged for 3 min at room temperature at 21000 x *g*. The supernatant was then pooled in a 2 ml reaction tube and 0.6 times the volume of isopropanol was added and centrifuged at 21000 x *g* for 2 min to precipitate the nucleic acids. The supernatant was discarded and the pellet was resolved in 600 µl TE buffer (5 mM, pH 7.5) containing 50 µg/ml RNAse A and incubated at 60 °C for 10 min in a water bath. Precipitates were removed by centrifugation at 21.000 x *g* for 3 min. 60 µl 3 M sodium acetate (pH 5.2) and 1.2 ml ethanol were added to the supernatant and the supernatant was spun down at 21000 x *g* for 5 min to precipitate the DNA. The pellet was washed with 70% ethanol and resuspended in 100 µl Tris buffer (5 mM). Nucleic acids concentrations were determined using a spectrophotometer (Nanophotometer, Implen).

A Poroshell 120 SB-C18 (2.7 µm, 2.1 × 150 mm; Agilent Technologies) was employed for chromatographic separation using a gradient employing solvent A (0.0075% formic acid in water) and solvent B (0.0075% formic acid in methanol) (0-3min min, 0% solvent B; 3.0-8.0 min 11% solvent B; 8.0-8.25 min, 50% solvent B; 8.25-10.25 min, 50% solvent B; 10.25-10.50 min, 0% solvent B until 13.5 min). A flowrate of 0.35 ml/min was applied and 10 µl digested RNA were injected. The source parameters were set as following: Positive mode, gas temperature 80°C, gas flow 13 L/min, nebulizer 30 psi, sheath gas temperature 275°C, sheath gas flow 11 L/min, capillary voltage 2500 V and nozzle voltage 500 V.

Sequence Analysis and Phylogenetic Analysis

Alignments were generated using the Muscle algorithm in MEGAX. The alignment was used to construct a maximum likelihood tree using MEGAX (Kumar et al., 2018). Initial tree(s) were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value (-5225.0). All positions with less than 95% coverage were eliminated, resulting in 173 positions in the final dataset. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.9102)). Numbers at branches indicate the percentage

of trees in which associated taxa clustered together and were calculated by a bootstrap analysis (1000 bootstraps). The alignment used to construct the phylogenetic tree can be found in Fig. S1. Conserved residue shading was performed using Boxshade (http://www.ch.embnet.org/software/BOX_form.html). Conserved domains were identified using information from crystallography studies of the human ITPA (Stenmark et al., 2007), as well as bioinformatics analysis using the Conserved Domain Database (Lu et al., 2020). Abbreviations and loci used for the phylogenetic analysis are as follows (Phytozome V12.1 locus identifiers for plant and algae sequences and NCBI locus identifiers for non-plant sequences are given): *Asp nig*, *Aspegillus niger* (XP_001398459.1); *Ara tha*, *Arabidopsis thaliana* (At4g13720); *Bac sub*, *Baccillus subitilis* (OIS57636.1); *Bot cin*, *Botrytis cinerea* (XP_024547817.1); *Cae ele*, *Caenorhabditis elegans* (NP_498121.1); *Chl rei*, *Chlamydomonas reinhardtii* (Cre02g095089.t1.1); *Dan rer*, *Danio rerio* (NP_001093456.1); *Dei rad*, *Deinococcus radiodurans* (WP_010886825.1); *Dro mel*, *Drosophila melanogaster* (NP_608890.1); *Esc col*, *Escherichia coli* (NAA10980.1); *Gal gal*, *Gallus gallus* (NP_001258859.1); *Gly max*, *Glycine max* (Glyma10G125100.1); *Hom sap*, *Homo sapiens* (AAK21848.1); *Mar pol*, *Marchantia polymopha* (Mapoly0026s0018.1); *Met jan*, *Methanocaldococcus jannaschii* (WP_064496428.1); *Med tru*, *Medicago truncatula* (Medtr1g047390.1); *Neu cra*, *Neurospora crassa* (XP_955963.1); *Ory sat*, *Oryza sativa* (Os10g31940.1); *Phy pat*, *Physcomitrium patens* (Pp3c16_19640V3.1); *Pha vul*, *Phaseolus vulgaris* (Phvul.007G236950.1); *Sac cer*, *Saccharomyces cerevisiae* (AJR66682.1); *Sel moe*, *Selangiella moellendorffii* (e_gw1.0.2702.1); *Sol lyc*, *Solanum lycopersicum* (Solyc09g091420.2.1); *Sol tub*, *Solanum tuberosum* (PGSC0003DMT400076168); *Sus scr*, *Sus scrofa* (XP_020933236.1); *Tri aes*, *Triticum aestivum* (KAF6994464); *Ust may*, *Ustilago maydis* (XP_011387918.1); *Vol car*, *Volvox carteri* (Vocar0027s0182.1); *Xen lae*, *Xenopus laevis* (XP_018114098.1); *Zea may*, *Zea mays* (GRMZM2G020295_T02).

SUPPORTING METHODS REFERENCES

Aliaga-Franco, N., Zhang, C., Presa, S., Srivastava, A.K., Granell, A., Alabadí, D., Sadanandom, A., Blázquez, M.A., and Minguet, E.G. (2019). Identification of transgene-free CRISPR-edited plants of rice, tomato, and arabidopsis by monitoring DsRED fluorescence in dry seeds. Frontiers in Plant Science **10**:

1150.Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant Journal for Cell and Molecular Biology **16** (6): 735–743.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant Journal for Cell and Molecular Biology **16** (6): 735–743.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. Molecular Biology and Evolution **35** (6): 1547–1549.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C.J., and Marchler-Bauer, A. (2020). CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Research **48** (D1): D265-D268.

Rinne, J., Witte, C.-P., and Herde, M. (2021). Loss of MAR1 Function is a Marker for Co-Selection of CRISPR-Induced Mutations in Plants. Frontiers in Genome Editing **3**: 723384.

Stenmark, P., Kursula, P., Flodin, S., Gräslund, S., Landry, R., Nordlund, P., and Schüler, H. (2007). Crystal structure of human inosine triphosphatase. Substrate binding and implication of the inosine triphosphatase deficiency mutation P32T. Journal of Biological Chemistry **282** (5): 3182–3187.

Wang, Z.-P., Xing, H.-L., Dong, L., Zhang, H.-Y., Han, C.-Y., Wang, X.-C., and Chen, Q.-J. (2015). Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biology **16**: 144.

3. Summarizing discussion and future perspectives

Summary

Many aspects of nucleotide metabolism in plants are not yet well understood, partially owing to difficulties in the analysis of nucleotides from plant material as plants have a complex metabolome and contain many phosphatases (Straube et al., 2021b). The development of a method for the sensitive and comprehensive quantification of nucleotides in various plants provided a deeper insight into plant nucleotide metabolism (Straube et al., 2021a; Straube and Herde, 2022). It could be shown that the ratios of deoxyribonucleotide triphosphates (dNTPs) in plants and algae are correlated with the GC content of the respective genomes and that enzymes involved in the degradation of purine and pyrimidine degradation not only catabolize ribonucleosides but also deoxyribonucleosides, providing the first evidence for the degradation of deoxyribonucleotides *in planta*. Furthermore, it was demonstrated that dNTP concentrations were increased when deoxyribonucleosides accumulated due to a genetic block of their degradation (Straube et al., 2021a). In another study, I contributed to the development of a method for the rapid purification of highly pure mitochondria, which allows the relative comparison of metabolites, for example between a mutant and a wild type plant (Niehaus et al., 2020). Using the developed tools, we analyzed the nucleotide metabolism during germination with a special emphasis on thymidilates, utilizing a comprehensive collection of mutants abolished in enzymes involved in thymidylate biosynthesis. It was shown that especially during germination, the organellar salvage of deoxythymidine is of major importance for chloroplast DNA (cpDNA) synthesis, as mutants lacking THYMIDINE KINASE 1B exhibit phenotypic abnormalities, delayed cpDNA synthesis and, intriguingly, an accumulation of all dNTPs except deoxythymidine triphosphate (dTTP, Niehaus et al., 2022). This could be due to stalling of DNA polymerization in chloroplasts, which consumes all dTTP molecules above a certain concentration, while the other dNTPs accumulate because their metabolic sink, the DNA, is not made. The method for the analysis of nucleotides also allowed the detection and quantification of rare nucleotides with concentrations as low as a few picomoles per gram fresh weight, like xanthosine monophosphate (XMP) in wild type and mutants of *XANTHOSINE MONOPHOSPHATE PHOSPHATASE* (XMPP; Heinemann et al., 2021). Besides enzymes of the canonical metabolism, we characterized an enzyme that is involved in

3. Summarizing discussion and future perspectives

the metabolite damage repair system in plants, the INOSINE TRIPHOSPHATE PYROPHOSPHATASE (ITPA). Biochemical characterizations showed that this enzyme has similar kinetic parameters to already characterized ITPases from other organisms. We successfully quantified the damaged metabolites inosine di- and triphosphate (IDP and ITP) in *ITPA* mutants, which are even less abundant than XMP (Straube et al., 2022). Subsequently, we could show that the concentrations of IDP and ITP were further increased by abiotic stress factors. Treating wild type and mutant plants with cadmium significantly increased IDP in mutant plants and ITP in wild type and mutant plants. It could be shown that these deaminated purines are incorporated by polymerases to a greater extent into nucleic acids in plants lacking ITPA, as the concentration of deoxyinosine and inosine was higher in *itpa* plants when compared to the wild type, in DNA and RNA, respectively. Interestingly, the effect of an abolished ITPA function on DNA clearly indicated that dITP is a substrate of ITPA *in vivo* although we were not able to detect dITP. Thus, sensitivity must still be increased to enable the quantification of all nucleotides that are physiologically relevant. Plants that lack ITPA also show slightly earlier senescence and a higher abundance of transcripts associated with senescence, systemic acquired resistance, and salicylic acid (SA). Analyzing phytohormones in wild type and *itpa* plants revealed that mutant plants contained higher concentrations of SA (Straube et al., 2022). It has recently been shown that a product of DNA degradation, deoxyadenosine, has a pathogen-associated signaling function (Dunken et al., 2022). We hypothesize that potential DNA damage in *itpa* plants could lead to an increase of deoxyadenosine and thus increased translation of pathogen-associated genes.

Future perspectives

Although many areas of nucleotide metabolism have been researched for a long time, several questions remain. Many biochemical functions are known to exist *in planta*, like guanosine/inosine kinase or GMP phosphatase activity, but the responsible genes remain hitherto undiscovered. This is true for nucleotide metabolism in general, but even more so for reactions involving the damage control system. If only a fraction of the proposed 5000 damaged molecules (Lerma-Ortiz et al., 2016) are real metabolites, then many unknown enzymes for their degradation and thus detoxification should exist, waiting to be discovered. Another exciting possibility is that in accordance with the adage "new wine in old wineskins", already characterized enzymes may exhibit additional activities with new substrates, as we described for the enzymes of

3. Summarizing discussion and future perspectives

ribonucleoside degradation (Straube et al., 2021a). This is especially true, if the physiological and biochemical peculiarities of an organism increase the amount of possible damaged metabolites, as a more complex metabolism likely leads to more susceptible metabolites. Recently, it has been postulated that this is the case with plants and microbes because, unlike mammals, they have the biosynthetic ability to synthesize multiple vitamins, cofactors, and amino acids. Many of these metabolites are inherently reactive, making them susceptible to damage (Niehaus and Hillmann, 2020).

A technique to identify and annotate new damaged metabolites is so-called credentialing, which can filter for artifacts and assigns an empirical formula to features obtained in an untargeted metabolomics experiment by comparison of stable isotope-labeled samples and unlabeled samples (Griffith et al., 2021). Using this approach for the detection of damaged metabolites relies on a high sensitivity of the method, as many of them are likely of low abundance in the plant cell. In a recent publication, credentialing was used to annotate metabolites from human erythrocytes that were infected with *Plasmodium falciparum*. This approach led to the discovery of novel non-canonical metabolites like 2-hydroxyglutarate, 4-phosphoerythronate, or orotidine and resulted in the characterization of previously unknown enzymes (Cobbold et al., 2021).

The detection of damaged metabolites may be accelerated by using biotic or abiotic stresses, as I showed in this thesis that stress leads to increased concentrations of at least some damaged metabolites (Straube et al., 2022). It will be interesting to see which other metabolite groups are detectable in the fractions of the developed metabolite extraction method. We already know that in the polar fraction, nucleotide sugars are detectable and showed for the first time that plants possess measurable amounts of deoxythymidine diphosphate-glucose, which potentially is a damaged metabolite derived from a promiscuous activity of UDP-GLUCOSE PYROPHOSPHORYLASE (Straube, unpublished). Furthermore, the method likely allows quantifying phosphorylated intermediates of several metabolic pathways, like thiamin, riboflavin, or isoprenoid biosynthesis. An upscaling may also allow the detection of damaged dNTPs, like dITP or 8-oxo deoxyguanosine triphosphate, which are likely to exist in plants (Yoshimura et al., 2007; Straube et al., 2022). In the semipolar fraction, we already detected 5-methyl uridine and 5-methyl cytosine (Straube, unpublished) and it will be interesting to search for other modified and damaged nucleosides like 8-oxo guanosine.

3. Summarizing discussion and future perspectives

Another interesting possibility for the detection of metabolites is selective derivatization of a certain group of metabolites, for example, nucleotides, or a selective turnover of target analytes via enzymatic reactions. Treating the nucleotides containing a fraction of the nucleotide extraction method with a nucleotide-specific phosphatase should result in nucleosides, which can be chromatographically separated by a reverse phase chromatography, while polar metabolites are not retained. This could further increase the sensitivity by reducing the matrix effect and increasing the ionization of target molecules, as nucleosides are better ionized in the source of the mass spectrometer when compared to nucleotides. Also, it limits the available space of metabolites that can be taken into account for potential candidates behind an unknown feature in a metabolic profile. These measures could lead to the detection of novel damaged nucleotides, but also novel damaged metabolites in general.

A fascinating aspect of metabolite damage is the possibility of the evolutionary adaption of plants to these metabolites. It will be interesting to uncover whether damaged metabolites can become canonical by organisms adapting to these compounds. A good example where reactions from the damage repair system likely became canonical are the specialized metabolites vicine and convicine in *Vicia faba*. The synthesis of these metabolites presumably involves an *N*-glycosidase (COG3236) that hydrolyzes the highly reactive intermediates 2,5-diamino-6-ribosylamino-4-pyrimidinone 5´-phosphate and 5-amino-6-ribosylamino-2,3-pyrimidinedione 5´-phosphate of the riboflavin synthesis to the less reactive pyrimidine intermediates 5,6-diaminouracil (DAU) and 4-hydroxy-2,5,6-triaminopyrimidine (HTP), respectively. While the metabolic fate of DAU and HTP is unknown in most plants, it has been shown that in *V. faba* these metabolites are likely utilized for the synthesis of vicine and convicine (Björnsdotter et al., 2021). This is potentially the first evidence for a concept, where plants utilized intermediates of the damage control system for the biosynthesis of specialized metabolites (Kreis and Munkert, 2019). Further elucidation of the damage repair system in plants may lead to the discovery of more of these examples. Besides becoming canonical metabolites, damaged metabolites may also serve as signaling molecules, which has been shown for example in mammals (Boldogh et al., 2012). Furthermore, for many intermediates of the damage control system, like DAU, HTP, or 8-oxo guanosine monophosphate, the metabolic fate is unknown. It will be interesting to further elucidate how these intermediates are connected to canonical metabolism.

3. Summarizing discussion and future perspectives

Regarding deoxyribonucleotide metabolism, several questions remain to be answered in the future. It will be interesting to see why chloroplast DNA apparently suffers the most when faced with abnormal concentrations or ratios of deoxyribonucleotide triphosphates (dNTPs). A prime candidate for this research will be the homolog of the mammalian dNTP triphosphatase SAMHD1 in plants, *VENOSA4* (Yoshida et al., 2018; Xu et al., 2020). Mutants in *VENOSA4* have been phenotypically characterized and display altered chloroplast size and photosynthetic traits. However, the underlying molecular and biochemical mechanisms of a mutation in this gene remain elusive. If this gene has a similar function to the mammalian homolog, a null mutation should lead to increased concentrations of dNTPs. It is known that reduced concentrations of dNTPs impair the development of chloroplasts, but such an effect is not described for increased concentrations of dNTPs (Wang and Liu, 2006; Garton et al., 2007; Xu et al., 2015; Niehaus et al., 2022). Such a result would be indicative of a need for fine-tuned ratios of dNTPs in plants, as even slight deviations lead to detrimental effects in plants.

Indeed, we could show that the ratios of dNTPs in several different plant and algae species correlated positively with the GC content in DNA and that apparently dNTP concentrations and DNA replication are tightly interconnected (Straube et al., 2021a; Niehaus et al., 2022). Some authors suggest a metabolite-driven evolution (Lorenzo, 2014), and a determination of the composition of DNA by the ratios of dNTPs available to the polymerases (Vetsigian and Goldenfeld, 2009; Šmarda et al., 2014). The results presented in this thesis could be a hint that this is indeed the case. It is also likely that a regulation mechanism adapts dNTP synthesis to the need of the polymerase.

The comprehensive analysis of the nucleotide metabolome also enables for the first time to study of the regulation of the interplay of nucleotide biosynthesis, degradation, and salvage. This could be used to further study the influence of environmental stimuli, like cold, different light regimes, diurnal effects, or other stressors, like salt, drought, malnutrition, or pathogens, in order to identify biological contexts where nucleotides play an enhanced role or metabolites are increasingly damaged. These contexts could then be used to study the effects of mutants that under normal growth conditions show almost no phenotypical or metabolic effects. This data could also explain differences in absolute concentrations of nucleotides and nucleosides that were reported, where small deviations in sampling time, tissues harvested or growth conditions led to differences in measured concentrations.

3. Summarizing discussion and future perspectives

Summarizing, in this thesis we developed novel tools to study the metabolism in plants and described the nucleotide metabolome in unprecedented completeness. We uncovered metabolic routes for the degradation of deoxyribonucleotides, as well as a novel interplay of nucleoside degradation and nucleotide concentrations. Novel enzymes of pant nucleotide metabolism were characterized, in particular an enzyme of the metabolic damage repair system.

4. References

Alamdari, K.; Fisher, K. E.; Tano, D. W.; Rai, S.; Palos, K.; Nelson, A. D. L.; Woodson, J. D. (2021): Chloroplast quality control pathways are dependent on plastid DNA synthesis and nucleotides provided by cytidine triphosphate synthase two. In: *The New Phytologist* 231 (4), P. 1431–1448. DOI: 10.1111/nph.17467.

Ashihara, H.; Ludwig, I. A.; Crozier, A. (2020): Plant nucleotide metabolism. Biosynthesis, degradation and alkaloid formation. Chichester, Oxford: Wiley.

Atkins, C. A.; Smith, P. M. C.; Storer, P. J. (1997): Reexamination of the intracellular localization of *de novo* purine synthesis in cowpea nodules. In: *Plant Physiology* 113 (1), P. 127–135. DOI: 10.1104/pp.113.1.127.

Baccolini, C.; Witte, C.-P. (2019): AMP and GMP catabolism in arabidopsis converge on xanthosine, which is degraded by a nucleoside hydrolase heterocomplex. In: *The Plant Cell* 31 (3), P. 734–751. DOI: 10.1105/tpc.18.00899.

Bellin, L.; Del Caño-Ochoa, F.; Velázquez-Campoy, A.; Möhlmann, T.; Ramón-Maiques, S. (2021): Mechanisms of feedback inhibition and sequential firing of active sites in plant aspartate transcarbamoylase. In: *Nature Communications* 12 (1), P. 947. DOI: 10.1038/s41467-021-21165-9.

Björnsdotter, E.; Nadzieja, M.; Chang, W.; Escobar-Herrera, L.; Mancinotti, D.; Angra, D.; Xia, X.; Tacke, R.; Khazaei, H.; Crocoll, C.; Vandenberg, A.; Link, W.; Stoddard, F. L.; O´Sullivan, D. M.; Stougaard, J.; Schulman, A. H.; Andersen, S. U.; Geu-Flores, F.(2021): VC1 catalyses a key step in the biosynthesis of vicine in faba bean. In: *Nature Plants* 7 (7), P. 923–931. DOI: 10.1038/s41477-021-00950-w.

Boldogh, I.; Hajas, G.; Aguilera-Aguirre, L.; Hegde, M. L.; Radak, Z.; Bacsi, A.; Sur, S.; Hazra, T. K.; Mitra, S. (2012): Activation of ras signaling pathway by 8-oxoguanine DNA glycosylase bound to its excision product, 8-oxoguanine. In: *The Journal of Biological Chemistry* 287 (25), P. 20769–20773. DOI: 10.1074/jbc.C112.364620.

Brychkova, G.; Alikulov, Z.; Fluhr, R.; Sagi, M. (2008): A critical role for ureides in dark and senescence-induced purine remobilization is unmasked in the Atxdh1 Arabidopsis mutant. In: *The Plant Journal* 54 (3), P. 496–509. DOI: 10.1111/j.1365-313X.2008.03440.x.

Busche, M.; Scarpin, M. R.; Hnasko, R.; Brunkard, J. O. (2021): TOR coordinates nucleotide availability with ribosome biogenesis in plants. In: *The Plant Cell* 33 (5), P. 1615–1632. DOI: 10.1093/plcell/koab043.

Chen, M.; Herde, M.; Witte, C.-P. (2016): Of the nine cytidine deaminase-like genes in Arabidopsis, eight are pseudogenes and only one is required to maintain pyrimidine homeostasis *in vivo*. In: *Plant Physiology* 171 (2), P. 799–809. DOI: 10.1104/pp.15.02031.

Clausen, A. R.; Girandon, L.; Ali, A.; Knecht, W.; Rozpedowska, E.; Sandrini, M. P. B.; Andreasson, E.; Munch-Petersen, B.; Piškur, J. (2012): Two thymidine kinases and one multisubstrate deoxyribonucleoside kinase salvage DNA precursors in Arabidopsis thaliana. In: *The FEBS Journal* 279 (20), P. 3889–3897. DOI: 10.1111/j.1742-4658.2012.08747.x.

4. References

Clausen, A. R.; Mutahir, Z.; Munch-Petersen, B.; Piškur, J. (2014): Plants salvage deoxyribonucleosides in mitochondria. In: *Nucleosides, Nucleotides & Nucleic Acids* 33 (4-6), P. 291–295. DOI: 10.1080/15257770.2013.853782.

Cobbold, S. A.; Tutor , M.; Frasse, P.; McHugh, E.; Karnthaler, M.; Creek, D. J.; John, A. O.; Tilley, L.; Ralph, S. A.; McConville, M. J. (2021): Non-canonical metabolic pathways in the malaria parasite detected by isotope-tracing metabolomics. In: *Molecular Systems Biology* 17 (4), e10023. DOI: 10.15252/msb.202010023.

Cornelius, S.; Witz, S.; Rolletschek, H.; Möhlmann, T. (2011): Pyrimidine degradation influences germination seedling growth and production of Arabidopsis seeds. In: *Journal of Experimental Botany* 62 (15), P. 5623–5632. DOI: 10.1093/jxb/err251.

Crécy-Lagard, V.; Haas, D.; Hanson, A. D. (2018): Newly-discovered enzymes that function in metabolite damage-control. In: *Current Opinion in Chemical Biology* 47, P. 101–108. DOI: 10.1016/j.cbpa.2018.09.014.

Dahncke, K.; Witte, C.-P. (2013): Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in Arabidopsis. In: *The Plant Cell* 25 (10), P. 4101–4109. DOI: 10.1105/tpc.113.117184.

D'Ari, R.; Casadesús, J. (1998): Underground metabolism. In: *Bioessays* 20 (2), P. 181–186. DOI: 10.1002/(SICI)1521-1878(199802)20:2<181::AID-BIES10>3.0.CO;2-0.

Davies, O.; Mendes, P.; Smallbone, K.; Malys, N. (2012): Characterisation of multiple substrate-specific (d)ITP/(d)XTPase and modelling of deaminated purine nucleotide metabolism. In: *BMB Reports* 45 (4), P. 259–264. DOI: 10.5483/bmbrep.2012.45.4.259.

Dixon, R. A. (2003): Phytochemistry meets genome analysis, and beyond. In: *Phytochemistry* 62, P. 815–816.

Doremus, H. D.; Jagendorf, A. T. (1985): Subcellular localization of the pathway of de novo pyrimidine nucleotide biosynthesis in pea leaves. In: *Plant Physiology* 79 (3), P. 856–861. DOI: 10.1104/pp.79.3.856.

Dubois, E.; Córdoba-Cañero, D.; Massot, S.; Siaud, N.; Gakière, Bertrand; Domenichini, S.; Guérard, F.; Roldan-Arjona, T.; Doutriaux, M.-P. (2011): Homologous recombination is stimulated by a decrease in dUTPase in Arabidopsis. In: *PloS one* 6 (4), e18658. DOI: 10.1371/journal.pone.0018658.

Dunken, N.; Widmer, H.; Balcke, G. U.; Straube, H.; Langen, G.; Charura, N. M.; Saake, P.; Leson, L.; Rövenich, H.; Wawra, S.; Djamei, A.; Tissier, A.; Witte, C.-P.; Zuccaro, A. (2022): A fungal endophyte-generated nucleoside signal regulates host cell death and promotes root colonization. *bioRxiv*, DOI: https://doi.org/10.1101/2022.03.11.483938

Faivre-Nitschke, S. E.; Grienenberger, J. M.; Gualberto, J. M. (1999): A prokaryotic-type cytidine deaminase from Arabidopsis thaliana gene expression and functional characterization. In: *European journal of biochemistry* 263 (3), P. 896–903. DOI: 10.1046/j.1432-1327.1999.00591.x.

Fang, C.; Fernie, A. R.; Luo, J. (2019): Exploring the diversity of plant metabolism. In: *Trends in Plant Science* 24 (1), P. 83–98. DOI: 10.1016/j.tplants.2018.09.006.

4. References

Galperin, M. Y.; Moroz, O. V.; Wilson, K. S.; Murzin, A. G. (2006): House cleaning, a part of good housekeeping. In: *Molecular Microbiology* 59 (1), P. 5–19. DOI: 10.1111/j.1365-2958.2005.04950.x.

Garton, S.; Knight, H.; Warren, G. J.; Knight, M. R.; Thorlby, G. J. (2007): Crinkled leaves 8 - a mutation in the large subunit of ribonucleotide reductase - leads to defects in leaf development and chloroplast division in *Arabidopsis thaliana*. In: *The Plant Journal* 50 (1), P. 118–127. DOI: 10.1111/j.1365-313X.2007.03035.x.

Griffith, C. M.; Walvekar, A. S.; Linster, C. L. (2021): Approaches for completing metabolic networks through metabolite damage and repair discovery. In: *Current Opinion in Systems Biology* 28. DOI: 10.1016/j.coisb.2021.100379.

Halimi, Y.; Dessau, M.; Pollak, S.; Ast, T.; Erez, T.; Livnat-Levanon, N.; Karniol, B.; Hirsch, J. A.; Chamovitz, D. A. (2011): COP9 signalosome subunit 7 from Arabidopsis interacts with and regulates the small subunit of ribonucleotide reductase (RNR2). In: *Plant Molecular Biology* 77 (1-2), P. 77–89. DOI: 10.1007/s11103-011-9795-8.

Hanson, A. D.; Henry, C. S.; Fiehn, O.; Crécy-Lagard, V. (2016): Metabolite damage and metabolite damage control in plants. In: *Annual Review of Plant Biology* 67, P. 131–152. DOI: 10.1146/annurev-arplant-043015-111648.

Heinemann, K. J.; Yang, S.-Y.; Straube, H.; Medina-Escobar, N.; Varbanova-Herde, M.; Herde, M.; Rhee, S.; Witte, C.-P. (2021): Initiation of cytosolic plant purine nucleotide catabolism involves a monospecific xanthosine monophosphate phosphatase. In: *Nature Communications* 12 (1), P. 6846. DOI: 10.1038/s41467-021-27152-4.

Hesberg, C.; Hänsch, R.; Mendel, R. R.; Bittner, F. (2004): Tandem orientation of duplicated xanthine dehydrogenase genes from *Arabidopsis thaliana*: differential gene expression and enzyme activities. In: *The Journal of Biological Chemistry* 279 (14), P. 13547–13554. DOI: 10.1074/jbc.M312929200.

Hickl, D.l; Scheuring, D.; Möhlmann, T. (2021): CTP Synthase 2 from *Arabidopsis thaliana* is required for complete embryo development. In: *Frontiers in Plant Science* 12, P. 652434. DOI: 10.3389/fpls.2021.652434.

Jung, B.; Hoffmann, C.; Möhlmann, T. (2011): Arabidopsis nucleoside hydrolases involved in intracellular and extracellular degradation of purines. In: *The Plant Journal* 65 (5), P. 703–711. DOI: 10.1111/j.1365-313X.2010.04455.x.

Kafer, C.; Zhou, L.; Santoso, D.; Guirgis, A.; Weers, B.; Park, S.; Thornburg, R. (2004): Regulation of pyrimidine metabolism in plants. In: *Frontiers in Bioscience* 9, S. 1611–1625. DOI: 10.2741/1349.

Keller, M. A.; Piedrafita, G.; Ralser, M. (2015): The widespread role of non-enzymatic reactions in cellular metabolism. In: *Current Opinion in Biotechnology* 34, P. 153–161. DOI: 10.1016/j.copbio.2014.12.020.

Kreis, W.; Munkert, J. (2019): Exploiting enzyme promiscuity to shape plant specialized metabolism. In: *Journal of Experimental Botany* 70 (5), P. 1435–1445. DOI: 10.1093/jxb/erz025.

4. References

Kumar, D.; Abdulovic, A. L.; Viberg, J.; Nilsson, A. K.; Kunkel, T. A.; Chabes, A. (2011): Mechanisms of mutagenesis in vivo due to imbalanced dNTP pools. In: *Nucleic Acids Research* 39 (4), P. 1360–1371. DOI: 10.1093/nar/gkq829.

Law, S. R.; Narsai, R.; Taylor, N. L.; Delannoy, E.; Carrie, C.; Giraud, E.; Millar, A. H.; Small, I.; Whelan, J. (2012): Nucleotide and RNA metabolism prime translational initiation in the earliest events of mitochondrial biogenesis during Arabidopsis germination. In: *Plant physiology* 158 (4), P. 1610–1627. DOI: 10.1104/pp.111.192351.

Lerma-Ortiz, C.; Jeffryes, J. G.; Cooper, A. J. L.; Niehaus, T. D.; Thamm, A. M. K.; Frelin, O.; Aunins, T.; Fiehn, O.; Crécy-Lagard, V.; Henry, C. S.; Hanson, A. D. (2016): 'Nothing of chemistry disappears in biology': the Top 30 damage-prone endogenous metabolites. In: *Biochemical Society Transactions* 44 (3), P. 961–971. DOI: 10.1042/BST20160073.

Linster, C. L.; van Schaftingen, E.; Hanson, A. D. (2013): Metabolite damage and its repair or pre-emption. In: *Nature Chemical Biology* 9 (2), P. 72–80. DOI: 10.1038/nchembio.1141.

Lorenzo, V. (2014): From the selfish gene to selfish metabolism: revisiting the central dogma. In: *Bioessays* 36 (3), P. 226–235. DOI: 10.1002/bies.201300153.

Ma, X.; Wang, W.; Bittner, Fl.; Schmidt, N.; Berkey, R.; Zhang, L.; King, H.; Zhang, Y.; Feng, J.; Wen, Y.; Tan, L.; Li, Y.; Zhang, Q.; Deng, Z.; Xiong, X.; Xiao, S. (2016): Dual and Opposing Roles of Xanthine Dehydrogenase in Defense-Associated Reactive Oxygen Species Metabolism in Arabidopsis. In: *The Plant Cell* 28 (5), P. 1108–1126. DOI: 10.1105/tpc.15.00880.

Narsai, R.; Law, S. R.; Carrie, C.; Xu, L.; Whelan, J. (2011): In-depth temporal transcriptome profiling reveals a crucial developmental switch with roles for RNA processing and organelle metabolism that are essential for germination in Arabidopsis. In: *Plant Physiology* 157 (3), P. 1342–1362. DOI: 10.1104/pp.111.183129.

Niehaus, M.; Straube, H.; Künzler, P.; Rugen, N.; Hegermann, J.; Giavalisco, P.; Eubel H.; Witte, C.-P., Herde, M. (2020): Rapid Affinity Purification of Tagged Plant Mitochondria (Mito-AP) for Metabolome and Proteome Analyses. In: *Plant Physiology* 182 (3), P. 1194–1210. DOI: 10.1104/pp.19.00736.

Niehaus, M.; Straube, H.; Specht, A.; Baccolini, C.; Witte, C.-P.; Herde, M. (2022): The nucleotide metabolome of germinating Arabidopsis thaliana reveals a central role of thymidine for chloroplast development. *Submitted to The Plant Cell.*

Niehaus, T. D.; Hillmann, K. B. (2020): Enzyme promiscuity, metabolite damage, and metabolite damage control systems of the tricarboxylic acid cycle. In: *The FEBS Journal* 287 (7), P. 1343–1358. DOI: 10.1111/febs.15284.

Niu, M.; Wang, Y.; Wang, C.; Lyu, J.; Wang, Y.; Dong, H.; Long, W.; Wang, D.; Kong, W.; Wang, L.; Gua, X.; Sun, L.; Hu, T.; Zhai, H.; Wang, H.; Wan, J. (2017): ALR encoding dCMP deaminase is critical for DNA damage repair, cell cycle progression and plant development in rice. In: *Journal of Experimental Botany* 68 (21-22), P. 5773–5786. DOI: 10.1093/jxb/erx380.

Pedroza-García, J.-A.; Nájera-Martínez, M.; Mazubert, C.; Aguilera-Alvarado, P.; Drouin-Wahbi, J.; Sánchez-Nieto, S.; Gualberto, J. M.; Raynaud, C.; Plasencia, J.

4. References

(2019): Role of pyrimidine salvage pathway in the maintenance of organellar and nuclear genome integrity. In: *The Plant Journal* 97 (3), P. 430–446. DOI: 10.1111/tpj.14128.

Poli, J.; Tsaponina, O.; Crabbé, L.; Keszthelyi, A.; Pantesco, V.; Chabes, A.; Lengronne, A.; Pasero, P. (2012): dNTP pools determine fork progression and origin usage under replication stress. In: *The EMBO Journal* 31 (4), P. 883–894. DOI: 10.1038/emboj.2011.470.

Ronceret, A.; Gadea-Vacas, J.; Guilleminot, J.; Lincker, F.; Delorme, V.; Lahmy, S.; Pelletier, G.; Chabouté, M.-E.; Devic, M. (2008): The first zygotic division in Arabidopsis requires de novo transcription of thymidylate kinase. In: *The Plant Journal* 53 (5), P. 776–789. DOI: 10.1111/j.1365-313X.2007.03372.x.

Sabina, R. L.; Paul, A.-L.; Ferl, R. J.; Laber, B.; Lindell, S. D. (2007): Adenine nucleotide pool perturbation is a metabolic trigger for AMP deaminase inhibitor-based herbicide toxicity. In: *Plant Physiology* 143 (4), P. 1752–1760. DOI: 10.1104/pp.107.096487.

Sauge-Merle, S.; Falconet, D.; Fontecave, M. (1999): An active ribonucleotide reductase from Arabidopsis thaliana cloning, expression and characterization of the large subunit. In: *European Journal of Biochemistry* 266 (1), P. 62–69. DOI: 10.1046/j.1432-1327.1999.00814.x.

Schroeder, R. Y.; Zhu, A.; Eubel, H.; Dahncke, K.; Witte, C.-P. (2018): The ribokinases of *Arabidopsis thaliana* and *Saccharomyces cerevisiae* are required for ribose recycling from nucleotide catabolism, which in plants is not essential to survive prolonged dark stress. In: *The New Phytologist* 217 (1), P. 233–244. DOI: 10.1111/nph.14782.

Šmarda, P.; Bureš, P.; Horová, L.; Leitch, I. J.; Mucina, L.; Pacini, E.; Tichý, L.; Grulich, V.; Rotreklová, O. (2014): Ecological and evolutionary significance of genomic GC content diversity in monocots. In: *Proceedings of the National Academy of Sciences of the United States of America* 111 (39), E4096-102. DOI: 10.1073/pnas.1321152111.

Smith, P. M. C.; Atkins, C. A. (2002): Purine Biosynthesis. Big in cell division, even bigger in nitrogen assimilation. In: *Plant Physiology* 128 (3), P. 793–802. DOI: 10.1104/pp.010912.

Soltabayeva, A.; Srivastava, S.; Kurmanbayeva, A.; Bekturova, A.; Fluhr, R.; Sagi, M. (2018): Early Senescence in older leaves of low nitrate-grown *Atxdh1* uncovers a role for purine catabolism in N Supply. In: *Plant Physiology* 178 (3), P. 1027–1044. DOI: 10.1104/pp.18.00795.

Straube, H.; Herde, M. (2022): Purification and analysis of Nucleotides and Nucleosides from Plants. Accepted in: *Methods in Molecular Biology*.

Straube, H.; Niehaus, M.; Zwittian, S.; Witte, C.-P.; Herde, M. (2021a): Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. In: *The Plant Cell* 33 (2), P. 270–289. DOI: 10.1093/plcell/koaa028.

Straube, H.; Straube, J.; Rinne, J.; Niehaus, M.; Rose, H.; Witte, C.-P.; Herde, M. (2022): An inosine triphosphate pyrophosphatase safeguards nucleic acids from

4. References

aberrant purine nucleotides and prevents a constitutive salicylic acid response. *bioRxiv*. DOI: https://doi.org/10.1101/2022.02.24.481826

Straube, H.; Witte, C.-P.; Herde, M. (2021b): Analysis of nucleosides and nucleotides in plants: An update on sample preparation and LC-MS techniques. In: *Cells* 10 (3). DOI: 10.3390/cells10030689.

Tang, C.; Tang, S.; Zhang, S.; Luo, M.; Jia, G.; Zhi, H.; Diao, X. (2019): SiSTL1, encoding a large subunit of ribonucleotide reductase, is crucial for plant growth, chloroplast biogenesis, and cell cycle progression in *Setaria italica*. In: *Journal of Experimental Botany* 70 (4), P. 1167–1182. DOI: 10.1093/jxb/ery429.

Tawfik, D. S. (2010): Messy biology and the origins of evolutionary innovations. In: *Nature Chemical Biology* 6 (10), P. 692–696. DOI: 10.1038/nchembio.441.

Tintemann, H.; Wasternack, C.; Benndorf, R.; Reinbothe, H. (1985): The rate-limiting step of uracil degradation in tomato cell suspension cultures and Euglena gracilis in vivo studies. In: *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* 82 (4), P. 787–792. DOI: 10.1016/0305-0491(85)90526-7.

Torrents, E. (2014): Ribonucleotide reductases: essential enzymes for bacterial life. In: *Frontiers in Cellular and Infection Microbiology* 4, P. 52. DOI: 10.3389/fcimb.2014.00052.

Urarte, E.; Esteban, R.; Moran, J. F.; Bittner, F. (2015): Established and proposed roles of Xanthine Oxidoreductase in oxidative and reductive pathways in plants. In: *Reactive Oxygen and Nitrogen Species Signaling and Communication in Plants*, Bd. 23. Cham: Springer International Publishing, P. 15–42.

van Schaftingen, E.; Rzem, R.; Marbaix, A.; Collard, F.; Veiga-da-Cunha, M.; Linster, C. L. (2013): Metabolite proofreading, a neglected aspect of intermediary metabolism. In: *Journal of Inherited Metabolic Disease* 36 (3), P. 427–434. DOI: 10.1007/s10545-012-9571-1.

Vértessy, B. G.; Tóth, J. (2009): Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. In: *Accounts of Chemical research* 42 (1), P. 97–106. DOI: 10.1021/ar800114w.

Vetsigian, K.; Goldenfeld, N. (2009): Genome rhetoric and the emergence of compositional bias. In: *Proceedings of the National Academy of Sciences of the United States of America* 106 (1), P. 215–220. DOI: 10.1073/pnas.0810122106.

Vincenzetti, S.; Cambi, A.; Neuhard, J.; Schnorr, K.; Grelloni, M.; Vita, A. (1999): Cloning, expression, and purification of cytidine deaminase from Arabidopsis thaliana. In: *Protein Expression and Purification* 15 (1), P. 8–15. DOI: 10.1006/prep.1998.0959.

Walsh, T. A.; Green, S. B.; Larrinua, I. M.; Schmitzer, P. R. (2001): Characterization of plant beta-ureidopropionase and functional overexpression in Escherichia coli. In: *Plant Physiology* 125 (2), P. 1001–1011. DOI: 10.1104/pp.125.2.1001.

Wang, C.; Liu, Z. (2006): Arabidopsis ribonucleotide reductases are critical for cell cycle progression, DNA damage repair, and plant development. In: *The Plant Cell* 18 (2), P. 350–365. DOI: 10.1105/tpc.105.037044.

4. References

Wang, S.; Alseekh, S.; Fernie, A. R.; Luo, J. (2019): The Structure and Function of Major Plant Metabolite Modifications. In: *Molecular Plant* 12 (7), P. 899–919. DOI: 10.1016/j.molp.2019.06.001.

Watanabe, S.; Kounosu, Y.; Shimada, H.; Sakamoto, A. (2014): Arabidopsis xanthine dehydrogenase mutants defective in purine degradation show a compromised protective response to drought and oxidative stress. In: *Plant Biotechnology* 31 (2), P. 173–178. DOI: 10.5511/plantbiotechnology.14.0117a.

Werner, A. K.; Witte, C.-P. (2011): The biochemistry of nitrogen mobilization: purine ring catabolism. In: *Trends in Plant Science* 16 (7), P. 381–387. DOI: 10.1016/j.tplants.2011.03.012.

Witte, C.-P.; Herde, M. (2020): Nucleotide Metabolism in Plants. In: *Plant Physiology* 182 (1), P. 63–78. DOI: 10.1104/pp.19.00955.

Xu, D.; Leister, D.; Kleine, T. (2020): VENOSA4, a human dNTPase SAMHD1 homolog, contributes to chloroplast development and abiotic stress tolerance. In: *Plant Physiology* 182 (2), P. 721–729. DOI: 10.1104/pp.19.01108.

Xu, J.; Deng, Y.; Li, Q.; Zhu, X.; He, Z. (2014): STRIPE2 encodes a putative dCMP deaminase that plays an important role in chloroplast development in rice. In: *Journal of Genetics and Genomics* 41 (10), P. 539–548. DOI: 10.1016/j.jgg.2014.05.008.

Xu, J.; Zhang, L.; Yang, D.-L.; Li, Q.; He, Z. (2015): Thymidine kinases share a conserved function for nucleotide salvage and play an essential role in Arabidopsis thaliana growth and development. In: *The New Phytologist* 208 (4), P. 1089–1103. DOI: 10.1111/nph.13530.

Yabuki, N.; Ashihara, H. (1991): Catabolism of adenine nucleotides in suspension-cultured plant cells. In: *Biochimica et Biophysica Acta* 1073 (3), P. 474–480. DOI: 10.1016/0304-4165(91)90218-6.

Yoo, S.-C.; Cho, S.-H.; Sugimoto, H.; Li, J.; Kusumi, K.; Koh, H.-J.; Iba, K.; Paek, N.-C. (2009): Rice virescent3 and stripe1 encoding the large and small subunits of ribonucleotide reductase are required for chloroplast biogenesis during early leaf development. In: *Plant Physiology* 150 (1), P. 388–401. DOI: 10.1104/pp.109.136648.

Yoshida, Y.; Sarmiento-Mañús, R.; Yamori, W.; Ponce, M. R.; Micol, J. L.; Tsukaya, H. (2018): The Arabidopsis phyB-9 mutant has a second-site mutation in the VENOSA4 gene that alters chloroplast size, photosynthetic traits, and leaf growth. In: *Plant Physiology* 178 (1), P. 3–6. DOI: 10.1104/pp.18.00764.

Yoshimura, K.; Ogawa, T.; Ueda, Y.; Shigeoka, S. (2007): AtNUDX1, an 8-oxo-7,8-dihydro-2'-deoxyguanosine 5'-triphosphate pyrophosphohydrolase, is responsible for eliminating oxidized nucleotides in Arabidopsis. In: *Plant & Cell Physiology* 48 (10), P. 1438–1449. DOI: 10.1093/pcp/pcm112.

Zrenner, R.; Stitt, M.; Sonnewald, U.; Boldt, R. (2006): Pyrimidine and purine biosynthesis and degradation in plants. In: *Annual Review of Plant Biology* 57, P. 805–836. DOI: 10.1146/annurev.arplant.57.032905.105421.

5. Appendix

5.1 *Curriculum vitae*

<div align="center">

Henryk Straube

Born 27.03.1992 in Soltau/ Germany

Adress: Am kleinen Felde 4a, 30167 Hannover

Henryk.straube@web.de

</div>

Education

2018 –2022 **Ph.D. student in Plant Biochemistry and Molecular Biology**

Department of Molecular Nutrition and Biochemistry of Plants, Leibniz University Hannover

- Project: Investigation of the metabolism of deoxyribonucleotides and damaged nucleotides in plants
- Establishment of analytical methods to monitor low abundant polar metabolites in plants
- Using complementary approaches as protein purification, determination of enzyme kinetics, genome editing and targeted and untargeted metabolite analysis to investigate the function of enzymes

2015 -2017 **Master studies in Plant Biotechnology**

Institute of Plant Genetics, Department of Plant Proteomics, Leibniz University Hannover

- Analysis of the mitochondrial proteome of heterotrophic tissues by large-pore blue-native gels and mass spectrometry

2012-2015 **Bachelor studies in Plant Biotechnology**

Institute for Plant Genetics, Department of Plant Proteomics, Leibniz University Hannover

- Analysis of the proteome of the outer mitochondrial membrane using chemical crosslinking

2004-2011 **University Entrance Degree (Abitur) at KGS Schneverdingen**

Scientific stays

2021 **Stay at the Bordeaux Metabolome Facility**

Centre INRA de Nouvelle Aquitaine – Bordeaux

- Untargeted metabolome analysis of *Arabidopsis* and *Lycopersicum* samples involving isotope-labeled material to identify novel metabolites

5. Appendix

Scholarships

2021   PROCOPE-MOBILITÄT 2021 scholarship from the Department of Science and Technology of the French Embassy in Germany

2018-2019   Exposé scholarship from the Leibniz University of Hannover

5.2 List of publications

Straube, Henryk; Straube, Jannis; Rinne, Jannis; Fischer, Lisa; Niehaus, Markus; Witte, ClausPeter; Herde, Marco (2022): An inosine triphosphate pyrophosphatase safeguards plant nucleic acids from aberrant purine nucleotides. In: New Phytologist 237, 1759-1775. DOI: 10.1111/nph.18656

Straube, Henryk; Herde, Marco (2022): Purification and analysis of Nucleotides and Nucleosides from Plants. In: Methods in Molecular Biology, New York, NY: Springer US, 2022. 145-155. DOI: 10.1007/978-1-0716-2176-9_11.

Straube, Henryk; Witte, Claus-Peter; Herde, Marco (2021): Analysis of nucleosides and nucleotides in plants: An update on sample preparation and LC-MS techniques. In: Cells, 2021. DOI: 10.3390/cells10030689

Straube, Henryk; Niehaus, Markus; Zwittian, Sarah; Witte, Claus-Peter; Herde, Marco (2020): Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plant and algae revealing connections between nucleoside and deoxynucleoside metabolism. In: The Plant Cell, 33, P. 270-289. DOI: 10.1093/plcell/koaa028

Dunken, Nick; Widmer, Heidi; Balcke, Gerd U.; **Straube, Henryk**.; Langen, Gregor; Charura, Nyasha M.; Saake, Pia; Leson, Lisa; Rövenich, Hanna; Wawra, Stephan; Djamei, Armin; Tissier, Alain; Witte, Claus-Peter; Zuccaro, Alga: A fungal endophyte-generated nucleoside signal regulates host cell death and promotes root colonization. DOI: 10.1101/2022.03.11.483938

Niehaus, Markus; **Straube, Henryk**; Specht, André; Baccolini, Chiara; Witte, Claus-Peter; Herade, Marco: The nucleotide metabolome of germinating Arabidopsis thaliana reveals a central role of thymidine for chloroplast development. In: The Plant Cell, 34, P. 3790-3813. DOI: 10.1093/plcell/koac207

Heinemann, Katharina J.; Yang, San-Young; **Straube, Henryk**; Escobar, Nieves-Medina; Varbanova-Herde, Marina; Herde, Marco; Rhee, Sangkee; Witte, Claus-Peter (2021): Initiation of cytosolic plant purine nucleotide catabolism involves a monospecific xanthosine monophosphate phosphatase. In: Nature Communications 12, 6846. DOI: 10.1038/s41467-021-27152-4

Niehaus, Markus; **Straube, Henryk**; Künzler, Patrick; Rugen, Nils; Hegermann, Jan; Giavalsico, Patrick; Eubel, Holger; Witte, Claus-Peter; Herde, Marco (2020): Rapid affinity purification of tagged plant mitochondria (Mito-AP) for metabolome and proteome analyses. In: Plant Physiology 182, P. 1194-1210. DOI: 10.1104/pp.19.00736

5. Appendix

Rugen, Nils; **Straube, Henryk**; Franken, Linda E.; Braun, Hans-Peter; Eubel, Holger (2019): Complexome profiling reveals association of PPR proteins with ribosomes in the mitochondria of plants. In: Molcular & Cellular Proteomics 18, P. 1345-1362. DOI: 10.1074/mcp.RA119.001396

5.3 Conference contribution

Poster Presentations

International Plant Science Conference (Botanikertagung)
September, 2019, Rostock (Germany)
Straube, H.; Niehaus, M.; Witte, C.-P. and Herde, M.: "A toolbox for the quantification of rare deoxynucleotide species"

Molecular Biology of Plants Conference
February, 2020, Darbinghausen (Germany)
Straube, H.; Niehaus, M.; Witte, C.-P. and Herde, M.: „On the elucidation of the deoxyribonucleotide metabolism"

6. Acknowledgements

First of all, I would like to thank Dr. Marco Herde, without whom I would never have decided to do a PhD in the department of "Plant Nutrition". I really appreciate that you always had an open ear, even when the time was short, and I certainly wouldn't know what I know today about plant biology and science in general if it hadn't been for our discussions, even when the public was already upset that we "weren't working". You always took ideas seriously and never judged. I have rarely met a person as kind and brilliant at the same time as you are.

My special thanks belong to Prof. Dr. Claus-Peter Witte, who gave me the opportunity to join the lab and supported me in everything I tried to achieve. I learned a lot from you and I would like to thank you for all the discussions we had, be it at the lunch table or about science, the hours you spent turning our stories until everybody understood what we wanted to say, and generally for enabling what is to come.

I would like to express my gratitude to the whole group, for bearing with me even if it was not always easy. Thank you, Markus Niehaus, for introducing me to molecular biology, for still talking and arguing with me; although we agree to disagree, and for the numerous times you helped me in the lab. I would like to thank Jannis Rinne for the many enjoyable evenings we had, for sometimes losing to me at table soccer, and for supporting me in science. Without you, life would have been less funny and interesting. The same holds true for Luisa Voß, thank you as well for nice evenings, playing Carcassonne or Mario Party and nice hikes in the Harz, and being definitely the person with the best eye for tasty clothes in the lab. To Xiaoguang Chen I would like to say thank you for being my office mate, discussing interesting views on politics, and introducing me to real Chinese food. I would furthermore like to thank Dr. Nieves Medina-Escobar for always having an open ear and helping me with problems whenever you could, Iris Wienkemeier for always helping with bureaucracy and documents, Hildegard Thölke for being a generally uplifting spirit, many conversations at the lunch table and for keeping an overview about lab stuff and André Specht for always jumping into the breach when a machine broke down and fixing it with patience. I also like to express my gratitude to all the people that have left the lab and that supported me in several ways to finish this thesis: Katharina Heinemann, Marina Varbanova-Herde, Adiel Lopez, Chiara Baccolini, Nabila Firdoos, and especially Lisa Fischer and Sarah Zwittian, two brilliant students who trusted me and helped tremendously finishing parts of this thesis.

I also like to mention all my friends that still beard with me over the years, regardless of my confusion, I owe you one: Alicia, Ana, Andi(s), Anil, Carsten, Ceddi, Daniel, David, Henno, HG, Holste, Janina, Jelco, Jonathan, Kevin, Kira, Laura, Lena, Marc, Max, die Michels, Olli, Pia, and Raphael.

I am endlessly grateful to my family, for supporting me in every aspect and being the great individuals they are, without you nothing would have been possible and I can't tell you how happy I am to have you in my life. Especially to my brother Jannis, I would like to express my thankfulness, we had a great time living, studying, and, to a certain degree, doing the PhD together.

Finally, I would like to thank Denise, my partner in crime and the love of my life, you know what you did. Without you everything is nothing, thank you for everything.