# Characterization of three Uridine Monophosphate Kinases from *Arabidopsis thaliana*

Von der Naturwissenschaftlichen Fakultät

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation von

Jannis Rinne, M. Sc.

2024

Referent: Prof. Dr. Claus-Peter Witte

Korreferent: Prof. Dr. rer. nat Hans-Peter Braun

Tag der Promotion: 15.03.2024

# 1.    Abstract

*De novo* synthesis of pyrimidine nucleotides is an essential pathway of plant primary metabolism that ends with UMP. This UMP serves as a precursor for the synthesis of all other pyrimidine nucleotides in plants. For this purpose, UMP is first phosphorylated by UMP KINASEs (UMKs), of which two families exist in plants: Eubacterial and AMP KINASE (AMK)-like UMKs. It is suggested that most UMK activity *in vivo* is catalyzed by AMK-like UMKs, of which *Arabidopsis thaliana* possesses three isoforms called UMK1, UMK2 and UMK3. Although the phosphorylation of UMP is of high importance for plant metabolism, the family has only been sparsely described so far. The aim of this thesis was to characterize the family of AMK-like UMKs of Arabidopsis and elucidate the specific roles of the three UMKs in pyrimidine nucleotide metabolism. While not all vascular plants possess a *UMK1* gene, *UMK2* and *UMK3* are evolutionary conserved throughout the plant kingdom. Biochemical characterization showed that all three UMKs can phosphorylate UMP, CMP and dCMP. UMK3 stood out in this analysis as it possessed the highest catalytic efficiencies for UMP and CMP. Studies on subcellular localization showed that UMK1 and UMK3 are cytosolic enzymes, while UMK2 resides in mitochondria. The analysis of CRISPR/Cas9 mutant lines revealed a central function of UMK3 in pyrimidine nucleotide metabolism. A weak mutant of *UMK3,* encoding a partially defective enzyme, showed growth deficiencies and an altered pyrimidine nucleotide content, whereas null mutant alleles of *UMK3* could not be inherited. While UMK3 is essential during the reproductive phase and for embryo development, mutants of *UMK1* and *UMK2* were phenotypically normal. In the background of the weak *UMK3* mutant, *in vivo* functions of the other UMKs were unmasked. UMK2 mainly operates as dCMP kinase, and during germination, UMK2 activity is required for normal mitochondrial DNA replication. Also UMK3 was shown to be important for mitochondrial DNA replication during germination, indicating that mitochondria are capable of importing pyrimidine nucleotides from the cytosol. The data also suggest that mitochondria are able to export pyrimidine nucleotides, implying the existence of a mitochondrial pyrimidine nucleotide transporter. The role of UMK1 was less clear. UMK1 appears to be mainly involved in the phosphorylation of CMP, although this function was only observed in the background of the weak *UMK3* mutant. It is possible that UMK1 is more important when the activity of UMK3 is downregulated. That such regulation may exist is indicated by the inhibition of UMK3 at high substrate concentrations. In such a situation, the activity of UMK1 could be important for the recycling of NMPs from DNA/RNA breakdown or nucleoside salvaging. Although the Arabidopsis AMK-like UMKs are all ubiquitously expressed and have similar enzymatic activities, this work shows that they are not redundant or at best partially redundant, but fulfill specific biological functions.

# Table of Contents

Table of contents

# 2.  Introduction

The results of this thesis were also used to write the publication '**Three UMP kinases of *Arabidopsis thaliana* have different roles in pyrimidine nucleotide synthesis and (deoxy)CMP salvage**', which, by the time this thesis is submitted, is under review in the journal 'The Plant Cell'.

## 2.1.  Functions and structure of nucleotides

Already in school, we are taught that nucleotides are the building blocks of DNA and RNA. In this function, they are indispensable for life and their sequence decides whether we become bacterium or plant, man or mouse. Beyond that, nucleotides fulfill a plethora of other roles without which life as we know it would not be possible. Adenosine triphosphate (ATP) is the universal energy carrier in all living organisms and the cells in our body consume almost our own body weight of ATP every day. Uridine diphosphate glucose (UDP-glucose) serves as precursor for the synthesis of cellulose, which is the main component of plant cell walls and the most abundant organic compound on earth. Examples of other essential nucleotide derived molecules are nicotinamide adenine dinucleotide (NADH), flavin adenine dinucleotide or coenzyme A, which are involved as cofactors in various enzymatic reactions of primary and specialized metabolism. One could say, our life depends on nucleotides.

A nucleotide consists of a variable nucleobase linked via a glycosidic bond to a pentose sugar moiety, and one to three phosphate groups. Without a phosphate group, the sugar and nucleobase are referred to as a nucleoside. The sugar moiety is either a ribose in ribonucleotides, which are used to form RNA, or a deoxyribose in deoxynucleotides, which are incorporated into DNA. Depending on the number of phosphate groups, nucleotides are called nucleoside monophosphates (NMPs), nucleoside diphosphates (NDPs) or nucleoside triphosphates (NTPs). In the following, NMP, NDP or NTP is used if both ribonucleotides and deoxynucleotides are meant and these abbreviations always refer to nucleotides with the phosphate attached to the 5'-carbon. If only the respective ribonucleotide is referred to, the abbreviations rNMP, rNDP or rNTP are used and dNMP, dNDP or dNTP are used for the deoxynucleotide. Nucleobases are divided in two groups, purines and pyrimidines. Characteristic for purine bases is a heterocyclic double ring structure with four nitrogen atoms, while pyrimidine bases consist of a single ring with two nitrogen atoms. The canonical purine bases of DNA are adenine and guanine and the pyrimidine bases are thymine and cytosine. In RNA, the pyrimidine base uracil is found instead of thymine.

Plants possess the necessary enzymes for the synthesis and degradation of all nucleotides. The nucleotide metabolism can generally be subdivided into the three routes of *de novo* synthesis, nucleoside salvage and nucleotide degradation. The *de novo* synthesis of adenosine monophosphate (AMP) is the main source of purine nucleotides in the cell and AMP can serve as precursor for the synthesis of all purine nucleotides. Likewise, *de novo* synthesized uridine monophosphate (UMP) is the precursor for all pyrimidine nucleotides (Witte and Herde, 2020). Nucleotides can also be obtained from nucleoside salvage. The plant is able to take up nucleosides from the environment (Girke et al., 2014) or recycle them from metabolic processes like RNA turnover (Li et al., 2018) or DNA base excision repair (Córdoba-Cañero et al., 2009). The process opposing *de novo* synthesis and salvage is the nucleotide and nucleoside degradation, which releases the nitrogen bound in the rings of the nucleobase and can support the plant in situations of nitrogen deficiency (Soltabayeva et al., 2018).

## 2.2. Pyrimidine nucleotide synthesis in Arabidopsis

The first step of the UMP *de novo* synthesis is catalyzed by the enzyme ASPARTATE CARBAMOYLTRANSFERASE (ATC) in plastids (Chen and Slocum, 2008; Witz et al., 2012). ATC is also a key regulator in UMP *de novo* synthesis, as it can be inhibited by UMP binding to its active site (Bellin et al., 2021a). Subsequent steps of UMP *de novo* synthesis are located in the cytosol and mitochondria (Witz et al., 2012). The last step is catalyzed by the enzyme UMP SYNTHASE, which is located in the cytosol, but seems to also be associated with plastids (Doremus and Jagendorf, 1985; Witz et al., 2012). To exhibit its feedback regulation on ATC, UMP has to be channeled from the cytosol into the plastids, but this mechanism is not yet well characterized. It is possible, that the UMP first has to be degraded to uracil, before it can be transported into plastids by the PLASTIDIC URACIL TRANSPORTER (PLUTO; Witz et al., 2012) where the enzyme URACIL PHOSPHORIBOSYLTRANSFERASE (UPP) again catalyzes the formation of UMP (Mainguet et al., 2009). A transport mechanism for uracil into the plastid appears logical, as the first step of the uracil degradation pathway is also located there (Cornelius et al., 2011). It is questionable however, if PLUTO is the only plastidic uracil transporter, as mutation of *PLUTO* did not influence the concentration of the uracil degradation pathway product beta-alanine, whereas mutation of other genes in the pathway did (my unpublished data). This suggests the existence of another yet uncharacterized plastidic uracil transporter. Recently, PLASTIDIC NUCLEOSIDE KINASE 1 (PNK1) has been described and was linked to the feedback inhibition of ATC. PNK1 catalyzes the formation of UMP from uridine in plastids and a knockout of *PNK1* increases the global UTP concentration, when the gene encoding the uridine

degradation enzyme NUCLEOSIDE HYDROLASE 1 (NSH1) is also mutated (Chen et al., 2023). An ATC feedback inhibition via UMP production from PNK1 also seems possible, but would require a plastidic transport mechanism for uridine.

For the synthesis of other pyrimidine nucleotides, the *de novo* synthesized UMP must first be phosphorylated (**Figure 1**). This reaction is catalyzed by UMP KINASEs (UMKs), which are characterized in this thesis and described in detail in a later chapter of this introduction. The product of the UMK reaction is uridine diphosphate (UDP), which is phosphorylated again to uridine triphosphate (UTP) by enzymes of the NUCLEOSIDE DIPHOSPHATE KINASE (NDPK) family. Arabidopsis possesses five NDPKs with different subcellular localizations (Dorion and Rivoal, 2015). They catalyze the transfer of a phosphate group from any NTP to any NDP (Parks and Aganwal, 1973). The NDPK family can be subdivided into four types. Arabidopsis possesses one gene encoding a type I NDPK (NDPK1), which localizes to the cytosol, but also shows association with peroxisomes and the nucleus (Reumann et al., 2009). Studies in potato showed that NDPK1 is responsible for the majority of NDPK activity in cellular extracts (Dorion et al., 2006). Thus, in Arabidopsis, NDPK1 could catalyze UTP formation downstream of a cytosolic UMK. Type II NDPK localizes to plastids (Bölter et al., 2007) and type III, of which Arabidopsis possesses two isoforms, shows a dual-targeting to plastids and mitochondria (Sweetlove et al., 2001; Spetea et al., 2004). Interestingly, NDPK3, which is a type III enzyme, has not been shown to reside in the mitochondrial matrix, but in the mitochondrial intermembrane space or anchored to the inner membrane (Sweetlove et al., 2001; Knorpp et al., 2003). Type IV NDPK putatively localizes to the endoplasmic reticulum (Dorion and Rivoal, 2015).
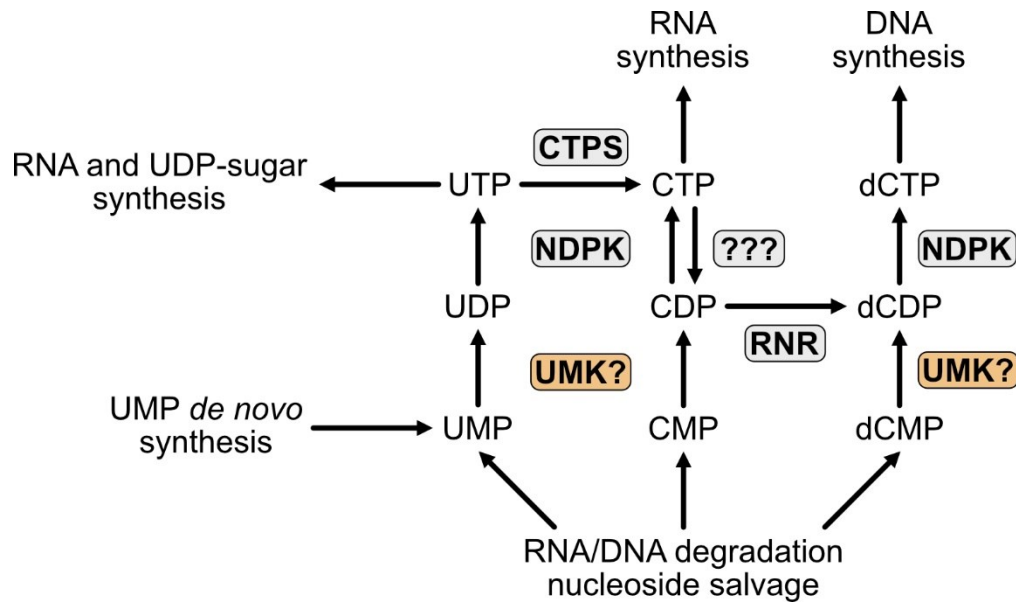
**Figure 1. Simplified model of pyrimidine nucleotide metabolism.**
The *de novo* synthesis of UMP is the starting point for the synthesis of all other pyrimidine nucleotides. The subcellular localizations of the enzymes are not indicated. A revised model based on the results of this thesis is presented in **Figure 30**. Enzymes: UMK, UMP KINASE; NDPK, NUCLEOSIDE DIPHOSPHATE KINASE; CTPS, CTP SYNTHASE; RNR, RIBONUCLEOTIDE REDUCTASE; ???, unknown phosphatase.

UTP is the precursor for the synthesis of UDP-sugars, cytidine triphosphate (CTP) or is directly incorporated into RNA. Synthesis of UDP-sugars is essential for plant growth and development. The most abundant UDP-sugar is UDP-glucose, which serves as a precursor for cellulose synthesis. Cellulose makes up the largest part of the plant cell wall. Different classes of enzymes are able to synthesize UDP-glucose *in vivo*. UDP GLUCOSE PYROPHOSPHORYLASE (UGP) or the more promiscuous UDP SUGAR PYROPHOSPHORYLASE (USP) can generate UDP-glucose from glucose-1-phosphate and UTP. The reactions catalyzed by UGP and USP are reversible and dependent on the intracellular concentrations of substrates and products (Kleczkowski et al., 2010). Thus, the continuous production of UTP is necessary to maintain the synthesis of UDP-sugars by UGP and USP for plant development. While UGP is specific for the synthesis of UDP-glucose, USP is also able to utilize other sugar-1-phosphates for production of the respective UDP-sugars (Kotake et al., 2007). UDP-glucose can also be obtained in the cell by the breakdown of sucrose, catalyzed by the enzyme SUCROSE SYNTHASE, of which Arabidopsis possesses multiple isoforms with distinct expression patterns (Bieniawska et al., 2007). In addition to the reaction catalyzed by USP, different UDP-sugars can also be produced from UDP-glucose by various enzymatic reactions. Some of these reactions are in equilibrium with each other. Thus, some UDP-sugars are interconvertible (Kotake et al., 2009). Sugar polymers generated from different UDP-sugars other than UDP-glucose are called hemicellulose, which is

an important building block of plant cell walls. Mainly in photosynthetically inactive tissues, UDP-glucose can also be generated from the breakdown of sucrose catalyzed by the enzyme SUCROSE SYNTHASE

UTP is not only required for the synthesis of UDP-sugars, but also for the production of CTP. The further processing of UTP thus represents a branching point in the plant's primary metabolism. The conversion of UTP to CTP is catalyzed by CTP SYNTHASEs (CTPSs), of which Arabidopsis possesses five isoforms which are all located in the cytosol. These isoforms are highly homologues among each other and are likely all functional. An *in vitro* activity as a CTPS was detected for one of the isoforms (Daumann et al., 2018). The CTPS enzymes show various levels of regulation, as they are activated by guanosine triphosphate (GTP) and inactivated by their product CTP. Associated with a CTP-induced inactivation is the formation of inactive filamentous protein structures (Noree et al., 2014; Daumann et al., 2018). It is unclear why Arabidopsis possesses five *CTPS* genes. Phenotypical characterization of knockout lines revealed no abnormalities for four of the five isoforms compared to wild type plants, which raises questions about redundancy (Bellin et al., 2021b). A specific role during embryo development was found for CTPS2 (Hickl et al., 2021). It was not possible to generate homozygous knockout lines of *CTPS2* due to embryo lethality, but knockdown lines were negatively affected in their growth, indicating a function of CTPS2 that cannot be performed by other members of the CTPS family (Bellin et al., 2021b; Hickl et al., 2021). CTP is also a substrate for the production of cytidine diphosphate (CDP) diacylglycerol (CDP-DAG) and CDP-diphosphoethanolamine (CDP-Etn), which are two important precursors for the synthesis of phospholipids. Phospholipids are essential components of biological membranes and signal transduction cascades in plants (Nakamura, 2017).

In addition to pyrimidine rNTPs, the plant also requires dNTPs for DNA replication. The dNTPs are generated by the enzyme complex RIBONUCLEOTIDE REDUCTASE (RNR), which is, like the other enzymes involved in this pathway, also located in the cytosol (Lincker et al., 2004). The enzyme complex catalyzes the rate-limiting step in the synthesis of dNTPs and requires rNDPs as substrates (Wang and Liu, 2006). Consequently, the CTP generated from the CTPS reaction has to be dephosphorylated to CDP first, before it can serve as a substrate for the RNR complex. It has been suggested that a specific phosphatase for this reaction might exist. Alternatively, CDP can be obtained from nucleoside salvage, the phosphorylation of CMP stemming from RNA turnover or the CDP-containing precursors of phospholipid synthesis (Witte and Herde, 2020). It is unclear however, whether enough CDP is generated from these processes to fuel the RNR reaction *in vivo*. Not all dNDPs can be synthesized directly by the RNR complex. The synthesis of deoxythymidine diphosphate (dTDP) first requires a reduction of the ribose hydroxyl group of UDP to generate dUDP. The dUDP must then be dephosphorylated to serve as a substrate for THYMIDILATE

SYNTHASE, which generates dTMP (Lazar et al., 1993; Gorelova et al., 2017). Finally, the dNDPs are phosphorylated by the NDPKs to generate dNTPs for DNA replication. Overall, all these reactions rely on an initial phosphorylation of *de novo* synthesized UMP, which is catalyzed by UMKs.

## 2.3. Salvage and recycling in pyrimidine metabolism

Turnover of cellular RNA molecules like messenger RNA (mRNA) or ribosomal RNA (rRNA) is an essential process to regulate gene expression and cellular functions (Li et al., 2018; Floyd et al., 2015). It is constantly happening in every cell and releases nucleotides or nucleosides, which can be recycled. mRNAs are protected from premature degradation by a 5' $m^7$guanosine diphosphate (GDP) cap and a 3' poly(A) tail. These protective elements are removed by decapping enzymes (Xu et al., 2006) or poly(A)-specific ribonucleases (Yan, 2014). The turnover of 'unprotected' mRNA in plants occurs by two conserved mechanisms. Degradation from 5'- to 3'-end is catalyzed by EXORIBONUCLEASE (XRN) proteins, which reside in the nucleus and cytoplasm, while degradation from the 3'-end is catalyzed by exosomes, which are also localized in the nucleus and cytosol (Chiba and Green, 2009). XRNs and exosomes hydrolyze the phosphodiester bond in the backbone of RNA, which releases rNMPs. Hydrolysis can occur at two sites, either at the 5'- or the 3'-phosphoester bond between ribose and the phosphates. Consequently, either nucleoside-3'-monophosphates or -5'-monophosphates are generated. The latter could directly be utilized by nucleoside monophosphate kinases (NMKs) for the production of rNDPs. It is unclear, what form is released by XRNs and exosomes, but the formation of 5'-NMP was observed for a bacterial ribonuclease (Kim et al., 2019), of which a probable homolog also exists in Arabidopsis, indicating that nucleoside-5'-monophosphates are formed during the RNA turnover of Arabidopsis. The vacuoles are also an important sink for RNA. It is estimated that more than 70 % of ribonuclease activity takes place there (Abel and Glund, 1987). The NMPs obtained from vacuolar RNA turnover are further metabolized inside the vacuole by phosphatases to release nucleosides and phosphate (Witte and Herde, 2020). For the latter, vacuoles are a major storage organ (Yang et al., 2017). The resulting nucleosides can be exported to the cytosol by the vacuolar EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 1 (ENT1) to make them available for nucleoside salvaging reactions (Bernard et al., 2011).

In addition to the turnover of RNA, the plant can also obtain nucleosides or nucleobases by taking them up from the environment (Girke et al., 2014). The nucleosides and nucleobases can then be converted to nucleotides by the plant in a process called salvaging. The majority of pyrimidine salvaging is performed by two dual-specific URIDINE CYTIDINE KINASEs (UCKs), which are located in the cytosol and catalyze the

formation of UMP and CMP (Ohler et al., 2019). As already mentioned before, plastids also possess enzymes, which catalyze the formation of UMP from uridine (PNK1; Chen et al., 2023) or uracil (UPP; Mainguet et al., 2009). However, an *in vivo* relevance of the activity of UPP for nucleotide production is questionable (Ohler et al., 2019), suggesting that pyrimidine salvage occurs mainly via the nucleoside and not via the nucleobase. In addition to ribonucleoside salvaging enzymes, the plant also possesses a DEOXYNUCLEOSIDE KINASE (dNK), which catalyzes the formation of all dNMP except dTMP from deoxynucleosides and resides inside the mitochondria (Clausen et al., 2012; Clausen et al., 2014) and chloroplasts (our unpublished data). During seed imbibition and germination, a high turnover of parental RNA has been observed, indicating that salvaging may have an increased role during early development (Li et al., 2006). The NMPs generated from salvaging can serve as potential substrates of the UMKs.

## 2.4. UMP Kinases

UMKs catalyze the transfer of a phosphate group from NTPs to different NMPs. **Figure 2** shows the reaction with UMP and ATP as substrates. In this case the products of the reaction are UDP and adenosine diphosphate (ADP). Generally, NMKs show a high substrate specificity in their NMP binding site (Yan and Tsai, 1999). This high specificity has been observed for Arabidopsis GUANOSINE MONOPHOSPHATE (GMP) KINASE, which is only active with GMP and dGMP (Kumar et al., 2000). UMKs however, generally show a broader substrate spectrum. Human UMK is highly active with UMP, CMP and dCMP and also shows some activity with dUMP, AMP and dAMP (van Rompay et al., 1999), as well as with various nucleotide analogs (Pasti et al., 2003). Yeast UMK crystal structure analysis suggests, that the broader substrate spectrum of UMKs is based on a slightly larger NMP binding site compared to other NMKs like for example ADENOSINE MONOPHOSPHATE KINASEs (AMKs), which is lined by three conserved residues (alanine-47, isoleucine-75 and threonine-81; numbers refer to the positions in yeast UMK). This appears counterintuitive at first, as uracil is smaller than adenine. However, it has been shown that UMKs require the presence of a water molecule in their NMP binding site for activity with uridinylates. Thus, the NMP binding site is bigger to accommodate the water molecule (Müller-Dieckmann and Schulz, 1995). The NTP binding site generally shows a lower substrate specificity, but *in vivo* ATP is mainly utilized as a substrate due to its higher abundance (Yan and Tsai, 1999).
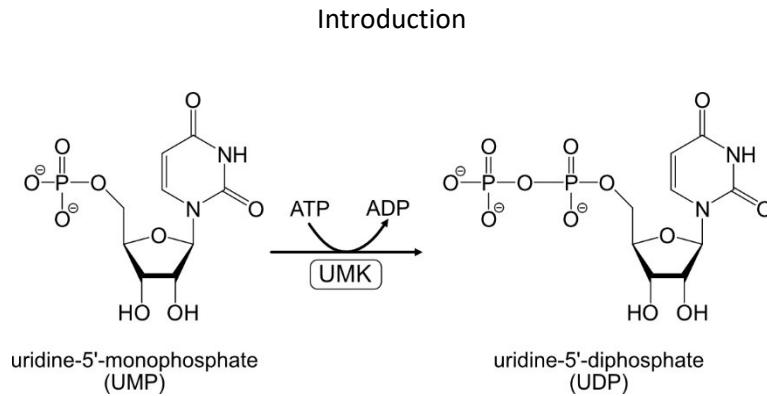
**Figure 2. Reaction catalyzed by UMKs with UMP as substrate.**

Plant UMKs can be divided into two families: eubacterial and AMK-like UMKs. Eubacterial UMKs show similarities to UMKs from prokaryotic organisms (Serina et al., 1995). Arabidopsis possesses two genes encoding eubacterial UMKs, of which one has been described in the literature as PLASTID UMP KINASE (PUMPKIN). PUMPKIN shows UMP phosphorylation activity *in vitro*, but is also involved in RNA-processing *in vivo*. Plants lacking PUMPKIN are pale and show growth deficits, which is caused by defects in plastid translation and photosynthesis (Schmid et al., 2019). A similar phenotype can be observed in rice plants lacking the homologous gene (Chen et al., 2018). It is unclear, whether the phenotypical abnormalities of the plants are caused by the absence of UMP phosphorylation activity in the plastids or by defects in RNA-processing. The other eubacterial UMK encoded at the locus At3g10030 has not been characterized yet.

Arabidopsis possesses three genes encoding potential UMKs belonging to the second family with similarities to AMKs. They were named *UMK1* (At3g60180), *UMK2* (At4g225280) and *UMK3* (At5g26667) by Lange et al., 2008 and this nomenclature was followed in this work. Also, from now on, UMK is referring to AMK-like UMKs. A biochemical characterization of UMK3 revealed that the enzyme is able to phosphorylate UMP, CMP and with less efficiency dCMP (Zhou et al., 1998). Using the protein sequence of UMK3 in BLASTp searches, protein sequences from other plants were recovered and a phylogenetic tree calculated by Claus-Peter Witte (**Figure 3**). The tree shows a clear distinction between UMK2 and UMK3, while UMK1 is very similar to UMK3. UMK2 and UMK3 are also highly conserved in all 23 analyzed plant species, while UMK1 is only found in Brassicaceae. A fourth group, consisting mainly of UMKs from monocotyledonous plants, was named UMK4. Another potential Arabidopsis AMK-like UMK, which is encoded at the locus At3g60961, is likely non-functional as it is missing more than 50 conserved N-terminal amino acids (**Figure A 1**). Therefore, this protein was excluded from the phylogenetic analysis and further molecular studies.
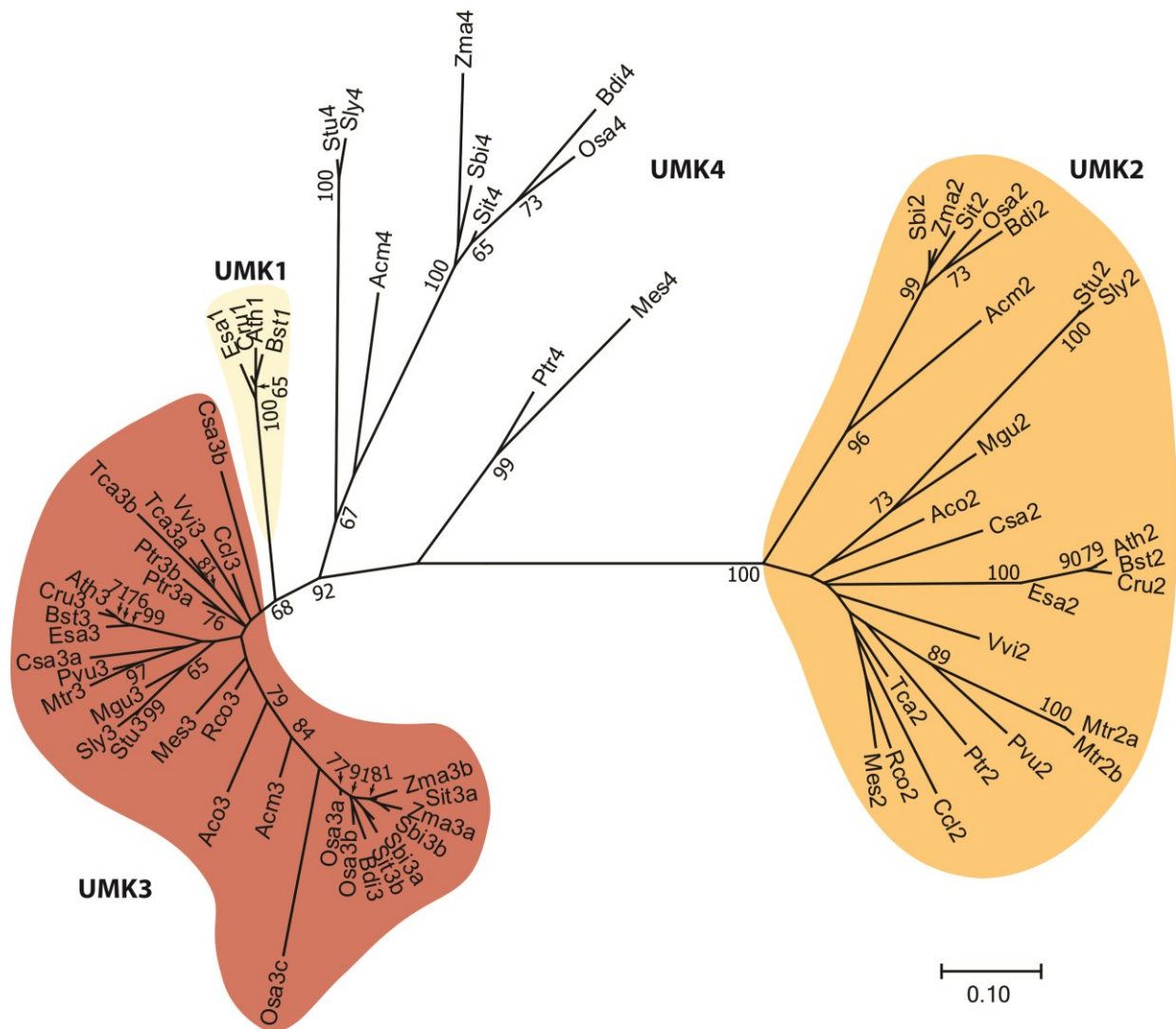
**Figure 3. Maximum likelihood tree constructed with AMK-like UMK sequences from 23 vascular plant species.**
The tree with the highest log likelihood is displayed. 1000 bootstraps were performed and only bootstrap values over 65% are shown. Branch lengths indicate the number of substitutions per site (see legend). Species names and accession numbers are given in **Table A 1** and the corresponding multiple alignment is shown in **Figure A 1**.

## 2.5. Aim of this thesis

While the basic function of AMK-like UMKs for plant pyrimidine metabolism is clear (they phosphorylate UMP), no thorough characterization of the gene family has been conducted yet. However, the importance of UMKs for plant pyrimidine metabolism justifies a detailed analysis. Not only the synthesis of nucleic acids depends on UMKs, but also the production of cellulose and thus of plant biomass. To gain insight into the function of the three AMK-like UMKs, the first aim of this thesis was to assess their substrate

specificities and kinetic parameters by biochemically characterizing the enzymes *in vitro*. For a better understanding of their roles *in vivo*, the subcellular localization of the UMKs needed to be investigated. An additional analysis of the metabolome of *UMK* mutant plants generated with the CRISPR/Cas9 system could contribute to the understanding of the specific function of UMKs in pyrimidine metabolism. The mutant plants were also to be analyzed during different growth phases of the plant in order to find developmental stages in which one of the UMKs could play an increased role.

# 3. Results

## 3.1. Biochemical characterization of the three AMK-like UMKs

UMKs accept a variety of substrates in their NMP binding pocket. Therefore, the first aim was to assess the substrate specificity of the three potential UMKs encoded at the loci At3g60180 (*UMK1*), At4g25280 (*UMK2*) and At5g26667 (*UMK3*). For UMK3 it has already been shown that an enzyme preparation from *Escherichia coli* cells is able to phosphorylate UMP and CMP (Zhou et al., 1998). The three *UMK* genes fused with a sequence encoding a C-terminal Strep-tag were expressed in *Nicotiana benthamiana* and purified via StrepTactin affinity chromatography. Successful purification was confirmed by Coomassie staining and Immunoblot using a Strep-tag specific antibody (**Figure 4 A and B**). The substrate specificity was tested in a coupled enzyme assay by incubating the purified proteins together with different NMPs and ATP. The coupled assay also included the enzymes pyruvate kinase and lactate dehydrogenase, as well as their substrates phosphoenolpyruvate (PEP) and NADH. This was necessary, as the reaction products of the UMK cannot be measured directly. In the coupled assay, the oxidation of NADH to $NAD^+$ is proportional to the production of ADP catalyzed by the UMK. ADP serves as a substrate for the pyruvate kinase, which produces pyruvate and ATP from PEP and ADP. The pyruvate and NADH are substrates of the lactate dehydrogenase, which produces $NAD^+$ and lactate and the decrease of NADH can be monitored in a photometer at a wavelength of 340 nm. All three UMKs showed activity with UMP, CMP and dCMP as phosphate acceptors (**Figure 4 C**). Minor activity was also detected with dUMP for UMK2 and UMK3 and AMP for UMK1 and UMK2. These results confirm the published data for UMK3 (Zhou et al., 1998), but show that the enzyme's substrate spectrum is even broader. The findings are in line with results for the human UMP kinase, which is also able to phosphorylate UMP, CMP and dCMP with high efficiency and various other NMPs with lower efficiency (van Rompay et al., 1999; Pasti et al., 2003).
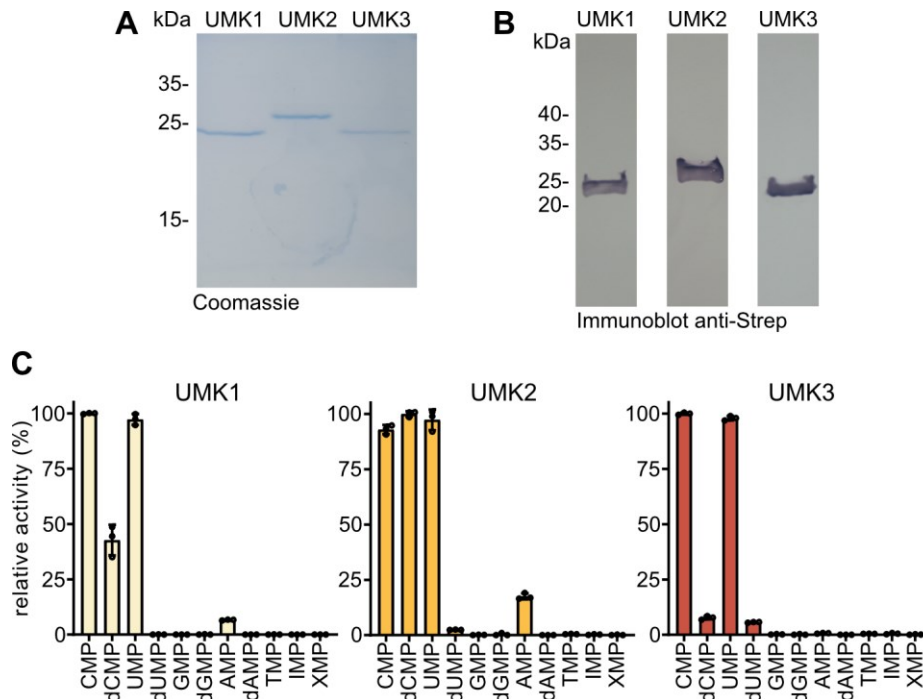
**Figure 4. Purification and substrate specificity scan of the three AMK-like UMP kinases from Arabidopsis. A)** Purity of the UMKs after affinity purification via the Strep tag shown on a Coomassie-stained SDS gel. 10 µl of the affinity elution fractions were loaded corresponding to 330 ng, 290 ng and 80 ng protein for UMK1, UMK2 and UMK3, respectively. The expected molecular masses are 23.1, 27.8 and 22.5 kDa for C-terminal HA-Strep-tagged UMK1, UMK2 and UMK3, respectively. **B)** Immunoblot of the purified UMKs detected with an anti-Strep antibody. **C)** Substrate screen using the purified enzymes with several nucleotide monophosphates (NMPs). The activity was determined with 1 mM NMP and 1 mM Mg-ATP at 22°C in three independent enzyme reactions using the same enzyme preparation. Error bars are SD.

To determine the kinetic constants of the three UMKs, the enzymatic activity with different substrate concentrations was determined for UMP, CMP and dCMP. Again, the coupled assay was employed with a fixed ATP concentration of 1 mM. Activity was determined using NMP concentrations of 0.025, 0.05, 0.1, 0.25, 0.5, 1, 2.5 and 10 mM with five replicates each. The results were fitted by the GraphPad software using the Michaelis-Menten equation (**Figure 5**), but for some substrates it was apparent that the enzymatic activity did not show the typical pattern of a Michaelis-Menten kinetic. For UMK3 in particular, no increase in activity was observed at higher substrate concentrations, but a strong reduction in activity at higher amounts of UMP or CMP. Thus, a different model for fitting was employed using an equation that takes substrate inhibition into account (Equation 5.44, Copeland, 2000). Weaker substrate inhibition was also observed for UMK1 with CMP and dCMP, as well as for UMK2 with CMP. In these cases, the results were fitted with both formulas and the resulting curve with the better fit, indicated by a higher $R^2$ value,

was used for determination of $K_M$ and $k_{cat}$. To confirm a good fit of the Michaelis-Menten curves, the data were linearized according to Hanes and fitted by linear regression.
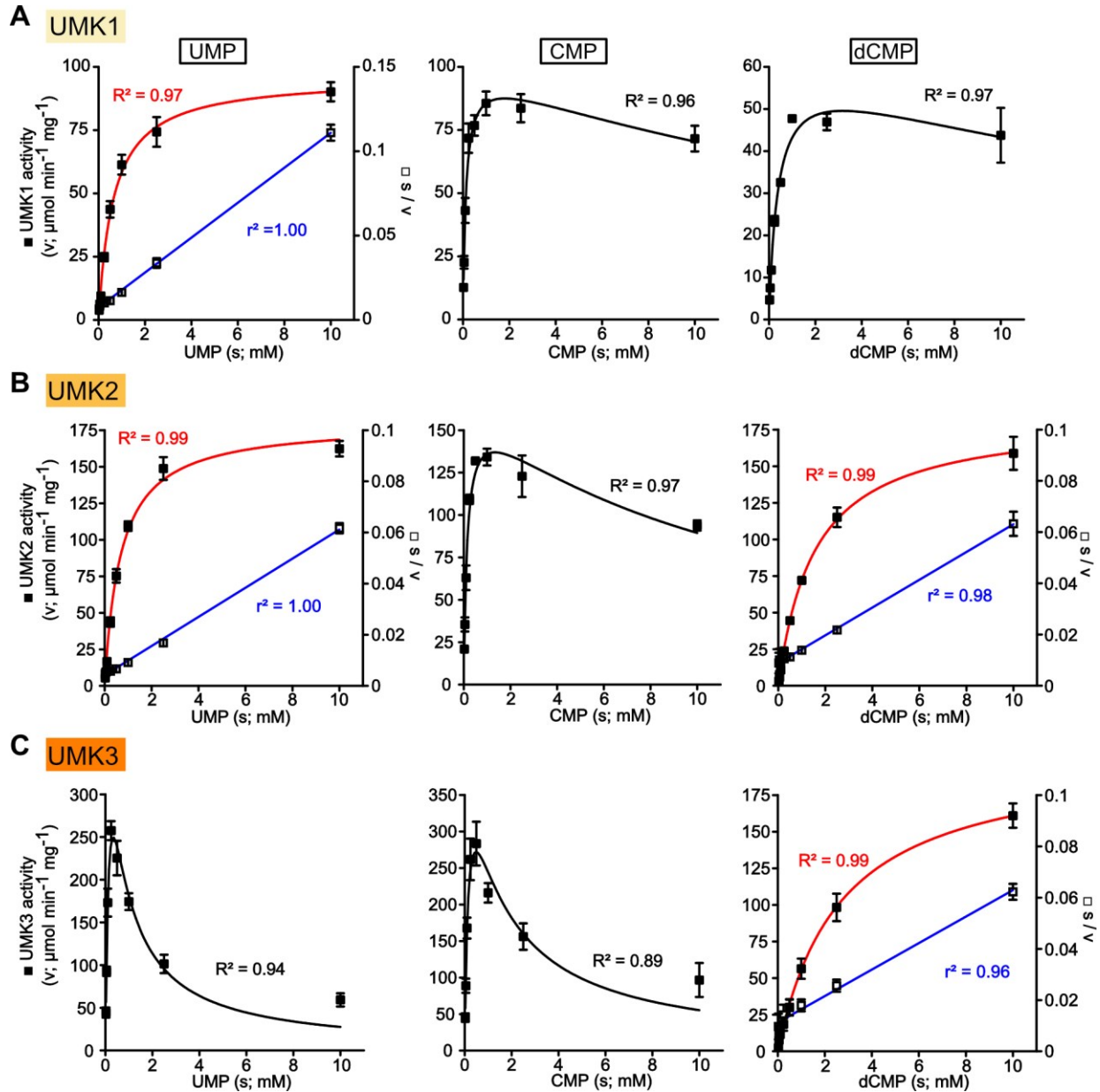


**Figure 5. Kinetic curves of the three UMKs with UMP, CMP and dCMP as phosphate acceptors.**
Kinetic data for **A)** UMK1, **B)** UMK2 and **C)** UMK3 with the substrates UMP, CMP and dCMP. For each substrate concentration, five independent enzymatic reactions were run using the same enzyme preparation. Error bars are SD. The data were fitted using the Michaelis-Menten equation (red curves) or an equation accounting for substrate inhibition (equation 5.44 from Copeland, 2000, black curves) or by linear regression for the Hanes plot (s/v over s, blue line) where applicable (right axis). s, substrate concentration; v, reaction velocity.

The $K_M$ and $k_{cat}$ values for the three enzymes with their substrates UMP, CMP and dCMP as well as the respective $k_{cat}/K_M$ values as an indicator for catalytic efficiency are displayed in **Table 1** for UMK1, **Table 2** for UMK2 and **Table 3** for UMK3. The $K_M$ value indicates the substrate concentration at which half of the enzymes are complexed with a substrate molecule. The $k_{cat}$ value indicates the number of enzymatic reactions that take place at the maximum velocity of the enzyme per second. If $k_{cat}$ is divided by $K_M$, the catalytic efficiency of an enzyme is obtained. A higher value here is an indicator for a low $K_M$ (good affinity to the substrate) and/or a high $k_{cat}$ (the reaction is catalyzed quickly). For UMP, it is evident that UMK3 possesses the best catalytic efficiency of the three UMKs, as it has the lowest $K_M$ and the highest $k_{cat}$ value. For CMP, the $K_M$ values are similar for all three enzymes, but again UMK3 has the best catalytic efficiency as it catalyzes the reaction faster than the other two. With dCMP as a substrate all three enzymes have significantly lower catalytic efficiencies, but in contrast to UMP and CMP, UMK1 and UMK2 possess the highest catalytic efficiency here. Overall, the activity with UMP as substrate is probably the most important function of the UMKs *in vivo*, as more UMP is produced within the cell than CMP and dCMP, which are only derived from RNA/DNA degradation or nucleoside salvage processes and are not synthesized *de novo*. The significantly better catalytic efficiency of UMK3 with UMP in comparison to UMK1 and UMK2 hints towards a key role of UMK3 in pyrimidine nucleotide metabolism. The results presented here deviate from the data presented for UMK3 by Zhou et al., 1998 who report similar $K_M$ values, but significantly lower $k_{cat}$ values for both UMP and CMP. They also did not report the strong substrate inhibition of UMK3 that was observed here.

**Table 1. Kinetic constants of UMK1 with the substrates UMP, CMP and dCMP.**

| Substrate | $K_M$ (mM) | $k_{cat}$ (s$^{-1}$) | $k_{cat}/K_M$ (mM$^{-1}$ s$^{-1}$) |
|---|---|---|---|
| UMP | **0.646** ± 0.037 | **41.1** ± 0.7 | 63.6 |
| CMP | **0.134** ± 0.013 | **43** ± 1.3 | 320.2 |
| dCMP | **0.391** ± 0.046 | **26.4** ± 1.3 | 67.5 |

**Table 2. Kinetic constants of UMK2 with the substrates UMP, CMP and dCMP.**

| Substrate | $K_M$ (mM) | $k_{cat}$ (s$^{-1}$) | $k_{cat}/K_M$ (mM$^{-1}$ s$^{-1}$) |
|---|---|---|---|
| UMP | **0.694** ± 0.036 | **95.1** ± 1.5 | 137 |
| CMP | **0.154** ± 0.014 | **89** ± 2.8 | 578.6 |
| dCMP | **1.555** ± 0.073 | **97.2** ± 1.6 | 62.5 |

**Table 3. Kinetic constants of UMK3 with the substrates UMP, CMP and dCMP.**

| Substrate | $K_M$ (mM) | $k_{cat}$ (s$^{-1}$) | $k_{cat}/K_M$ (mM$^{-1}$ s$^{-1}$) |
|---|---|---|---|
| UMP | **0.24** ± 0.054 | **252.1** ± 37 | 1050.2 |
| CMP | **0.179** ± 0.038 | **196.7** ± 22.4 | 1101 |
| dCMP | **2.653** ± 0.172 | **85.1** ± 2.2 | 32.1 |

When repeating the coupled assay with UMK3 and different amounts of ATP in the presence of a fixed amount of UMP, substrate inhibition was also observed (**Figure 6 A**). Further investigations revealed however, that this was not caused by the enzymatic properties of UMK3, but most likely by an ATP-induced inhibition of PK. A UMK assay with alternative HPLC/MS-detection of the reaction products was performed with 1 and 2.5 mM ATP and stopped after 0, 5 and 10 minutes and the amounts of UMP and UDP quantified (**Figure 6 B**). In the samples containing 2.5 mM ATP, where initially a lower activity was observed in the coupled assay, the detected amount of UDP was greater than in the samples which contained 1 mM ATP. This indicates that UMK3 is catalyzing the reaction faster in the presence of more ATP and not slower, as initially suspected. A possible and quite likely explanation for this lies in the properties of the enzymes used in the coupled enzyme assay. The PK, which catalyzes the transfer of a phosphate group from PEP to ADP, is inhibited by ATP, the product of this reaction (Carbonell et al., 1973). The coupled enzyme assay is therefore not suitable for higher ATP concentrations such as those used here. Also increasing the amount of magnesium (Wood, 1968) did not solve this problem. As this analysis was carried out towards the end of the project, it was not possible to optimize this HPLC/MS-detection approach for the determination of the kinetic parameters of the UMKs with ATP due to time constraints.
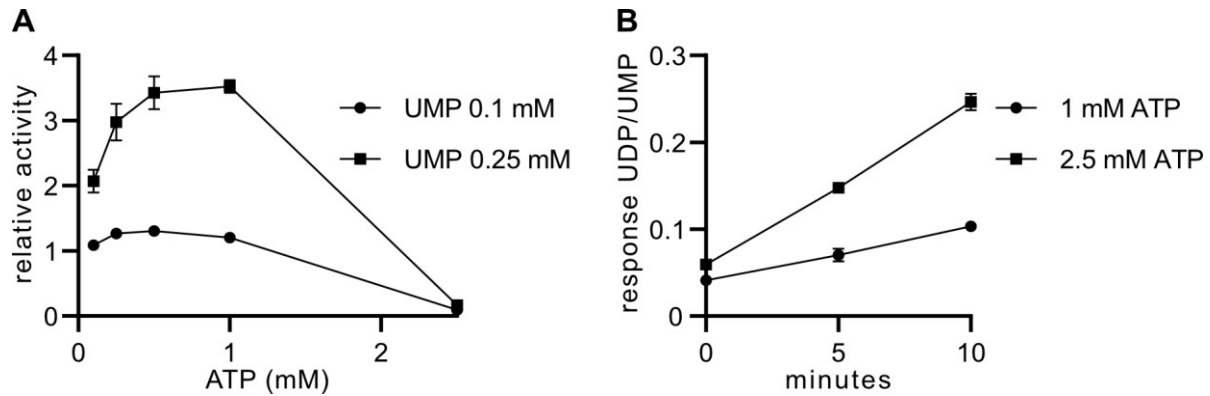
**Figure 6. Enzymatic assay of UMK3 with varying ATP amounts.**
**A)** Enzymatic assay with UMK3 in the presence of 0.1 or 0.25 mM UMP with varying ATP concentrations. For each substrate concentration, five independent enzymatic reactions were run using the same enzyme preparation. Error bars are SD. **B)** The same assay was repeated with 0.25 mM UMP and 1 or 2.5 mM ATP. Samples were taken at the start of the assay, after 5 and after 10 minutes. A 1:100 dilution of each reaction was quantified by LC-MS analysis to determine the amounts of UDP and UMP. The y-axis shows the ratio of UDP to UMP. For each substrate concentration, three independent enzymatic reactions were run using the same enzyme preparation. Error bars are SD.

## 3.2. Subcellular localization of the three UMKs

The subcellular localization of the three UMKs has not been assessed experimentally yet and prediction tools deliver inconclusive results. While cytosolic localization is suggested for UMK1, the results for UMK2 suggest cytosolic, mitochondrial or plastid localization and for UMK3 cytosolic or mitochondrial localization (Lange et al., 2008). To test the subcellular localization *in vivo*, constructs were generated containing either of the three *UMKs* fused with a gene encoding an mNeonGreen fluorescence protein for production of C-terminally fluorescence protein tagged UMKs. The fusion proteins were expressed in *Nicotiana benthamiana* and leaf discs of the infiltrated leaf areas examined with a confocal laser microscope. To assess the localization of the UMKs, mRuby marker proteins with known localizations were co-expressed. Example pictures and results of their respective Van Steensel cross correlation analysis (van Steensel et al., 1996) as well as the Pearson correlation coefficients for nine individual image sections are displayed in **Figure 7**. In a Van Steensel cross correlation analysis the two images of the protein of interest and the marker protein are aligned and shifted by a certain amount of pixels as indicated on the x-axis by the dX value. The amount of signal overlap is then calculated and given for each individual overlay with 1 being the highest value for a perfect overlay of the signals from both images. A good overlap is usually indicated by the fact that the highest value is associated with dX=0, i.e. both images have not been shifted. The value noted for dX=0 also represents the Pearson correlation coefficient. The UMK1-mNeonGreen

fusion protein was co-expressed with a cytosolic mRuby marker and it could be confirmed that the protein localizes to the cytosol (**Figure 7 A**). The same result was obtained with UMK3 (**Figure 7 C**). UMK2 showed an overlap with a mitochondrial mRuby marker and no cytosolic or plastid localization was apparent (**Figure 7 B**). The final step of UMP *de novo* synthesis takes place in the cytosol (Witz et al., 2012). UMK3, which showed a significantly higher catalytic efficiency with UMP than UMK1 and UMK2, also locates to the cytosol and could therefore catalyze further phosphorylation of the *de novo* synthesized UMP. Interestingly, UMK1 is also cytosolic, which raises the question of a possible redundancy of the two enzymes. The mitochondrial localization of UMK2 hints towards an involvement of it in mitochondrial nucleotide metabolism. Mitochondria require NTPs to fuel their transcriptional machinery and DNA replication. UMK2 would enable the production of those NTPs directly where they are needed. However, it is unclear in which form Arabidopsis mitochondria can import nucleosides and/or nucleotides. Nevertheless, a mitochondrial UMK should assist in the phosphorylation of salvaged nucleosides or recycled NMPs that are generated within the organelle.
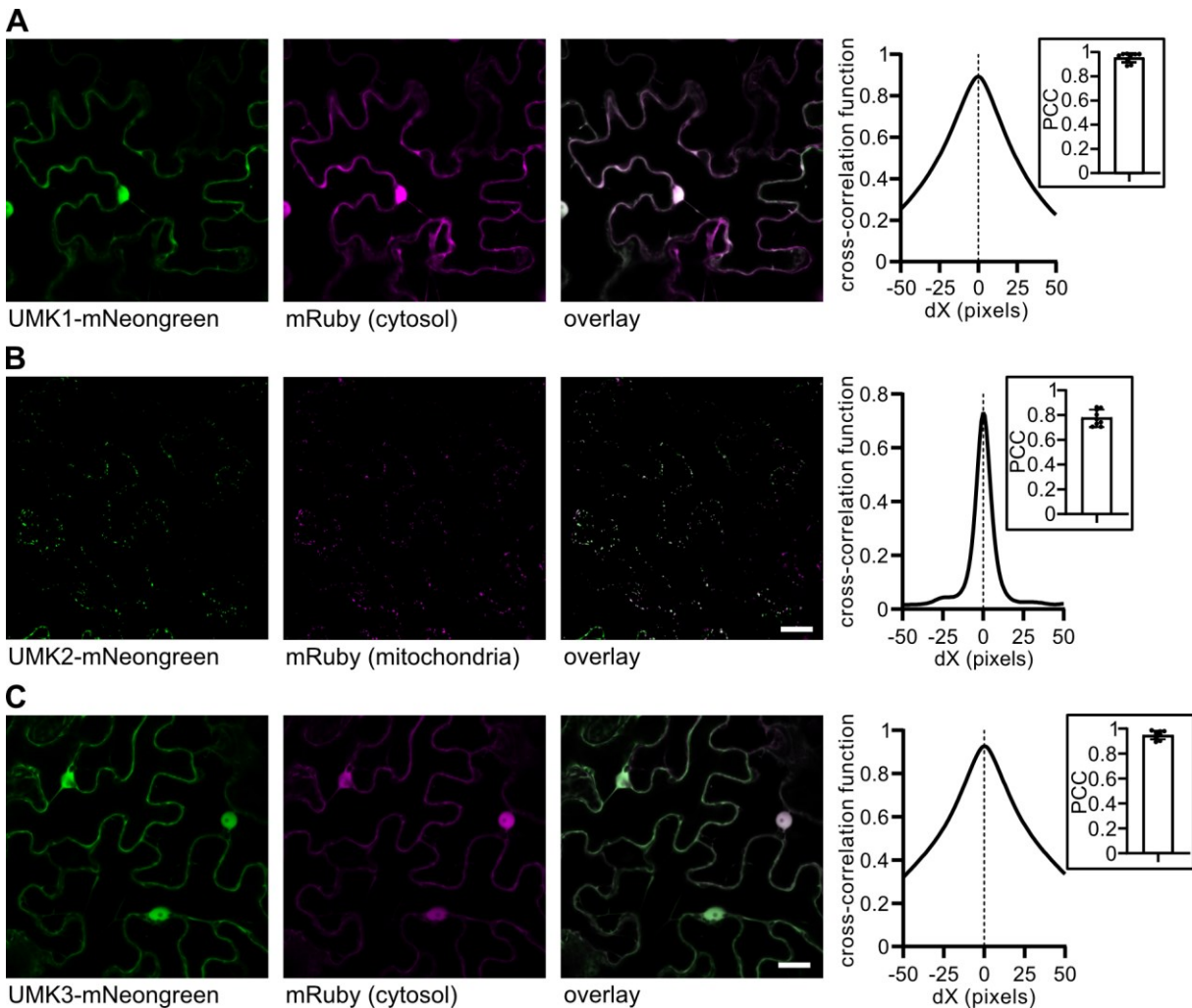
**Figure 7. Subcellular localization of the three UMKs.**
**A)** Confocal microscope images of a *Nicotiana benthamiana* leaf transiently co-expressing C-terminal mNeonGreen-tagged UMK1 (first left panel), cytosolic mRuby3 (second panel). The third panel shows an overlay of the mNeonGreen and mRuby3 images. Scale bar, 25 µm. Fourth panel, quantitative image analysis using the Van Steensel Cross Correlation Function with the images shown in the first and second panels. The inset in the fourth panel shows the Pearson Correlation Coefficients at dX = 0 for nine independent fluorescence images of different cells. **B)** as A, but for UMK2. In this case a mitochondrial variant of mRuby3 was used as subcellular marker. **C)** as A, but for UMK3.

## 3.3. Generation of *UMK* mutant plants using the CRISPR/Cas9 system

For characterization of *UMK* mutant plants, T-DNA insertion lines from the collection of the SALK Institute for Biological Studies and GABI-Kat of the University Bielefeld were analyzed. For *UMK1* the line SALK_092377, which supposedly carried a T-DNA insertion in the fifth exon of *UMK1*, was chosen. For *UMK2* the line GK723G02, which had a T-DNA insertion in the first exon of *UMK2*, was ordered. For *UMK3* there was no T-DNA line with an insertion inside an exon available. When genotyping the *UMK1* line

SALK_092377 it quickly became apparent that the T-DNA insertion site was not in the *UMK1* locus *At3g60180*, but in the locus *At3g60961*, which encodes the N-terminal truncated protein that was identified as likely non-functional in the phylogenetic analysis. As both loci are quite similar around the T-DNA insertion site and only distinguishable by a few single nucleotide polymorphisms, there was likely a mistake in the automatic annotation during the initial analysis of the line by SALK. Further characterization of the *UMK2* line GK723G02 also revealed contradicting results. It was possible to successfully identify homozygous T-DNA insertion lines, which were phenotypically distinct from heterozygous or wild type plants. New developing leaves of homozygous mutant plants were yellow and the plants grew slower (**Figure A 2**). It was attempted to complement the phenotype by transforming the plants with constructs that contained a *UMK2* transgene expressed by its native promoter. Several transgenic lines were generated harboring the T-DNA from the complementation construct, but the observed phenotype persisted in all lines. Further investigations revealed that a mutation in the neighboring locus At4g25270 causes the exact same phenotype (Chateigner-Boutin et al., 2011). It was therefore concluded that the T-DNA insertion in GK723G02 likely also affects At4g25270. As this strong phenotype would probably mask phenotypes caused by a mutation of *UMK2* and generally could affect further experiments, it was decided to discontinue work with that line.

Instead, it was decided to generate mutant lines using the CRISPR/Cas9 system, which can be used for the induction of base deletions or insertions (InDels) in the genome of plants. CRISPR/Cas9 induces double strand breaks in the genome, which are repaired by endogenous plant DNA repair systems. This can result in errors producing base pair InDels. When located in the open reading frame of a gene, InDels can cause a frameshift leading to an altered protein sequence and making the protein non-functional. The CRISPR/Cas9 system needs two components for the induction of a double strand break: a single guide RNA (sgRNA) and the Cas9 protein. The sgRNA can be specifically designed to target a gene of interest and guides the Cas9 protein to its target sequence where it induces a double strand break.

The sgRNAs for *UMK1* and *UMK3* were designed based on an internal ruleset proposed by Marco Herde in combination with an sgRNA efficiency prediction tool (Doench et al., 2014). For both genes two sgRNAs were selected for further work. An *in vitro* cleavage experiment with six sgRNAs was carried out for *UMK2* in order to test the efficiency of the sgRNAs experimentally in advance. As the *in vitro* cleavage assay was a novel technique in our laboratory, it was decided to also test all six sgRNAs *in vivo*, to assess whether the results were transferable to editing in plants. The structures of the three *UMK* loci with the chosen sgRNAs are depicted in **Figure 8**.
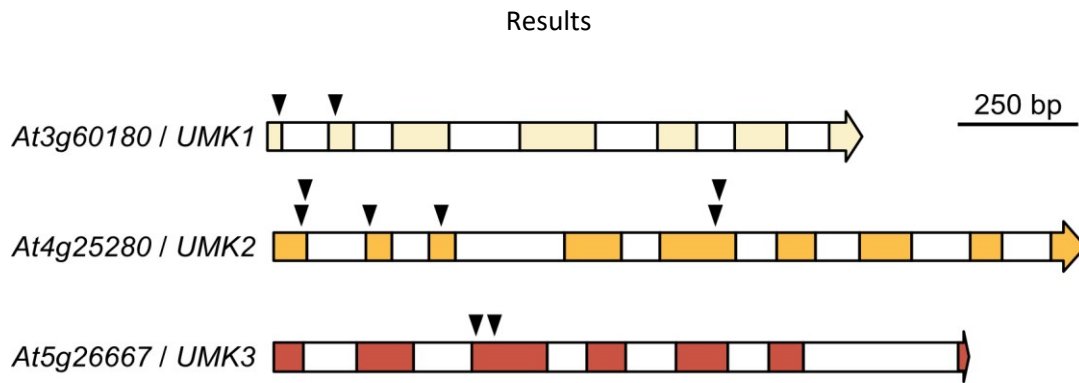
**Figure 8. Genomic organization of the *UMK* loci with sgRNA target sites.**
Genomic organization of the three *UMKs* from Arabidopsis. The coding regions of exons are represented by colored boxes while the introns are shown as white boxes. Arrows point to the sgRNA target sites.

### 3.3.1. Detection of editing events with a capillary sequencer

Since screening a large amount of plants for potential genome editing events is a laborious and cost-intensive procedure that requires a purification of PCR-amplified DNA fragments from an agarose gel and sending the fragments for out-of-house Sanger sequencing, a faster and cheaper workflow utilizing a capillary sequencer was established. A capillary sequencer is able to separate single stranded DNA fragments via gel electrophoresis in a thin capillary based on their size. The system was utilized for the screening of potential CRISPR/Cas9-induced mutations by comparing the size of a PCR amplicon from a wild type plant to the size of an amplicon from a potentially edited plant. The sequencer enables the differentiation of amplicons with a size difference of only one base pair, making the system perfectly suited for the detection of small InDels, as they typically occur in CRISPR/Cas9 experiments. The wild type amplicons can be distinguished from the sample amplicons by the use of differently dye-labelled primers in the PCR. The capillary sequencer has different detectors for the dyes and generates a chromatogram showing the size of the amplicons. Here, primers with the green dye 4-5-Dichlorocarboxyfluorescein (JOE) were used to generate the wild type amplicons and primers with the blue dye 6-Carboxyfluorescein (FAM) for generation of amplicons from potentially edited plants (**Figure 9**). Overall, the method is cheaper and faster than the previous workflow, in which PCR-amplified DNA fragments had to be purified from an agarose gel and sent out-of-house for Sanger sequencing. The results are also easier and quicker to interpret as the exact size of the InDels are directly visible compared to the laborious examination of the chromatograms of Sanger sequencing.
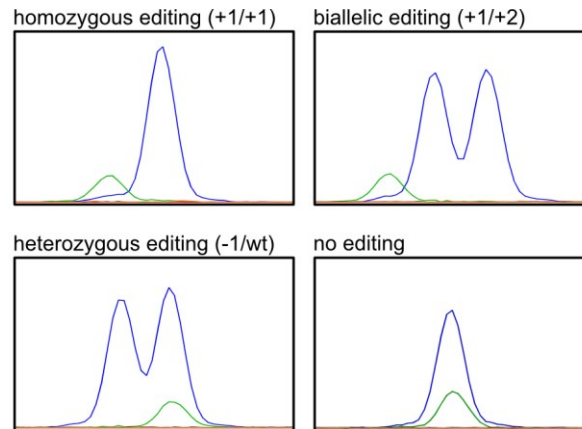
**Figure 9. Capillary sequencing chromatograms showing examples for different editing results.**
Green peaks correspond to fragments generated by PCRs with genomic wild type DNA, blue peaks originate from DNA of edited plants. The first panel shows an example of homozygous editing with both edited alleles carrying a one-base-pair insertion, the second panel shows biallelic editing with a one- and two-base-pair insertion, the third panel shows heterozygous editing with a one-base-pair deletion and a wild type allele, the last panel shows the result for an unedited plant.

### 3.3.2. Generation of *UMK1* mutant lines

*Arabidopsis thaliana* Col-0 were transformed with *Agrobacterium tumefaciens* carrying the constructs H773 or H774, each expressing an sgRNA targeting *UMK1*, by floral dip. Seeds of the transformed plants were harvested and transgenic seeds that glowed green due to the napinA::GFP cassette on the construct were selected. A total of 15 glowing seeds could be identified for the H773 transformation and 13 for the H774 transformation. The glowing seeds were germinated on soil and the plants genotyped for editing events using the capillary sequencer. Of the 15 H773 plants, four plants with editing events and of the 13 H774 plants one edited plant could be identified (**Table 4**).

**Table 4. Editing events detected in T1 plants transformed with the two *UMK1*-targeting CRISPR/Cas9 constructs H773 and H774.**

| Plant | Type of editing | Allele 1 | Allele 2 |
|---|---|---|---|
| H773 #01 | biallelic | insertion 1 bp | deletion 45 bp |
| H773 #04 | biallelic | insertion 1 bp | insertion 3 bp |
| H773 #10 | heterozygous | deletion 14 bp | wild type |
| H773 #15 | homozygous | insertion 1 bp | insertion 1 bp |
| H774 #11 | biallelic | insertion 1 bp | deletion 41 bp |

The seeds of all plants harboring editing events were harvested. These T2 seeds were re-analyzed for GFP fluorescence to identify seeds that no longer contained the T-DNA, as indicated by the absence of green fluorescence. Since the T1 plants should carry a heterozygous T-DNA integration, the transgene should be absent in 25% of the seeds of the T2 generation if only one integration of the T-DNA has taken place. It is also possible that multiple T-DNA integrations occurred, in which case a lower amount of non-glowing seeds was expected (6.25% for two integrations, 1.56% for three integrations, etc.). It was possible to find several non-glowing seeds for the plants H773 #04, H773 #10 and H774 #11. For H773 #01 and H773 #15 all T2 seeds were still glowing, indicating a high number of T-DNA integrations. This is not surprising, as it is quite common for the floral dip method to yield transgenic seeds carrying multiple T-DNA insertions (Buck et al., 2009). At least five non-glowing seeds of H773 #04, H773 #10 and H774 #11 were germinated on soil and the developing plants again genotyped. It was possible to identify a homozygous plant carrying a one-base-pair insertion on both alleles in the offspring of H773 #04, as well as a homozygous plant also carrying a one-base-pair insertion on both alleles in the offspring of H774 #11. The results were confirmed by Sanger sequencing and the H773 mutant line named *umk1-1* and the H774 mutant line *umk1-2* (**Figure 10**). Absence of the T-DNA was additionally confirmed by PCR (**Figure A 3**). All further experiments involving *UMK1* mutant plants were carried out with the offspring of these two lines.

**Figure 10. CRISPR/Cas9-induced mutations in *UMK1*.**
**A) + B)** Sanger sequencing chromatograms of *UMK1* mutant (left panel) and wild type (right panel) alleles. The inserted bases are indicated with arrows in the left panel. **C)** Genomic organization of the *UMK1* locus as in **Figure 8**. In addition, the sgRNA target sequences and edited allele sequences are shown in the boxes. Inserted nucleotides are written in red, protospacer adjacent motif (PAM) sequences are underlined.

### 3.3.3. Generation of *UMK2* mutant lines

An *in vitro* cleavage assay was performed to identify the most promising sgRNA candidates targeting *UMK2*. Each sample of the assay contained the linearized pJet construct with the *UMK2* gene (H522 cut with *Bsa*I), one of the six *in vitro* transcribed sgRNAs and Cas9 protein. The samples were incubated for 15 minutes, in which the target DNA is cut by the sgRNA/Cas9 complex with varying efficiencies depending on the properties of the sgRNA. The samples were then loaded onto an agarose gel and separated by gel electrophoresis to visualize the results. The signal strength of linearized construct indicates the cutting efficiency of the respective Cas9/sgRNA complexes. The lower the signal intensity, the more DNA was cut. The two sgRNAs in the samples seen in lane 1 and 6 showed the highest cutting efficiency in the assay, the ones in lane 4 and 5 showed a medium efficiency and for the other two almost no cleaving activity was visible (**Figure 11**). Lane 7 contains a sample without sgRNA as a control without any cleaving activity.
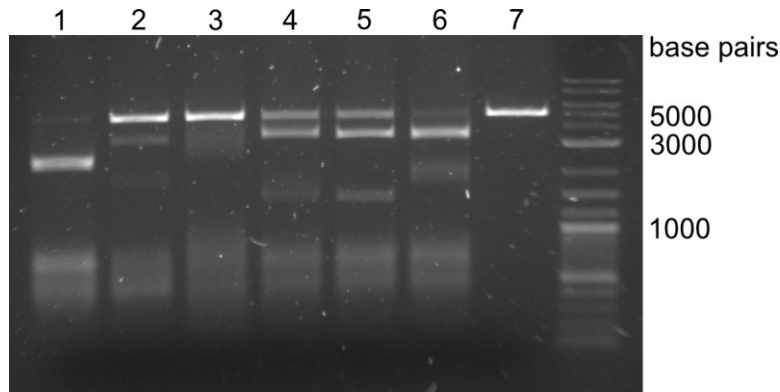
**Figure 11. *In vitro* cleavage assay with six sgRNAs targeting *UMK2*.**
*In vitro* cleavage assay reactions visualized on an agarose gel. Each reaction contained linearized H522 construct, Cas9 protein and an *in vitro* transcribed sgRNA targeting *UMK2*. Lane 1 to 6 contains reactions with six different sgRNAs (target sites indicated in **Figure 8**), lane 7 contains a control without sgRNA, lane 8 contains a DNA size standard (1 kb Plus DNA ladder, New England Biolabs).

A construct was generated for each of the six tested sgRNAs and Arabidopsis plants transformed via floral dip. The constructs were H1130, H1131, H1132, H1133, H1134 and H1135 in order of the sgRNAs in the assay. Between seven and 14 glowing seeds were found for each transformation for a total of 56 glowing seeds. The seeds were germinated on soil and all plants genotyped using the capillary sequencer. The overall editing rate at the *UMK2* locus was lower than for *UMK1*. Only three edited plants were identified from the 56 analyzed plants (5.4 % with editing). In comparison, five of 28 plants showed editing in the *UMK1* locus (17.6 % with editing). While this lower editing efficiency for *UMK2* may be related to the selected sgRNAs, it is also possible that other factors, such as lower accessibility of the *UMK2* locus for the sgRNA/Cas9 complex due to chromatin structure, reduced the amount of detected editing events. The detected editing events in *UMK2* are displayed in **Table 5**. It was possible to identify two editing events in plants that were transformed with H1133, which expresses the medium efficient sgRNA4 and one for H1135, which expresses the highly efficient sgRNA6. No editing was detected for any of the low efficient sgRNAs, but the same is true for the highly efficient sgRNA1. Overall, the results are inconclusive with regard to the transferability of the *in vitro* cleavage assay to the *in vivo* situation, as the editing rate in the *UMK2* locus was too low. To obtain a clearer picture, it would be necessary to either screen more T1 seeds or screen the T2 generation of some of the transformed plants.

**Table 5. Editing events detected in T1 plants transformed with the two *UMK2*-targeting CRISPR/Cas9 constructs H1133 and H1135.**

| Plant | Type of editing | Allele 1 | Allele 2 |
|---|---|---|---|
| H1133 #11 | heterozygous | insertion 1 bp | wild type |
| H1133 #12 | heterozygous | deletion 1 bp | wild type |
| H1135 #02 | homozygous | insertion 1 bp | insertion 1 bp |

The seeds of the three edited plants were harvested. At least eight non-glowing seeds were identified for each line, germinated on soil and the plants again genotyped. A homozygous plant carrying a one-base-pair insertion on both alleles was identified in the offspring of H1133 #11. In the offspring of H1135 #02, all plants carried a one-base-pair insertion. The progeny of H1135 #02 were genotyped by Sanger sequencing, which revealed that they segregate for the insertion of an adenine or thymine nucleotide. For further experiments a plant with a homozygous adenine nucleotide insertion was chosen. The line originating from the H1135 transformation was named *umk2-1* and the line from the H1133 transformation *umk2-2* (**Figure 12**). Again, the absence of T-DNA was confirmed by PCR (**Figure A 3**) and all further work involving *UMK2* mutant plants was carried out with these two lines.
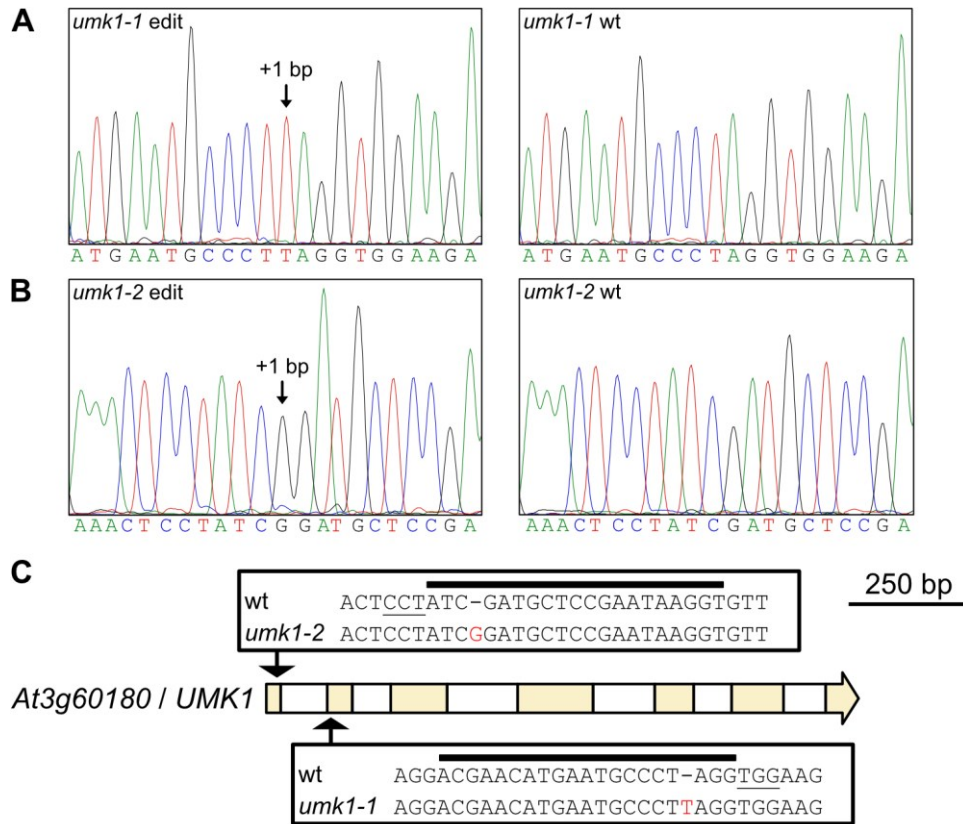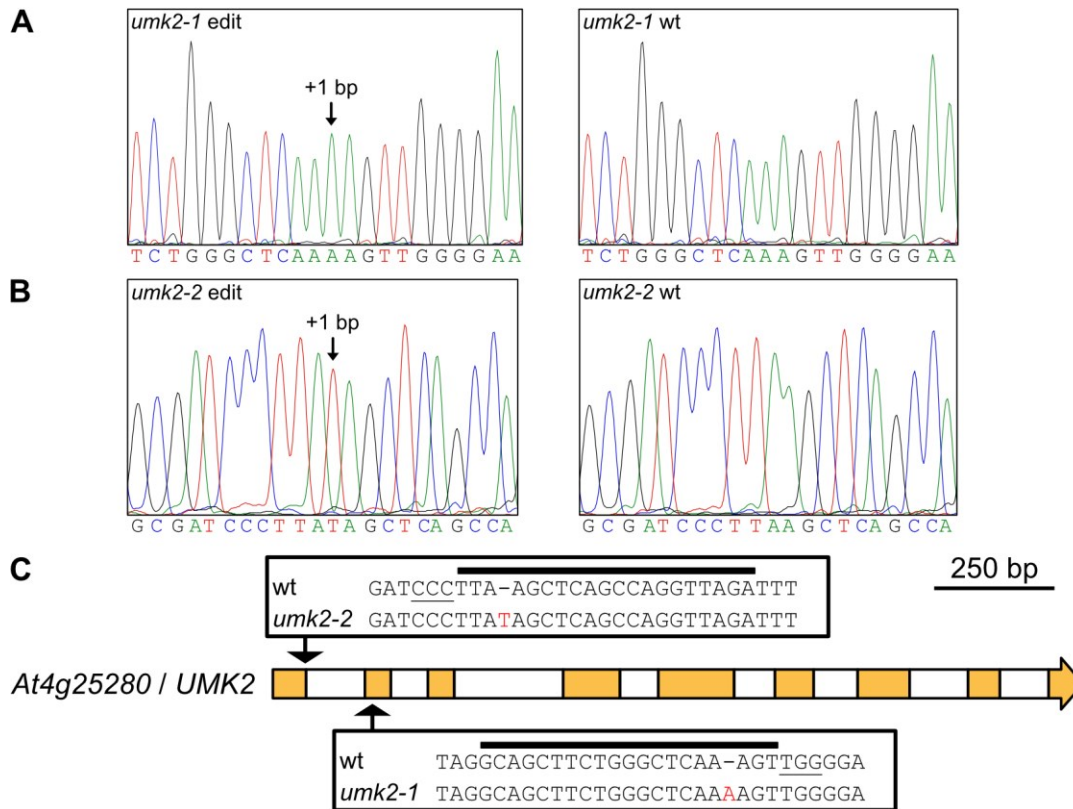
**Figure 12. CRISPR/Cas9-induced mutations in *UMK2*.**
**A) + B)** Sanger sequencing chromatograms of *UMK2* mutant (left panel) and wild type (right panel) alleles. The inserted bases are indicated with arrows in the left panel. **C)** Genomic organization of the *UMK2* locus as in **Figure 8**. In addition, the sgRNA target sequences and edited allele sequences are shown in the boxes. Inserted nucleotides are written in red, PAM sequences are underlined.

### 3.3.4. Generation of *UMK3* mutant lines

To generate *UMK3* mutants, Arabidopsis plants were transformed by floral dip with Agrobacteria carrying either the construct H982 or H983. The seeds of the transformed plants were harvested and eleven glowing seeds from the H982 transformation and three glowing seeds from the H983 transformation were collected. In contrast to the other two genes, no editing in the *UMK3* locus was detected in any of the T1 generation plants grown from these seeds. It was decided to investigate the T2 generation for genome editing. The gene encoding the Cas9 is expressed by an egg cell-specific promoter. It is therefore more likely for editing events to occur in the T2 generation where the system has more time for the induction of editing events compared to the T1 generation. One line of each of the two transformations was chosen to be analyzed in the T2 generation. The chosen lines had strongly glowing seeds to ensure a strong expression of the transgene and some non-glowing seeds in between indicating a low amount of T-DNA

insertions to facilitate selection of transgene-free plants later. Interestingly, there were also many dark and deformed seeds which only exhibited a faint glowing and did not germinate (**Figure 13**).
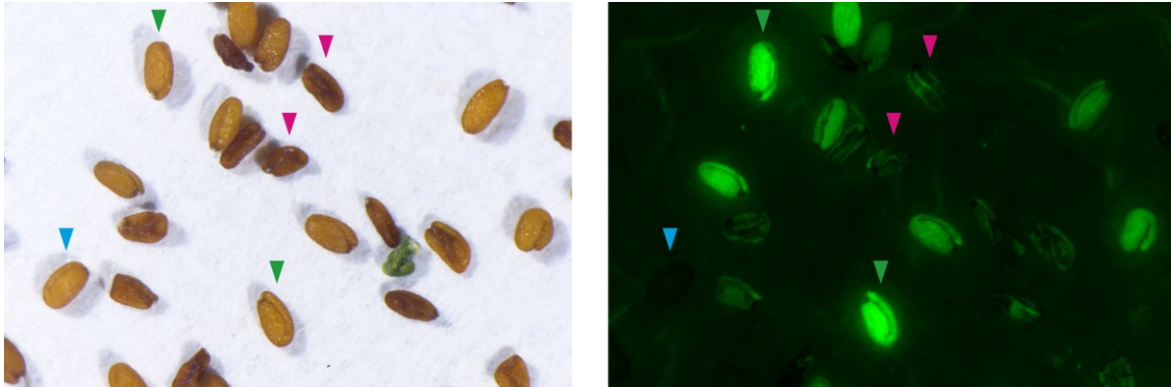


**Figure 13. Seeds of a T2 plant transformed with the *UMK3*-targeting CRISPR/Cas9 vector H982.**
The left panel shows a bright field image. The right panel shows the same image section taken through a Nikon P2-EFL GFP-L filter block to visualize GFP fluorescence. Green arrows indicate fluorescing and normal looking seeds, magenta arrows indicate dark and deformed seeds, the blue arrow indicates a non-glowing and likely non-transgenic seed.

24 glowing seeds of each line were selected, germinated on soil and genotyped. For the H982 transformation, a variety of different editing events was detected. In total, eleven of the 24 analyzed plants had InDels. Four edited plants were found in the H983 transformation. The detected editing events in five of the H982 plants which were chosen for further analysis and the four H983 plants are displayed in **Table 6**. Interestingly, no edited plant with a frameshift mutation on both alleles was detected. In the H982 line which showed editing in almost half of the analyzed plants a three-base-pair deletion, which does not result in a frameshift, was observed multiple times. For the H983 line, only one-base-pair insertions in combination with a wild type allele were detected.

**Table 6. Editing events detected in T2 plants transformed with the two *UMK3*-targeting CRISPR/Cas9 constructs H982 and H983.**

| Plant | Type of editing | Allele 1 | Allele 2 |
| --- | --- | --- | --- |
| H982 #02 | biallelic | insertion 1 bp | deletion 3 bp |
| H982 #03 | biallelic | deletion 10 bp | deletion 3 bp |
| H982 #04 | heterozygous | deletion 3 bp | wild type |
| H982 #08 | biallelic | insertion 1 bp | deletion 3 bp |
| H982 #12 | heterozygous | insertion 1 bp | wild type |
| H983 #15 | heterozygous | insertion 1 bp | wild type |
| H983 #21 | heterozygous | insertion 1 bp | wild type |
| H983 #22 | heterozygous | insertion 1 bp | wild type |
| H983 #24 | heterozygous | insertion 1 bp | wild type |

Seeds of the plants harboring editing events were harvested. In a regular segregation pattern, it should be possible to obtain plants with a homozygous frameshift mutation in the T3 generation, but the opposite was the case when the plants were genotyped. 16 of 16 plants originating from H982 #02, #03 and #08 were homozygous for the three-base-pair deletion and all frameshift mutations were completely eliminated from the population. Also backcrossing with wild type plants only yielded heterozygous plants carrying the three-base-pair deletion in combination with a wild type allele. A similar result was observed in the T3 generation of the heterozygous H983 lines. Of 20 analyzed plants, 18 were homozygous wild type plants and only two carried a heterozygous one-base-pair insertion. Since these plants were still transgenic, it is possible that these two mutations were newly induced by the still active CRISPR/Cas9 system. These results suggest that it is not possible to generate homozygous frameshift mutations in UMK3 or even to transfer frameshift mutations to the next generation.

**Figure 14 A** shows a silique of plant H982 #08, which was biallelic for a three-base-pair deletion and a one-base-pair insertion in the *UMK3* locus. It is apparent that about half of the silique is empty and many seeds were aborted during early development, possibly due to the frameshift mutation. The same was observed for the heterozygous plants obtained from the H983 transformation (a silique of the plant H983 #15 is shown, **Figure 14 B**), but not for a wild type plant grown in parallel (**Figure 14 C**). The plants that germinated from the viable seeds of H982 #02 all carried the homozygous three-base-pair deletion, indicating that this mutation can be transferred to the following plant generation. When examining the dried seeds, they appear normal and no dark and deformed seeds can be seen in between (**Figure 14 D**).

In comparison, some of the seeds harvested from plant H982 #12, which had a one-base-pair insertion and a wild type allele, are dark and deformed again (**Figure 14 E**).



**A** Silique H982 #08 (+1/-3 bp)
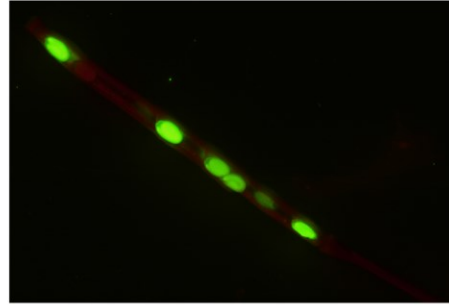
**B** Silique H983 #15 (+1/wt)
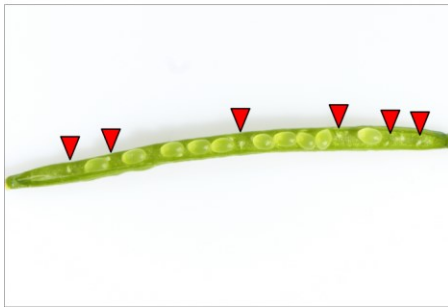
**C** Silique Col-0

**D** Seeds H982 #08 (+1/-3 bp)
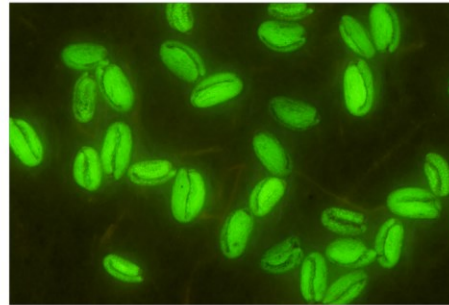
**E** Seeds H982 #12 (wt/+1 bp)
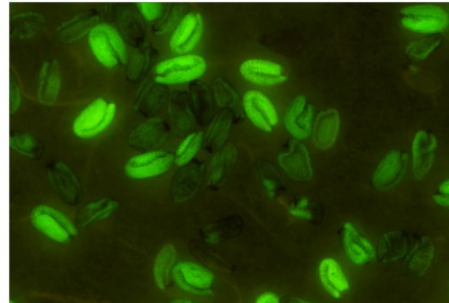
**Figure 14. Siliques and seeds of *UMK3* mutants and a wild type plant.**
Green siliques were opened on one side. Red arrows indicate aborted seeds. **A)** Silique of a T2 plant obtained from transformation with construct H982. The mother plant was a biallelic mutant with one allele

carrying a one-base-pair insertion and the other a three-base-pair deletion. Most of the viable seeds were still fluorescent indicating the presence of the transgene (right panel). **B)** Silique of a T2 plant obtained from transformation with construct H983. The mother plant was heterozygous having one wild type allele and one allele with a one-base-pair insertion. **C)** Silique of a wild type mother plant that had been grown together with the mutants. **D)** T3 seeds harvested from a T2 plant obtained from transformation with H982. The mother plant was heterozygous having one wild type allele and one allele with a one-base-pair insertion. **E)** As D, but the mother plant was biallelic having one allele with a three-base-pair deletion and one allele with a one-base-pair insertion.

As the H983 transformation did not yield any viable plants harboring editing events for further experiments, it was decided to only continue with the homozygous three-base-pair deletion line obtained from the H982 transformation. It was possible to select non-transgenic seeds of the H982 line which were analyzed by Sanger sequencing to determine the missing three base pairs (**Figure 15**). It turned out that the missing nucleotides result in the deletion of the glutamic acid on position 76 (E76) of the translated protein. The line was therefore named *umk3$_{ΔE76}$*. Absence of the T-DNA was confirmed by PCR (**Figure A 3**).
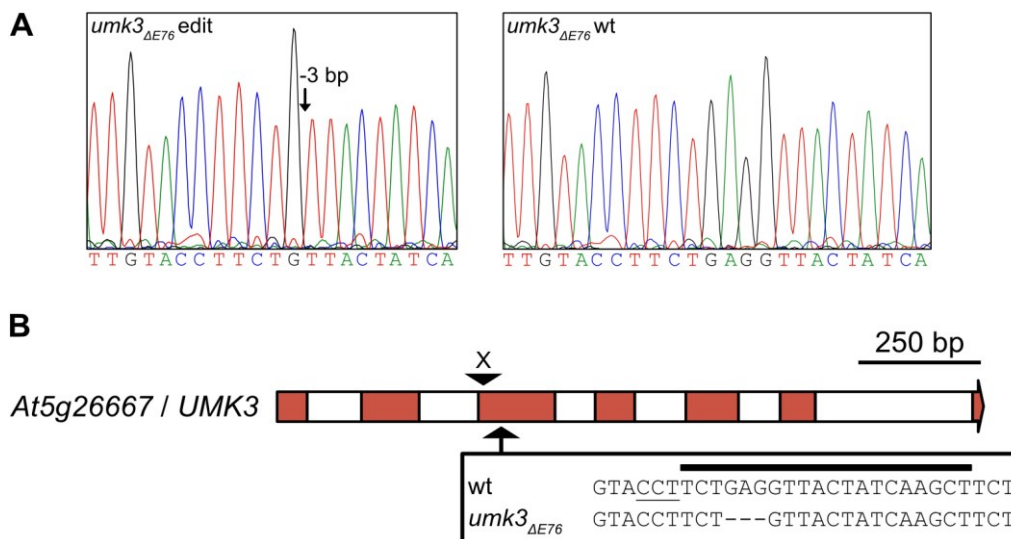


**Figure 15. CRISPR/Cas9-induced mutation in *UMK3*.**
**A)** Sanger sequencing chromatogram of the *UMK3* mutant (left panel) and wild type (right panel) allele. The location of the deleted bases is indicated with an arrow in the left panel. **B)** Genomic organization of the *UMK3* locus as in **Figure 8**. In addition, the sgRNA target sequence that led to the viable three base pair deletion allele and the edited allele sequence are shown in the box. The PAM sequence is underlined. The X indicates that no viable alleles could be generated with the sgRNA targeting this site.

As the three-base-pair deletion in the *UMK3* gene does not result in a frameshift and only causes the deletion of one glutamic acid in the translated protein, it is not clear whether this would impact the enzymatic activity of UMK3. To investigate this, *UMK3* and *UMK3$_{\Delta E76}$* fused with a sequence encoding a C-terminal Strep-tag were expressed in *Nicotiana benthamiana* from the constructs H554 and H1326, respectively, and the proteins purified via StrepTactin affinity chromatography. Success of the purification was confirmed by Coomassie stain (**Figure 16 A**). Enzymatic activity was measured using the coupled enzyme assay with 100 µM UMP, 100 µM CMP or 500 µM dCMP in the presence of 1 mM ATP (**Figure 16 B**). A lower activity could be observed for UMK3$_{\Delta E76}$ with all three substrates in comparison to UMK3. The specific activity was decreased 1.9-fold with UMP and 4.3-fold with CMP. For dCMP almost no activity could be measured. This means that it was possible to generate an impaired version of UMK3, which is not as active as the intact enzyme, but active enough to sustain the plant. For further investigation of the induced mutation, the amino acid sequence of UMK3$_{\Delta E76}$ was aligned with the sequences of the Arabidopsis UMKs and UMKs from other organisms (**Figure 16 C**). Except for the slime mold *Dictyostelium discoideum*, E76 is conserved in all analyzed organisms. The area around the missing glutamic acid is also highly conserved and it is located between the isoleucine at position 72 (I72) and threonine at position 78 (T78) which line the active site and are responsible for NMP binding (Yan and Tsai, 1999). To examine the influence on the secondary structure of the protein, the structures of UMK3$_{\Delta E76}$ and UMK3 were predicted using the AlphaFold ColabFold v1.5.2 webserver and aligned in PyMol (**Figure 16 D**). E76 is predicted to be located at the beginning of an alpha helix and deletion of the amino acid tightens the helical structure. In the model, this appears to have no influence on the position of the side chain of I72 and only slightly moves that of T78. Other amino acids in proximity to the NMP binding pocket, like S75 are moved more drastically. S75 could be involved the binding the water molecule, which is required for activity with pyrimidine NMPs (Müller-Dieckmann and Schulz, 1995; Scheffzek et al., 1996). Thus, it is possible that the shift of S75 causes the lowered enzymatic activity observed in the assay. However, this is only speculative, and determining the exact cause of the reduced UMK3$_{\Delta E76}$ activity is beyond the scope of this work. This coincidentally generated line turned out to be fortunate for the course of the project. As a full knockout of *UMK3* could not be obtained, this line still allowed the analysis of effects caused by partially compromised UMK3. Overall, the results support the hypothesis that UMK3 is the key enzyme in pyrimidine nucleotide metabolism. While the generation of *UMK1* and *UMK2* null-mutant lines was possible, null-mutation of *UMK3* is apparently lethal for the plant.
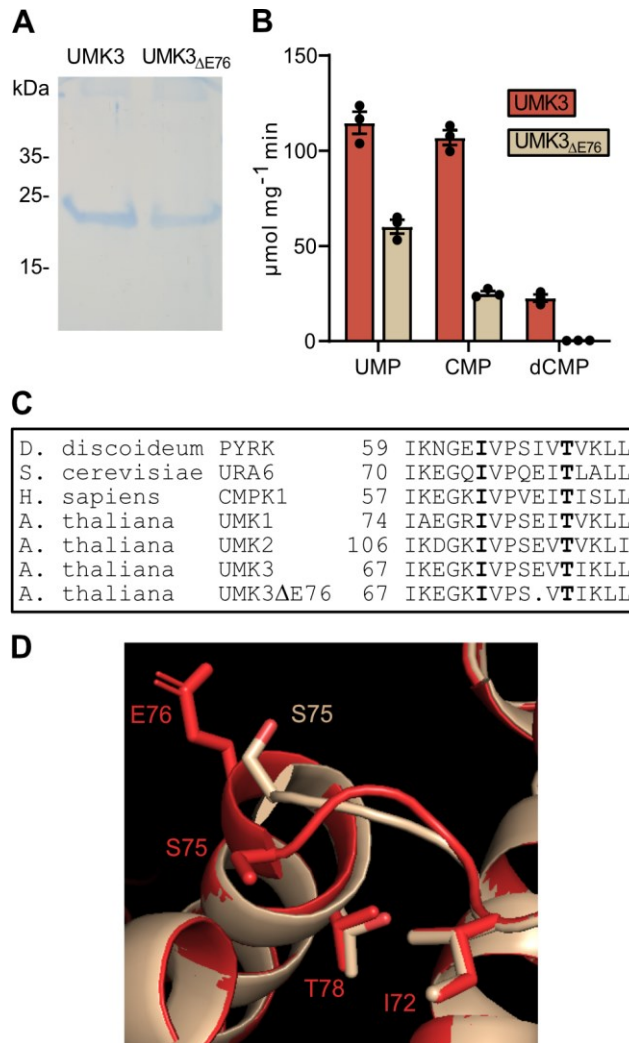
**Figure 16. Analysis and biochemical properties of UMK3$_{\Delta E76}$.**
**A)** Affinity purified UMK3 and UMK3$_{\Delta E76}$ (Coomassie-stained SDS gel). The C-terminal Strep-tagged proteins were transiently expressed in *Nicotiana benthamiana* and purified by Strep-Tactin affinity chromatography. 20 µl of the affinity elution fractions were loaded corresponding to 66 ng and 30 ng protein for UMK3 and UMK3$_{\Delta E76}$, respectively. **B)** Specific activities of UMK3 and UMK3$_{\Delta E76}$ with UMP, CMP (100 µM) and dCMP (500 µM) in the presence of 1 mM ATP. Three independent enzymatic reactions were run per substrate, error bars are SD. **C)** Sequence alignment of UMK3$_{\Delta E76}$ with UMKs from Arabidopsis and non-plant organisms near E76. I72 and T78, which line the active site, are shown in bold. **D)** Structure prediction of UMK3 (red) and UMK3$_{\Delta E76}$ (beige) with Alphafold2 ColabFold v1.5.2 visualized and aligned in PyMol v2.5.5. Both predicted structures had an average pLDDT score above 90, including around the E76 deletion site, indicating high confidence in the prediction.

## 3.4. Phenotypical characterization of the *UMK* mutants and crosses

All three UMKs are able to utilize the same substrates, and both UMK1 and UMK3 are localized in the cytosol. To assess the redundancy between the enzymes and to determine whether there are additive

effects when a plant lacks the function of several UMKs, double mutant lines and a triple mutant line were generated. The following crosses were made to generate the double mutants: *umk1-1* x *umk2-1*; *umk1-1* x *umk3$_{ΔE76}$* and *umk2-1* x *umk3$_{ΔE76}$*. For the triple mutant *umk1-1 umk2-1* was crossed with *umk1-1 umk3$_{ΔE76}$*. All plants were genotyped with the capillary sequencer and a homozygous line could be generated for each cross. The crosses were also analyzed for absence of T-DNA by PCR (**Figure A 3**).

### 3.4.1. Germination rates of *UMK* mutant lines

The germination rates of the mutant lines were evaluated using *UMK* mutant seeds from a uniform seed batch. Three repetitions of at least 100 seeds per line were placed on filter paper soaked with half strength MS-medium in Petri dishes. The Petri dishes were incubated for 48 hours at 4°C in the dark and then moved to long-day growth conditions. After another 48 hours the seeds were counted and classified as germinated if the radicle was visible (**Figure 17**). Approximately 95 % of wild type seeds germinated within 48 hours. Similar germination rates of at least 92 % were observed for the lines *umk1-1*, *umk1-2*, *umk2-1*, *umk2-2*, *umk3$_{ΔE76}$* and *umk1 umk2*. A significantly lower germination rate could be observed for the *umk1 umk3$_{ΔE76}$* (86 %) and *umk2 umk3$_{ΔE76}$* (89 %) lines and in tendency for the triple mutant (90 %; P < 0.07). Overall, the mutation of *UMK3* in combination with null-mutation of one of the other two *UMKs* seems to reduce the seed germination rate. A dormant seed is physiologically not very active and accumulates cellular damage to macromolecules including lipids, proteins or DNA (Waterworth et al., 2015). Upon imbibition, cellular DNA repair mechanisms are initiated before cell division in the embryo starts, which require dNTPs. The lower germination rates observed in this experiment could be an indicator for a lower availability of dNTPs, more specifically dCTP, caused by the mutation of multiple *UMKs*. A lower availability of dCTP could hinder the DNA repair resulting in a delay or stoppage of germination.
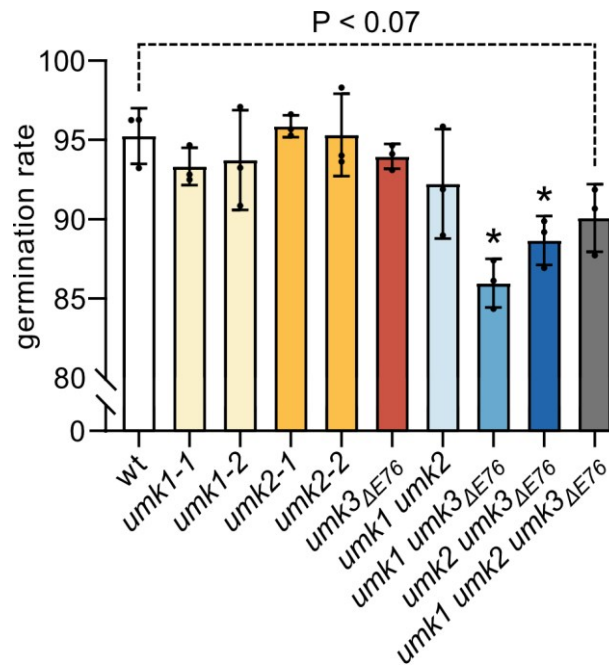
**Figure 17. Germination rates of the wild type and *UMK* mutant lines after 48 hours.**
Seeds were placed on filter paper soaked with half strength MS-medium, incubated for 48 hours at 4°C in the dark and then transferred to long-day growth conditions for 48 hours. Several mother plants of each genotype were grown in parallel under long-day conditions for seed production. Seeds of the same genotype were pooled and from each pool, three replicates of at least 100 seeds were evaluated. A seed was counted as germinated when the radicle was visible. Statistical analysis was performed by two-sided ANOVA coupled with Bonferroni posttest comparing to the wild type, asterisks indicate a statistical significance with P < 0.05. Error bars are SD.

### 3.4.2. Phenotypical characterization of *UMK* mutant seedlings

In addition to the analysis of germination rates, seedlings of all mutant lines were phenotypically characterized. 30 seeds of each line were placed on soil and imbibed for 48 hours at 4°C in the dark. The pots were then transferred to long-day growth conditions and the seedlings photographed after seven days. The leaf areas of the 15 seedlings that showed the best growth were quantified for each line using the Fiji extension of the ImageJ software (Schindelin et al., 2012). A representative seedling of each line whose leaf area corresponds to the median of the respective population is shown (**Figure 18 A**). No phenotypic abnormalities or growth disorders can be observed. Also the quantitative analysis of the leaf areas does not show any significant results, except for a significant difference between *umk1-1* and *umk2-2* with P = 0.017 (**Figure 18 B**). Because the plants of other lines like *umk1-2* and *umk2-1* do not show similar effects, this difference might have occurred by chance. Even the seedlings of the triple mutant, which would be most likely to show a growth defect if the mutations were to result in one, do not differ from the wild type. These results show, that loss of UMK1 or UMK2 or an impairment of UMK3 does not affect early

plant growth under the employed growth conditions. Also the combination of the three mutations causes no visible phenotypical abnormalities. It appears that the UMK3$_{\Delta E76}$ variant alone is able to sustain the pyrimidine nucleotide demand of the seedling. The fact that the *umk2* lines grow normally without their mitochondria having an active UMK for the phosphorylation of UMP, CMP or dCMP indicates that mitochondria are able to import pyrimidine nucleotides from the cytosol. The results slightly contradict the observations made for the germination rates (**Figure 17**), where a negative effect of the mutations on seed germination was observed. This may be explained by the fact that here the 15 seedlings that showed the best growth were used for leaf area quantification. Consequently, seeds that did not germinate or seedlings that showed a delayed development were not taken into account. Alternatively, it is possible that the *UMK* mutations do not influence the establishment of the seedling after germination.
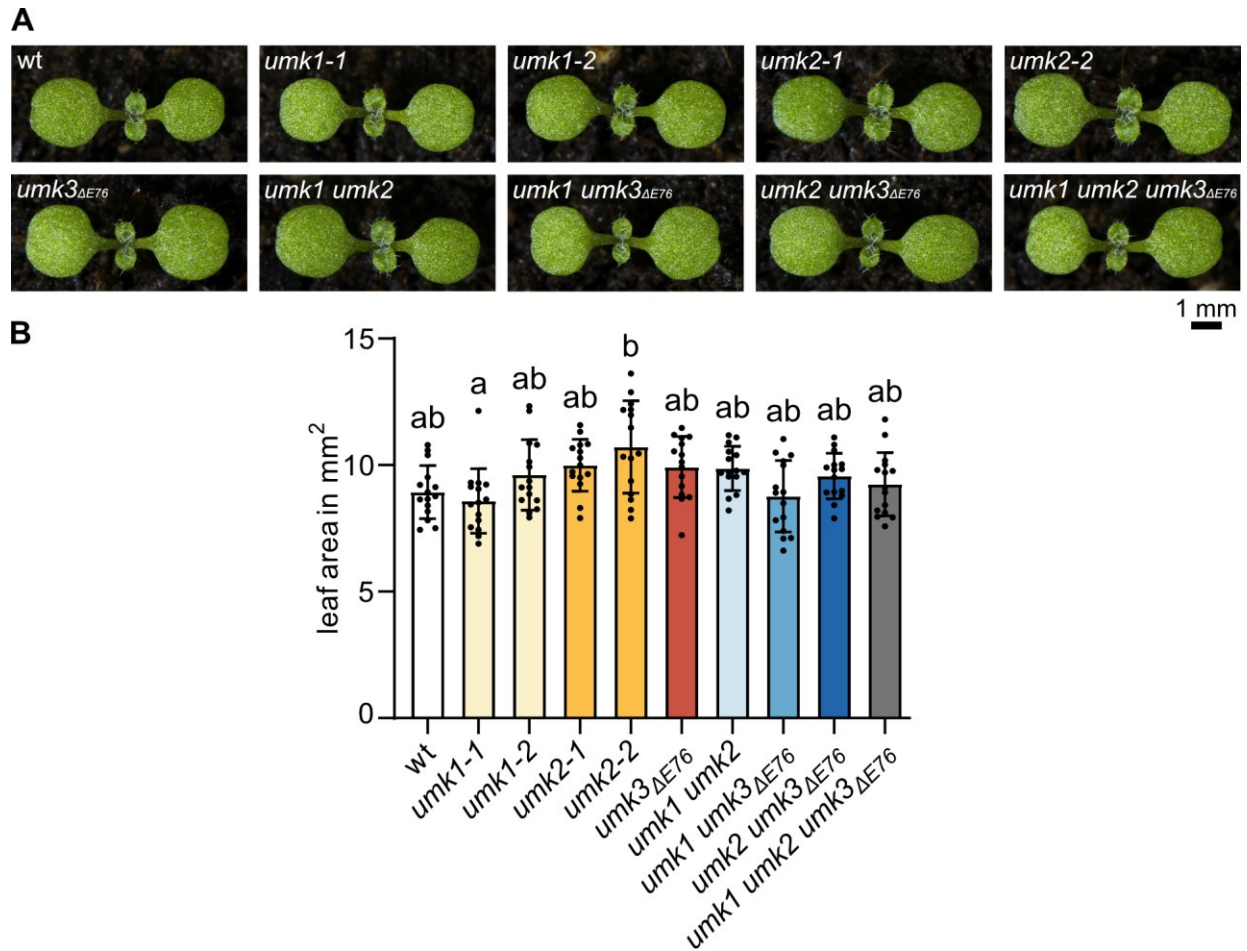


**Figure 18. Leaf area of seven-day-old *UMK* mutants and the wild type.**
Per genotype, 15 seedlings were analyzed. **A)** Representative seedlings. Of the 15 seedlings of each genotype, the plant with the median leaf area is shown. **B)** Quantified leaf areas. Statistical analysis was performed using two-sided Tukey's pairwise comparisons. Different letters indicate differences at $P < 0.05$. Error bars are SD.

### 3.4.3. Phenotypical characterization of 35-day-old *UMK* mutants

The leaf area quantification was repeated with 35-day-old plants (**Figure 19 A**). The leaf area was determined for five plants of each line. At this later growth stage, some phenotypical abnormalities can be associated with the *UMK3$_{\Delta E76}$* allele. While mutation of *UMK1* and/or *UMK2* does not appear to affect plant growth, the leaf morphology of the *umk3$_{\Delta E76}$* line and the double mutants containing the *UMK3$_{\Delta E76}$* allele is slightly different as the leaves appear a bit smaller and not as round. This indicates that normal vegetative plant growth can be maintained by UMK3 alone under the growth conditions used, but that partial impairment of UMK3 activity leads to a reduction in growth. This observation is confirmed by the quantification of the leaf area of these plants, as *umk3$_{\Delta E76}$* and the crosses containing *umk3$_{\Delta E76}$* have a reduced leaf area of around 20 % (**Figure 19 B**). The reduction is even more prominent in the triple mutant. This line shows a reduced leaf area of around 60 % compared to the wild type, indicating a strong influence of mutating all three *UMKs* on plant growth at later stages. This strong reduction does not occur in *umk1 umk3$_{\Delta E76}$* or *umk2 umk3$_{\Delta E76}$* plants, suggesting that UMK1 or UMK2 alone are able to partially complement the lower activity of UMK3$_{\Delta E76}$. The fact that plants of the *umk1 umk3$_{\Delta E76}$* line grow better than the triple mutant also shows, that mitochondria are likely also able to export pyrimidine nucleotides to the cytosol, where the synthesis of UDP-sugars required for biomass production takes place. The activity of UMK2 must be contributing to this production in the *umk1 umk3$_{\Delta E76}$* line, which requires an export of UDP or UTP from the mitochondria to the cytosol.
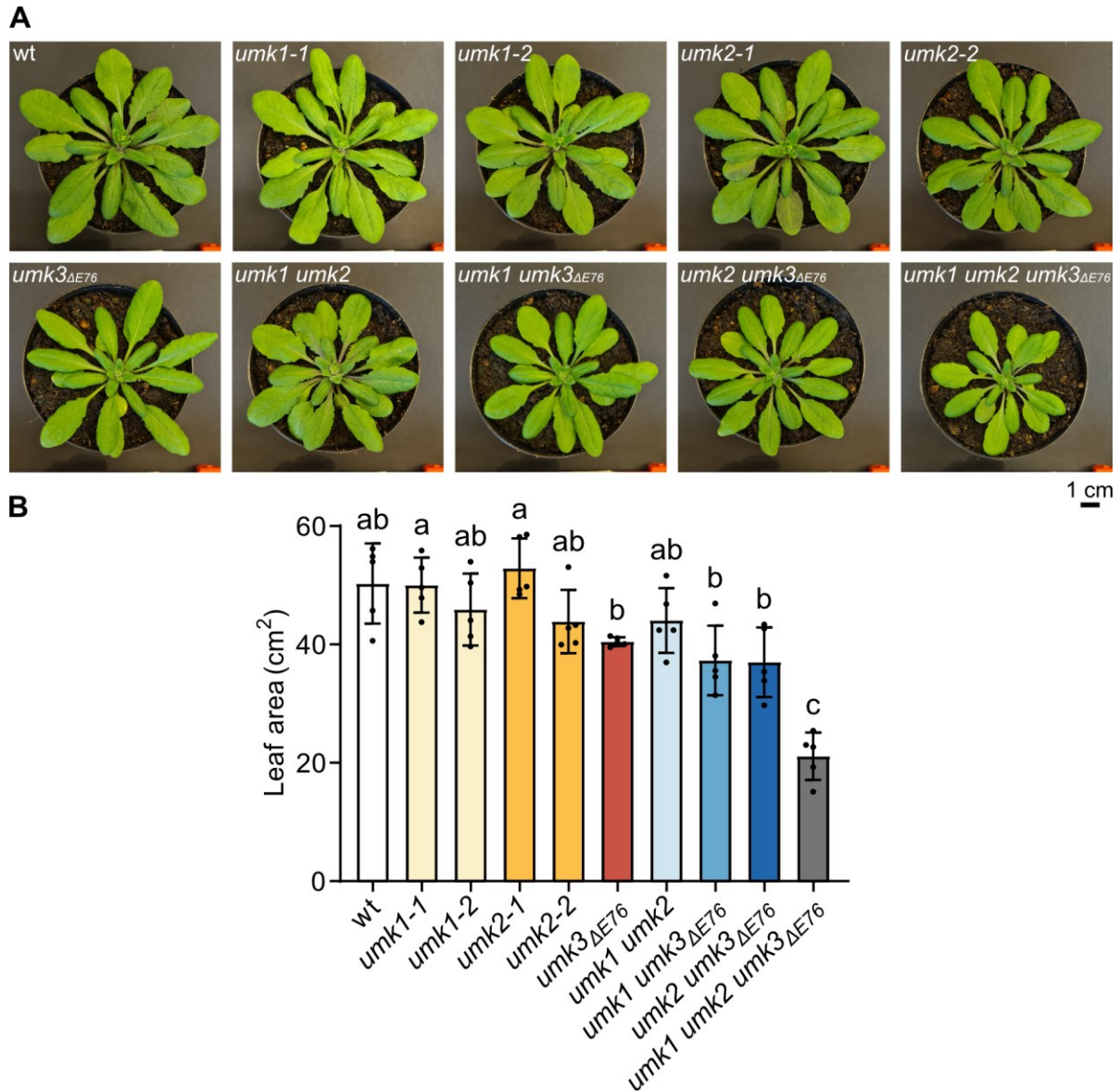
**Figure 19. Leaf area of 35-day-old *UMK* mutants and the wild type.**
Five plants were analyzed per genotype. **A)** Images of 35-day-old wild type and *UMK* mutant plants, the plant with the median leaf area is shown. **B)** Quantified leaf areas. Statistical analysis was performed using two-sided Tukey's pairwise comparisons. Different letters indicate differences at $P < 0.05$. Error bars are SD.

## 3.5. Metabolome analysis of *UMK* mutant plants

UMKs catalyze the transfer of a phosphate group from ATP to UMP, CMP or dCMP, resulting in ADP, as well as UDP, CDP or dCDP. The metabolome of the *UMK* mutant lines was therefore analyzed for alterations of nucleotide content with a focus on pyrimidine nucleotides. This analysis was made possible by the development of an optimized solid phase extraction method in our laboratory (Straube et al., 2021;

Straube et al., 2023). Plants of all *UMK* mutant lines were grown from a uniform seed batch on soil under long-day conditions for 18 days. Three whole plants without roots were harvested and pooled per sample with five biological replicates per mutant line. Following solid phase extraction, the metabolites were quantified by LC-MS analysis using an Agilent 6470 triple quadrupole mass spectrometer. Unfortunately, correct quantification of the UMP concentration was not possible with the used method, as small amounts of UDP-sugars were found to hydrolyze to UMP during extraction and in buffer before analysis (**Figure A 4**). Because UDP-sugars are highly abundant, this led to a false magnification of the small UMP pool. Because there were no significant differences between *umk1-1* and *umk1-2* as well as *umk2-1* and *umk2-2*, only the results for the *umk1-1* and *umk2-1* line are shown in the main figures for sake of simplicity. The results for all genotypes and measured metabolites, which were not included in a main figure, are shown in **Figure A 5**, **Figure A 6** and **Figure A 7**.

### 3.5.1. Quantification of pyrimidine NTPs and NDPs

UTP, CTP and dCTP are the products of the nucleotide phosphorylation pathway in which the UMKs are directly involved. Their concentrations were measured for from plant material of all *UMK* mutant lines (**Figure 20 A**). For the *umk1-1*, *umk2-1* and *umk1 umk2* lines, no differences in NTP contents can be detected compared to the wild type. This indicates that UMK1 and UMK2 do not contribute significantly to the global pyrimidine NTP pools in 18-day-old plants. The loss of UMK1 does not cause any metabolic alterations, because its loss can likely be compensated by the catalytically more efficient UMK3, which is also located in the cytosol. Also the loss of UMK2 does not lead to metabolic differences in the global pools. As UMK2 is located in mitochondria, it is probable that its loss only causes metabolic alterations of the mitochondrial nucleotide content. It has previously been observed that differences in mitochondrial nucleotide pools may not be detectable in whole cell extracts (Niehaus et al., 2022).

In contrast, the *umk3$_{\Delta E76}$* line and the crosses made with the *umk3$_{\Delta E76}$* line show a lowered UTP, CTP and dCTP content. This indicates a direct negative influence of the mutation of *UMK3* on pyrimidine NTP concentrations. For UTP, there are no differences detectable between the *umk3$_{\Delta E76}$* line and the *umk3$_{\Delta E76}$* crosses. All show a similar reduction in UTP content, indicating that UMP phosphorylation relies mostly on UMK3. This is also reflected in the kinetic parameters of the UMKs, as UMK3 has by far the best catalytic efficiency with UMP compared to the other two enzymes (**Table 3**).

For CTP however, the *umk1 umk3$_{\Delta E76}$* line and the triple mutant have a significantly lower CTP content than the *umk3$_{\Delta E76}$* line. In general, lower UTP amounts should be reflected in lower CTP concentrations, as both

metabolites are directly coupled via the reaction catalyzed by the CTPS. The even lower CTP concentration associated with a mutation of both cytosolic UMKs additionally suggests an impact of cytosolic CMP phosphorylation on the CTP content. This effect becomes even more apparent when looking at the UTP to CTP ratios in the mutant lines (**Figure 20 B**). They are similar in all lines, except for the two lines lacking both cytosolic UMKs, where they are shifted towards UTP. Consequently, an involvement of UMK1 in the phosphorylation of CMP is evident. Although the *umk1* line does not have a lower CTP content than the wild type, the function of UMK1 as a CMP kinase is unmasked in the *umk3$_{\Delta E76}$* background. This shows a certain degree of redundancy between UMK1 and UMK3 in their function as CMP kinase.

A similar effect as for CTP can be observed for dCTP, which can be synthesized via the RNR from CTP via CDP. The *umk1 umk3$_{\Delta E76}$* line and the triple mutant show lower amounts of dCTP compared to the *umk3$_{\Delta E76}$* line. There are two possibilities for the lower dCTP amount associated with a mutation of both cytosolic UMKs. Either the lower dCTP content is a direct result of the lowered CTP concentration or the decrease in cytosolic dCMP phosphorylation causes the observed reduction. This question can be answered when examining the CMP and dCMP concentrations, which are presented in the next chapter.

The effects on the contents of UDP and CDP, which are the intermediates between the NMPs and NTPs, are similar to those observed for UTP and CTP (**Figure 20 C**). Overall, the *UMK3* mutant line and its crosses are the only lines that show significant metabolic changes compared to the wild type, but it appears that additional effects of the null-mutation of other *UMK* genes are apparent in the background of the weak *UMK3$_{\Delta E76}$* allele.
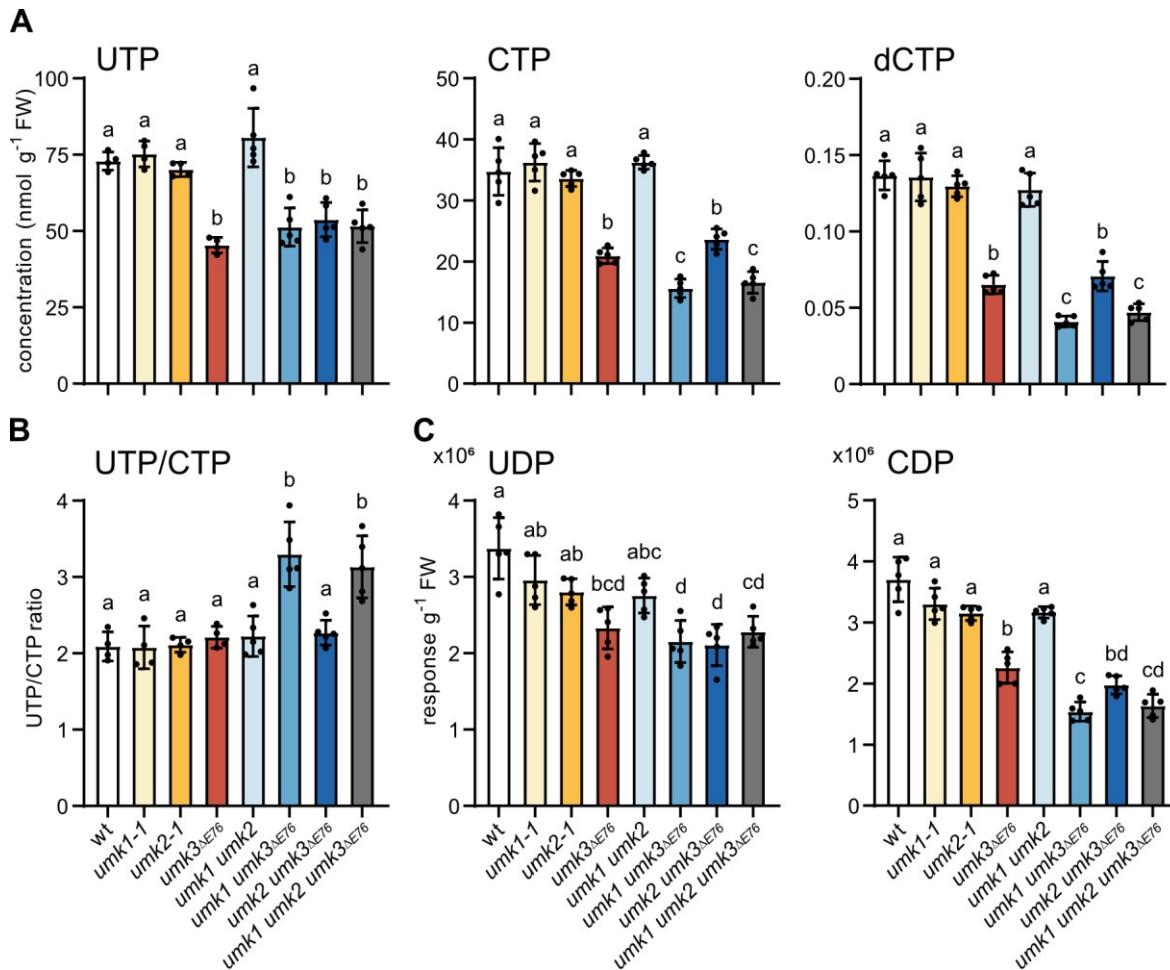
**Figure 20. Pyrimidine NTP and NDP content of 18-day-old *UMK* mutant lines and the wild type.**
Rosettes of 18-day-old plants grown on soil under long-day conditions in a phytochamber were harvested three hours after the onset of light. For each replicate, leaf material from three plants was pooled. Five biological replicates were analyzed per genotype, error bars are SD. Statistical analysis was performed by two-sided Tukey's pairwise comparisons. Different letters indicate differences at $P < 0.05$. To reduce complexity, only data of the *umk1-1* and *umk2-1* lines are shown. Data of *umk1-2* and *umk2-2* lines are very similar and can be accessed in **Figure A 5**, **Figure A 6** and **Figure A 7**. **A)** Pyrimidine NTPs, **B)** ratios of UTP to CTP, **C)** pyrimidine NDPs.

### 3.5.2. Quantification of pyrimidine NMPs

UMP, CMP and dCMP are the direct substrates of the UMKs. As mentioned earlier, it was unfortunately not possible to properly quantify the UMP concentration. As already observed for UTP, CTP and dCTP, there are also no significant metabolic differences for CMP and dCMP in the *umk1-1*, *umk2-1* and *umk1 umk2* lines in comparison to the wild type (**Figure 21**).

Consistent with the observations for the NTPs, which were reduced in the *UMK3* mutant line and its crosses, the NMPs accumulate in these lines. The *umk1 umk3~ΔE76~* line and the triple mutant have a

significantly elevated content of CMP in comparison to the wild type. This effect is not visible for the $umk3_{\Delta E76}$ line, which emphasizes that UMK1 is a CMP kinase *in vivo* and is able to compensate the lowered CMP kinase activity of UMK3$_{\Delta E76}$. The fact that there is no CMP accumulation visible in the $umk3_{\Delta E76}$ line also suggests that the lowered UTP concentration is indeed the sole reason for the lower CTP content in that line. The accumulation of CMP in the *umk1 umk3$_{\Delta E76}$* line and the triple mutant compared to *umk3$_{\Delta E76}$* is consistent with the lower amounts of CTP detected in these lines and confirms that the phosphorylation of CMP in the cytosol contributes towards the CTP content *in vivo*. Both processes proceed via the two cytosolic UMKs. This is plausible as Arabidopsis possesses two cytosolic uridine/cytidine kinases, which catalyze the majority of uridine and cytidine salvage *in vivo* and generate CMP in the cytosol (Ohler et al., 2019). Also CMP from RNA turnover is generated in the cytosol and could contribute to the observed accumulation.

The situation for dCMP is a bit more complex. When compared to the wild type, the accumulation of dCMP in plants of the lines containing the *umk3$_{\Delta E76}$* mutation is generally stronger than the accumulation of CMP. There are two factors which may contribute to a stronger accumulation. Firstly, the three UMKs have generally lower catalytic efficiencies for dCMP and secondly, UMK3$_{\Delta E76}$ has almost no residual activity with dCMP. The dCMP accumulation is stronger in the *umk2 umk3$_{\Delta E76}$* line in comparison to the *umk1 umk3$_{\Delta E76}$* line. This seems a bit contradictory at first as the *umk1 umk3$_{\Delta E76}$* line has less global dCTP than the *umk2 umk3$_{\Delta E76}$* line. Consequently, it could have been expected that *umk1 umk3$_{\Delta E76}$* accumulates more dCMP than *umk2 umk3$_{\Delta E76}$*. A logical explanation for this phenomenon would be that the dCTP content is dependent on the amount of CTP rather than the phosphorylation of dCMP derived from salvage or DNA degradation. Similar to how UTP and CTP are coupled via the CTPS reaction, CTP and dCTP are coupled via the RNR reaction. Consequently, the lower CTP content of the *umk1 umk3$_{\Delta E76}$* line and the triple mutant directly leads to a lower dCTP content, and there is no additional contribution of dCMP phosphorylation to the amount of dCTP, which answers the question from the previous chapter. Another conclusion that can be drawn from the dCMP results is that UMK2 functions primarily as a dCMP kinase *in vivo*. As already mentioned, there is a significant elevation in the global dCMP content visible for *umk2 umk3$_{\Delta E76}$* compared to *umk3$_{\Delta E76}$*. This elevation is likely the result of a dCMP accumulation in mitochondria caused by deoxycytidine salvage, which is catalyzed by the mitochondria localized enzyme dNK (Clausen et al., 2014) or from DNA repair processes. There is no accumulation of dCMP visible for the *umk2-1* line, which suggests that a transport mechanism exists channeling salvaged dCMP to the cytosol, where UMK3 is taking over the phosphorylation. In the *umk3$_{\Delta E76}$* background, dCMP is accumulating in the cytosol, which may negatively influence the channeling of dCMP from the mitochondria, resulting in the additional

accumulation visible in the *umk2 umk3~ΔE76~* line. While an additional accumulation of CMP is visible in *umk1 umk3~ΔE76~* compared to *umk3~ΔE76~*, this is not the case for dCMP, indicating that UMK1 is not involved in dCMP phosphorylation in the cells of 18-day-old Arabidopsis plants. Another interesting observation is the comparably small accumulation of dCMP in the *UMK* triple mutant. As UMK3~ΔE76~ has almost no activity with dCMP, the triple mutant can be considered almost unable to phosphorylate dCMP. However, the accumulation is rather weak, which suggests either the existence of a previously unknown dCMP phosphatase or another way for the cell to metabolize dCMP. A possible way would be via the plastid UMK PUMPKIN, which is active with UMP, but has not been examined with dCMP as substrate (Schmid et al., 2019), the yet undescribed eubacterial UMK encoded at the locus At3g10030 or by activity of another NMK. As NMKs generally possess a high substrate specificity (Yan and Tsai, 1999), it is unlikely that the dCMP is phosphorylated by an NMK that does not belong to the UMK families.
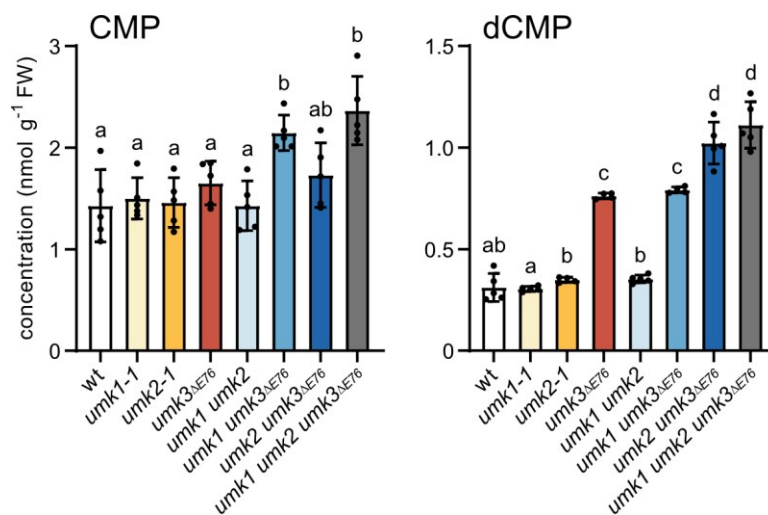


**Figure 21. Pyrimidine NMP content of 18-day-old UMK mutant lines and the wild type.**
Metabolite quantifications from the same samples as in **Figure 20**. Statistical analysis was performed the same way.

The quantification of pyrimidine nucleotide concentrations in the *UMK* mutant lines led to some interesting conclusions about the organization of Arabidopsis pyrimidine metabolism. As previous results already indicated, UMK3 seems to be the main UMK in Arabidopsis responsible for most pyrimidine NMP phosphorylation *in vivo*. All observed metabolic phenotypes during the growth phase are associated with a mutation of *UMK3*. In the *umk3~ΔE76~* background, evidence for the functions of the other UMKs can be found, suggesting that UMK1 acts mainly as a CMP kinase and UMK2 as a dCMP kinase *in vivo*. The results also suggest that UMP phosphorylation seems to be exclusively reliant on UMK3, as there is no additional

reduction of the UTP concentration when the other UMKs are mutated in $umk3_{\Delta E76}$ background. Apart from the functions of the UMKs, there is also evidence that not only CTP synthesis from UTP via the CTPS reaction, but also the cytosolic phosphorylation of CMP contributes to CTP pools *in vivo*. In contrast, the dCTP concentration is closely linked to the CTP concentration and phosphorylation of dCMP does not significantly contribute to the global dCTP amount in 18-days-old plants.

### 3.5.3. Quantification of UDP-sugars

In addition to the nucleotides, also UDP-sugars were quantified. Plants possess a variety of UDP-sugars involved in different biosynthetic reactions. The most abundant one is UDP-glucose, which is essential for cellulose synthesis. UDP-sugars were quantified from solid phase extraction samples as described by (Rautengarten et al., 2019). The results for UDP-glucose, UDP-galactose and UDP-arabinose as representative examples are displayed (**Figure 22**). The quantifications of other UDP-sugars led to similar results and can be found in **Figure A 7**.

Compared to the wild type, the amount of UDP-glucose is lower in plants of the $umk3_{\Delta E76}$ line, but not in *UMK1* or *UMK2* mutants. As already observed for UTP, there seems to be no additional reduction of UDP-glucose content when *UMK1* and/or *UMK2* are mutated in $umk3_{\Delta E76}$ background. The same is true for the other quantified UDP-sugars. Interestingly, the reduction associated with the *UMK3* mutation varies greatly between the different UDP-sugars. While the $umk3_{\Delta E76}$ line has 30 % less UDP-glucose than the wild type, UDP-galactose content is reduced by 40 % and UDP-arabinose content by almost 80 %. There seems to be no obvious pattern to the observed reductions. UDP-arabinose can be synthesized from UDP-glucose via three enzymatic steps, but the intermediates show different reductions. UDP-glucose (30 % reduction) is oxidized by UDP GLUCOSE-6-DEHYDROGENASE to UDP-glucuronate (80 % reduction), which is decarboxylated by UDP GLUCURONATE DECARBOXYLASE to UDP-xylose (30 % reduction). UDP ARABINOSE EPIMERASE then catalyzes the conversion from UDP-xylose to UDP-arabinose (80 % reduction). It is unlikely, that the availability of the respective sugar-1-phosphates causes the reductions, as the UMKs are not involved in sugar metabolism. There appears to be a regulative mechanism in UDP-sugar metabolism based on UTP availability, which has varying effects on the contents of the different UDP-sugars. Since they are the precursors for cell wall synthesis, the reduced amounts of UDP-sugars are probably the reason for the growth phenotypes observed in the $umk3_{\Delta E76}$ line and the crosses derived from the $umk3_{\Delta E76}$ line (**Figure 19**).
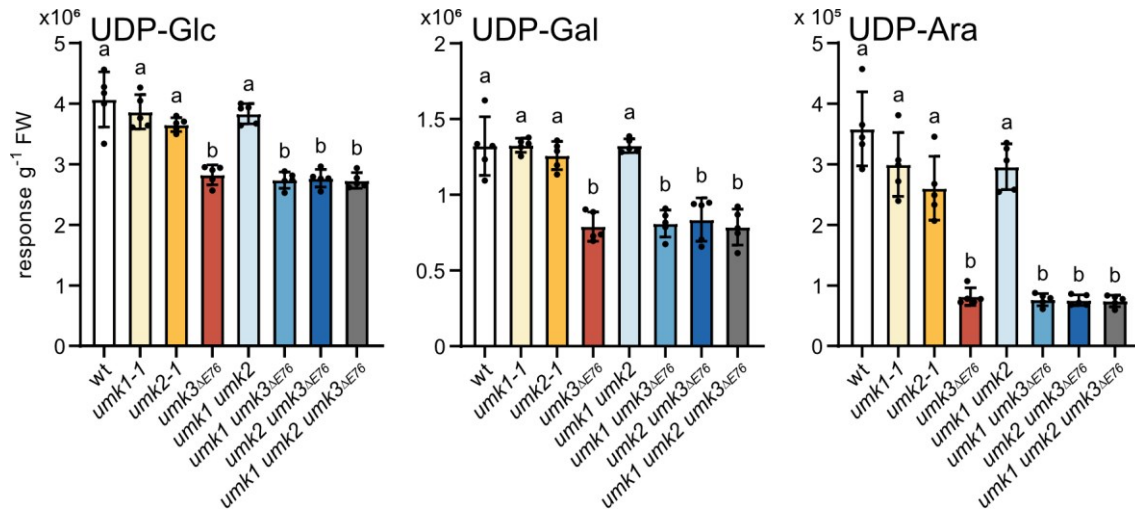
**Figure 22. UDP-sugar content of 18-day-old UMK mutant lines and the wild type.**
Metabolite quantifications from the same samples as in **Figure 20**. Statistical analysis was performed the same way. UDP-Glc, UDP-glucose; UDP-Gal, UDP-galactose; UDP-Ara, UDP-arabinose.

## 3.6. Complementation of the *UMK3* mutation

The *umk3$_{\Delta E76}$* line was transformed with the construct H1178, which expresses the *UMK3* gene fused with a sequence encoding a C-terminal Strep-tag by the native *UMK3* promoter. Nine transgenic plants were identified in the T1 generation by glufosinate selection, of which three were used for further experiments. The plants were grown together with wild type and *umk3$_{\Delta E76}$* plants and T2 seeds harvested. The T2 plants were grown in parallel and the *umk3$_{\Delta E76}$::UMK3* plants were again sprayed with glufosinate to eliminate the non-transgenic plants from the population. Five plants per line were grown for 35 days under long-day conditions and photographed as before (**Figure 23 A**) and the leaf area of these plants was determined (**Figure 23 B**). Plants of the *umk3$_{\Delta E76}$* line showed a reduced leaf area of about 30 % compared to the wild type, which is slightly more than observed before. The three complementation lines also showed reduced leaf areas between 10 and 20 % compared to the wild type, but they grew better overall compared to *umk3$_{\Delta E76}$*. It is possible that the early treatment with glufosinate, which was only done for the complementation lines, stressed the plants resulting in a slightly delayed growth. Samples were taken from the complementation lines and a wild type plant for StrepTactin affinity purification and immunoblot with an anti-Strep antibody. The immunoblot showed that the transgene was expressed in these lines (**Figure 23 C**). For metabolite analysis, four samples were taken of each line by pooling three plants per sample from 18-day-old plants. The nucleotide content of the three complementation lines, the wild type and the *umk3$_{\Delta E76}$* line were analyzed by extracting nucleotides via SPE and quantification by LC-MS analysis. The results show that it was possible to complement the metabolomic phenotypes of the *umk3$_{\Delta E76}$* line in the

three chosen *umk3_{ΔE76} ::UMK3* lines (**Figure 23 D**). The lower UTP-, CTP- and dCTP-concentrations of the *umk3_{ΔE76}* line could not be observed in either of the three *umk3_{ΔE76}::UMK3* lines. The same is true for the accumulation of dCMP, which was only observed in the *umk3_{ΔE76}* line. Consequently, the metabolic phenotype in the *umk3_{ΔE76}* line can be complemented by expression of an intact *UMK3* gene. It is evident, that the growth reduction and metabolic alterations observed in *umk3_{ΔE76}* are caused by the mutation of *UMK3*.
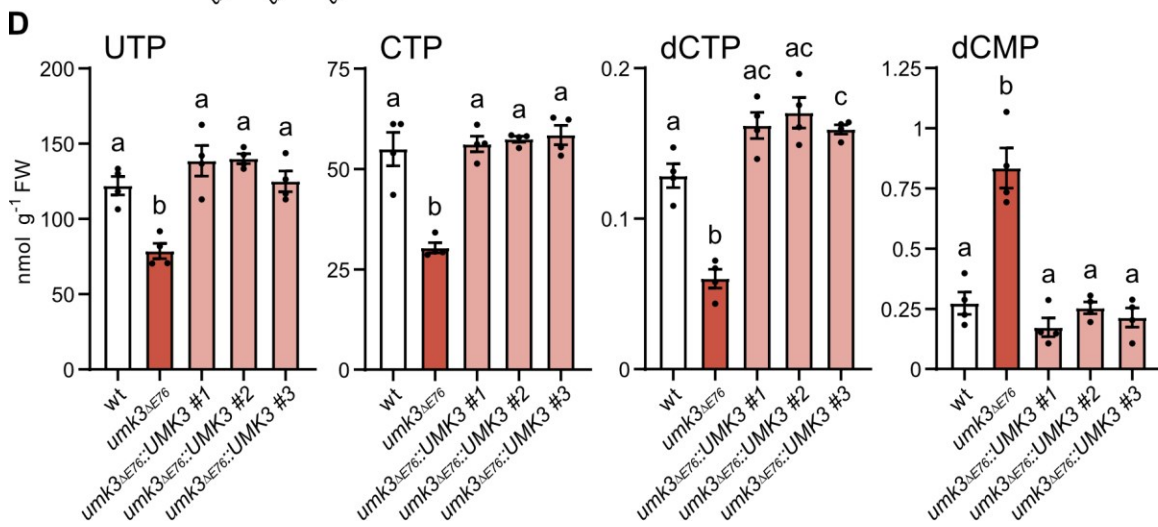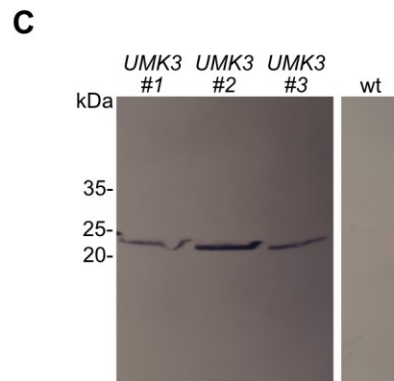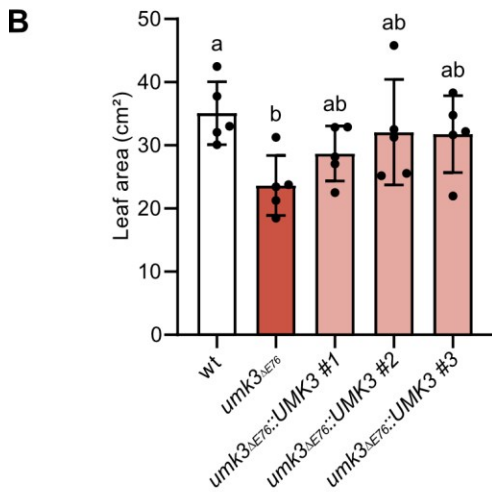
**Figure 23. Quantification of leaf area and pyrimidine nucleotides in *umk3$_{\Delta E76}$* and complementation lines compared to wild type.**
C-terminal HA-Strep-tagged UMK3 was expressed from a transgene driven by the native *UMK3*-promoter (construct H1178) in *umk3$_{\Delta E76}$* background. Three lines were chosen for analysis. **A)** Phenotypes of 35-day-old plants of the wild type (wt), *umk3$_{\Delta E76}$* and three complementation lines. Scale bar, 1 cm **B)** Quantification of rosette leaf areas of five plants of each genotype. Statistical analysis was performed using two-sided Tukey's pairwise comparisons. Different letters indicate differences at P < 0.05. Error bars are SD. **C)** Detection of the tagged UMK3 in leaves of the transgenic lines after affinity purification on an immunoblot using an anti-Strep antibody. **D)** Metabolite analysis using 18-day-old plants of the wild type, *umk3$_{\Delta E76}$* and the *umk3$_{\Delta E76}$* complementation lines. Statistical analysis was performed using the two-sided Tukey's pairwise comparison. Different letters indicate significant differences at P < 0.05. Error bars are SD.

## 3.7.  Expression profile of the *UMK* genes

In 18-day-old plants, metabolic alterations are primarily associated with the mutation of *UMK3*. When analyzing expression data of the three *UMKs*, it is apparent that *UMK3* is strongly expressed in leaves during the vegetative growth phase of the plants, while *UMK1* and *UMK2* are expressed weaker (**Figure 24 A**). The same can be observed in proteomic data, where UMK3 is about one order of magnitude more abundant than UMK1 or UMK2 (Mergner et al., 2020). Consequently, the analysis of 18-day-old plants may not be the ideal developmental stage to identify roles of *UMK1* or *UMK2*. Analysis of older leaves may be suited to detect metabolomic alterations in the *UMK1* mutant line, as *UMK1* is stronger expressed in mature leaves (harvested upon anthesis of the first flower) than *UMK3*. Additionally, the expression of *UMK3* is lower during early stages of germination. After seed imbibition at 4°C in the dark *UMK2* shows the strongest expression (**Figure 24 B**). The stronger expression persists until shortly after the seeds are transferred to the light, which appears to be a trigger for the expression of *UMK3*. Analysis of the mutant lines during germination could therefore provide information about the role of UMK2.
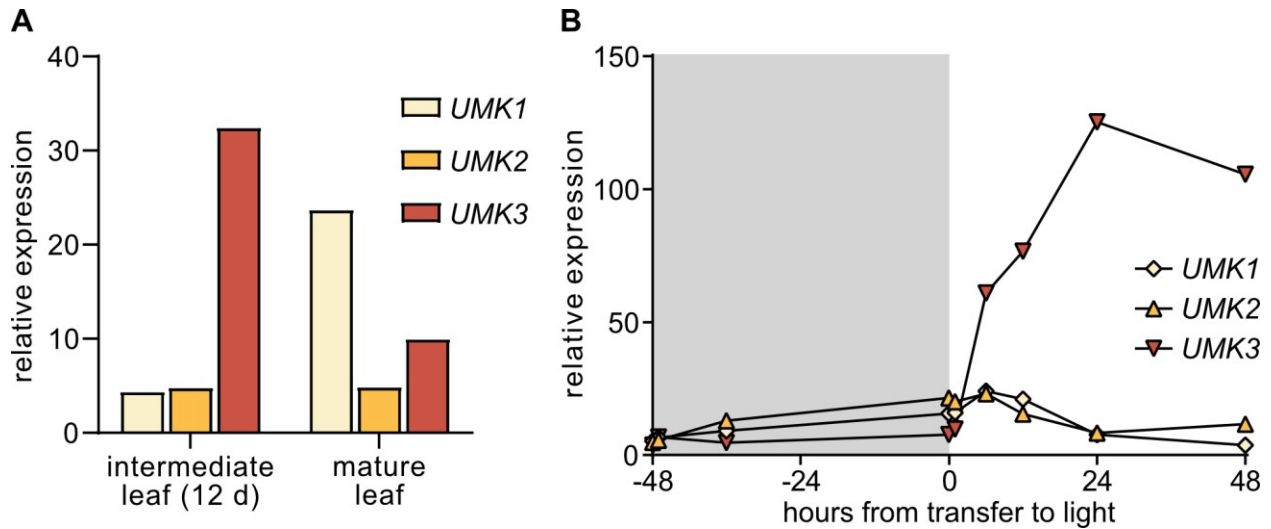
**Figure 24. Relative expression profile of the three UMKs.**
**A)** Relative expression of *UMK1*, *UMK2* and *UMK3* in leaves of developing (12 days) and mature (anthesis of the first flower) Arabidopsis plants (adapted from Klepikova et al., 2016). **B)** Relative expression during seed imbibition and germination. Imbibition of the seeds was started at the -48 hours time point. The shaded area refers to an incubation of the seeds at 4°C in the dark. At 0 hours the seeds were moved to long-day growth conditions (adapted from Narsai et al., 2011).

## 3.8. Quantification of pyrimidine NTPs in senescent *UMK* mutant plants

To assess whether mutation of *UMK1* causes metabolomic phenotypes in older leaves, the leaves of 2-month-old plants were harvested and the nucleotide content analyzed. Wild type, *umk1-1*, *umk3$_{\Delta E76}$* and *umk1 umk3$_{\Delta E76}$* plants were included in this experiment and the results for UTP, CTP and dCTP are displayed in **Figure 25**. In general, the concentrations of the three compounds are reduced in comparison to the 18-day-old plants. In senescent leaves the metabolism is switching towards a remobilization of compounds like nitrogen, lipids and nucleotides to make them available for younger or reproductive tissues (Feller and Fischer, 1994; Thompson et al., 1998; Diaz et al., 2008; Sakamoto and Takami, 2014). As cellular functions decline, transcription and DNA replication also decrease. Thus, less nucleotides are necessary to fuel these processes. This is particularly evident in the dCTP content, which is reduced more than five-fold in the senescent wild type leaves compared to the 18-day-old plants. When comparing the *umk1-1* line with the wild type, it is apparent that also in older leaves the *UMK1* single mutant does not show a metabolomic phenotype. Interestingly, the *umk3$_{\Delta E76}$* line also shows no differences in CTP and dCTP concentrations compared to the wild type, indicating that UMK1 can fully complement the lower activity of UMK3$_{\Delta E76}$ in older tissue for the production of these compounds, which was not observed in the 18-day old plants. Thus, UMK1 appears to have an elevated role during senescence. However, the mutation of *UMK3* still

seems to be more detrimental to the plant in this context because the UTP concentration in $umk3_{\Delta E76}$ is reduced compared to the wild type as it had already been observed in 18-day-old plants. In contrast to the situation in younger leaves, it appears that in the older leaves the CTP concentration is not that closely coupled to the UTP concentration anymore. It is possible that here salvaging and recycling of nucleotides has a greater influence on the nucleotide concentrations, specifically CTP and dCTP. In this context it is plausible that *UMK3* expression is reduced, as probably less UMP is produced from *de novo* synthesis.
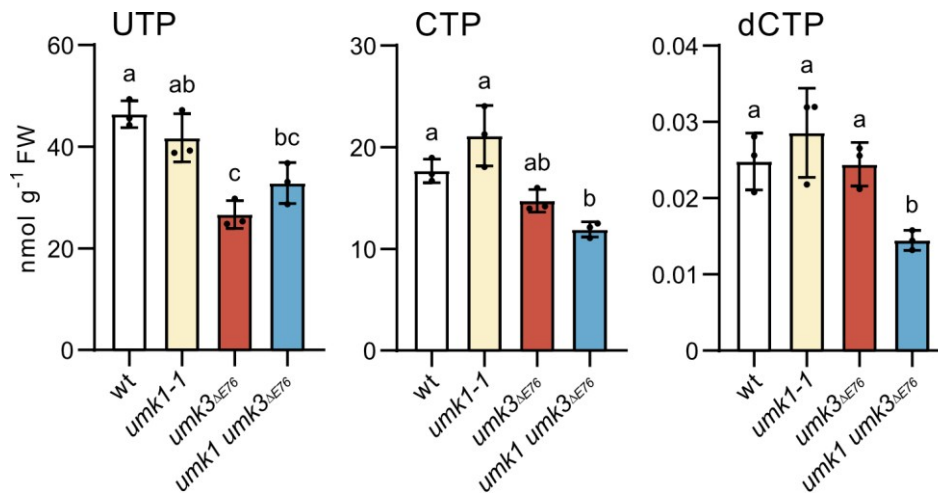


**Figure 25. Pyrimidine NTP content of 60-day-old *UMK* mutant lines and the wild type.**
Multiple leaves of 60-day-old plants grown on soil under long-day conditions in a phytochamber were harvested three hours after the onset of light. Three biological replicates were analyzed per genotype, error bars are SD. Statistical analysis was performed by two-sided Tukey's pairwise comparisons. Different letters indicate differences at P < 0.05.

## 3.9. Metabolome analysis of *UMK* mutant lines during germination

For analysis of *UMK* mutant lines during germination, 10 mg of wild type, *umk2-1*, *umk3_{\Delta E76}* and *umk2 umk3_{\Delta E76}* seeds were placed on filter paper soaked with half-strength MS-medium in Petri dishes. The Petri dishes were incubated for 48 hours in the dark at 4°C and then transferred to long-day growth conditions to replicate the conditions that were used to obtain the expression data (**Figure 24 B**, Narsai et al., 2011). Five replicates of each line were collected from dry seeds, seeds imbibed for 24 hours in the dark and imbibed seeds moved to the light for 1, 24 or 48 hours. Nucleotides were extracted from all samples by solid phase extraction and quantified via LC-MS analysis. The aim of this experiment was to find a role for the mitochondria-localized UMK2 during germination. During the early stages of germination, *UMK2* is stronger expressed than *UMK1* and *UMK3* (**Figure 24 B**). Additionally, mitochondrial activity is initiated

directly upon imbibition of the seeds accompanied by the synthesis of NTPs for DNA replication and transcription (Paszkiewicz et al., 2017). UMK2 may be involved in early mitochondrial development and mutation of the gene may result in altered nucleotide concentrations in germinating seeds. The development over time of the contents of the UMK2 substrates dCMP and CMP, as well as the end products of the phosphorylation pathway, in which UMK2 is involved, UTP, CTP and dCTP, are displayed in **Figure 26**.

A significant metabolomic effect, which is not dependent on the mutation of *UMK3*, is observed in the *UMK2* mutant line compared to the wild type. At all analyzed time points, the *umk2-1* line contains significantly more dCMP than the wild type (indicated by *). 24 hours before and one hour after moving the seeds to light the dCMP content is also significantly higher than in the *umk3$_{ΔE76}$* line (indicated by **). However, the global dCTP concentration is unaltered in the *umk2-1* line compared to the wild type at all time points. This indicates that also during germination, the phosphorylation of mitochondrial dCMP does not significantly contribute to the global dCTP content. Nevertheless, UMK2 appears to be involved in recycling this dCMP during the early stages of germination. After *UMK3* expression increases with the transfer of the seeds to light, the bulk of dCMP phosphorylation is probably catalyzed by UMK3.

The germinating seeds of the *umk2-1* line also have a significantly lower CTP content compared to the wild type at 24 hours before and one hour after transfer to the light. It was established earlier, that a lower CTP content is likely a result of a lower UTP content as both are directly coupled via the enzyme CTPS. But here, the *umk2-1* line does not show a lower UTP content in comparison to the wild type. This means that there must be another CTP source that is not accessible due to the mutation of *UMK2*. It has been shown that the total RNA content of seeds decreases rapidly upon imbibition (Li et al., 2006). This rapid degradation of parental RNA releases rNMPs which can be recycled. This indicates that UMK2 is not only involved in dCMP, but also CMP recycling early on. UMK2 has a comparatively good catalytic efficiency with CMP (**Table 2**). As a similar CTP reduction is visible in the *umk3$_{ΔE76}$* line, UMK2 and UMK3 likely play an equally important role here. The difference in CTP content in *umk2-1* disappears however, as soon as the *UMK3* expression increases.

The effect that is observed for the CTP content 24 hours before and one hour after transfer to light in the *umk2-1* or *umk3$_{ΔE76}$* line appears to have no influence on the dCTP content yet. This could be due to the fact that nuclear and organellar DNA replication has not yet started at these time points and therefore dCTP is not yet consumed (Niehaus et al., 2022). However, 24 hours after transfer to light, DNA replication is active resulting in a consumption of dCTP. From this time point on, the influence of the CTP content on

the dCTP concentration becomes evident, as in *umk3*$_{\Delta E76}$ both concentrations are lower compared to the wild type.

The pyrimidine nucleotide contents of the *umk3*$_{\Delta E76}$ and *umk2 umk3*$_{\Delta E76}$ lines show similar alterations at the 24 hour and 48 hour time points as the 18-day-old plants. Also during imbibition there are already small reductions in UTP and CTP content visible compared to the wild type. After transfer to the light, when *UMK3* expression increases, the metabolic difference between the *umk3*$_{\Delta E76}$ line and the wild type also increases. There are no metabolic differences between the *umk3*$_{\Delta E76}$ line compared to the *umk2 umk3*$_{\Delta E76}$ line, except for their dCMP content. This observation is consistent with the results obtained from the 18-day-old plants and again emphasizes that, also during germination, dCMP phosphorylation does not contribute to the production of dCTP.
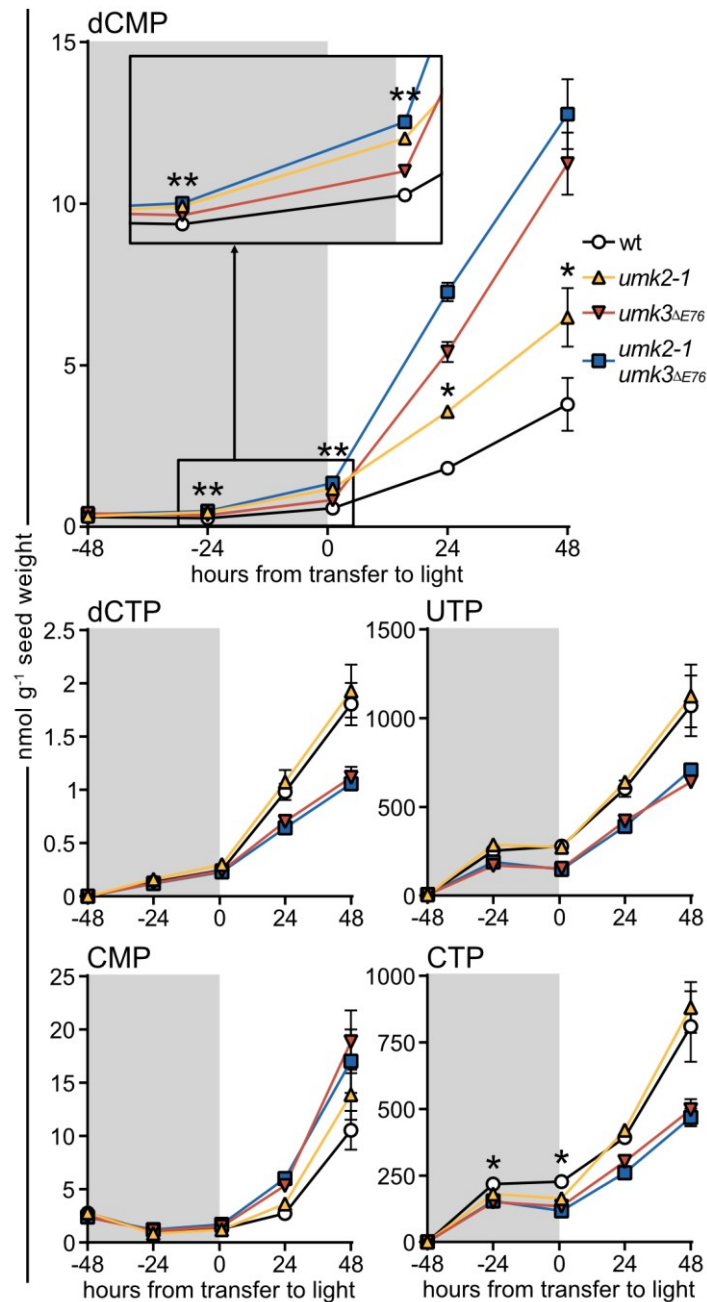
**Figure 26. Pyrimidine nucleotide content of *UMK* mutant seeds compared to the wild type during germination.**

Metabolite content of *umk2* (allele *umk2-1*), *umk3_{ΔE76}* and *umk2 umk3_{ΔE76}* seeds before and during germination in comparison to the wild type. -48 h, dry seeds; -24 h, seeds imbibed for 24 hours; all other time points, number of hours after moving seeds to light. 10 mg seeds were used per sample and five biological replicates from the same seed batch were used per time point, error bars are SD. Statistical analysis was performed by two-sided Tukey's pairwise comparisons. One asterisk indicates a difference of *umk2* to the wild type at P < 0.05. Two asterisks indicate higher dCMP content in *umk2* compared to the wild type and also *umk3_{ΔE76}* at P < 0.05.

To gain further insights into the organization of Arabidopsis pyrimidine metabolism during germination, the experiment was repeated with seeds from all mutant lines 24 hours after transfer to light. While the results for the pyrimidine nucleotide triphosphates look similar to what has been observed for the 18-day-old plants, the results for CMP and dCMP revealed interesting variations (**Figure 27**).

The results for CMP deviate in germinating seeds compared to the 18-day-old plants. In the 18-day-old plants only the *umk1 umk3*$_{\Delta E76}$ line and the triple mutant showed a slight accumulation. In the germinating seeds a similar accumulation is visible in *umk3*$_{\Delta E76}$, *umk1 umk2* and *umk2 umk3*$_{\Delta E76}$, as well as a more pronounced accumulation in *umk1 umk3*$_{\Delta E76}$ and the triple mutant. The higher amount of CMP during germination is likely originating from RNA degradation (Li et al., 2006). The results suggest that this initial wave of CMP is mainly phosphorylated by the cytosolic UMKs at this time point. The strong accumulation in the *umk1 umk3*$_{\Delta E76}$ line compared to *umk3*$_{\Delta E76}$ indicates a significant involvement of UMK1 in this early CMP phosphorylation. It is possible that the role of UMK1 is even more prominent during imbibition, where *UMK3* is not yet strongly expressed. The small difference between *umk1 umk3*$_{\Delta E76}$ and the triple mutant indicates that after 24 hours in the light UMK2 is not as strongly involved in CMP phosphorylation anymore as it had been during the imbibition phase.

The accumulation of dCMP is stronger in germinating seeds than in the 18-day-old plants. This can especially be observed in the triple mutant, which accumulates 9.2-fold more dCMP than the wild type, while the difference is only 3.6-fold in 18-day-old plants. This indicates that either there is more dCMP produced during early germination than in older plants, possibly from DNA repair, or that other pathways that would metabolize dCMP are not yet active. The second strongest accumulation of dCMP is visible for *umk1 umk3*$_{\Delta E76}$, which is different to the situation in 18-day-old plants, where the *umk2 umk3*$_{\Delta E76}$ line accumulated more dCMP (**Figure 21**). Most dCMP appears to be released in the cytosol, hence the stronger accumulation in the *umk1 umk3*$_{\Delta E76}$ line. However, the accumulation is even stronger in the triple mutant indicating that dCMP is able to relocate to mitochondria where it is phosphorylated by UMK2. It seems that the DNA repair during germination releases significant amounts of dCMP, requiring all three UMKs for effective recycling. Another surprising result from this experiment is the absolute concentration of dCMP in comparison to dCTP. The triple mutant contains about 40-fold more dCMP than dCTP (24.6 nmol g$^{-1}$ seed weight dCMP compared to 0.6 nmol g$^{-1}$ seed weight dCTP). One could imagine that phosphorylation of this dCMP would increase the cellular dCTP content. However, the *umk1 umk3*$_{\Delta E76}$ line shows the same dCTP content as the triple mutant, while having 40 % less dCMP. An amount of dCMP equivalent to this 40 % difference, was presumably phosphorylated to dCTP by UMK2 in the *umk1 umk3*$_{\Delta E76}$ line without this having any influence on the dCTP content. This raises the question of why the excess of

dCTP generated in the *umk1 umk3$_{\Delta E76}$* line compared to the triple mutant cannot be detected? One possibility is, that it is rapidly consumed for DNA replication. To investigate this theory, the number of nuclear and mitochondrial genomic DNA copies was quantified.
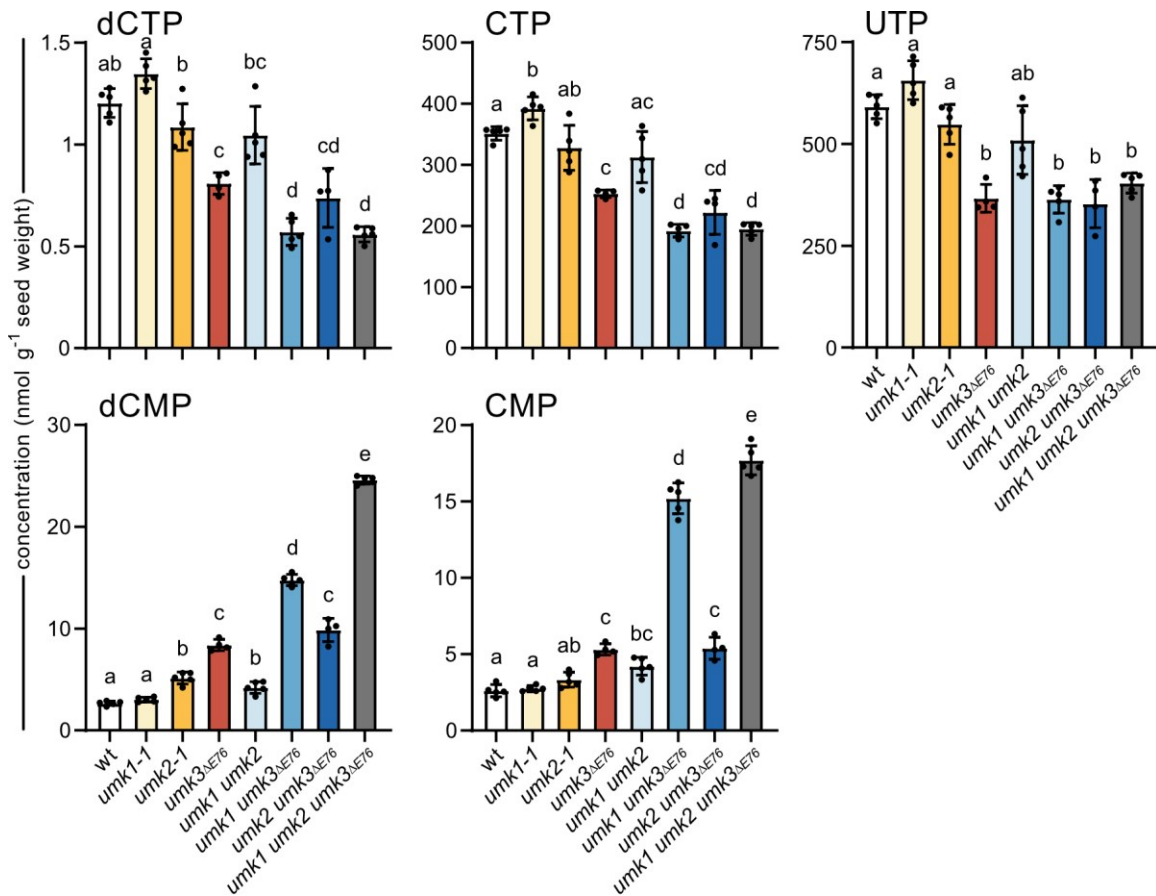


**Figure 27. Pyrimidine nucleotide content of *UMK* mutant seeds compared to the wild type after imbibition and 24 hours in the light.**
Content of pyrimidine nucleotides 24 hours after the seeds were moved to long-day growth conditions. Five biological replicates were analyzed per genotype, error bars are SD. Statistical analysis was performed by two-sided Tukey's pairwise comparisons. Different letters indicate differences at P < 0.05.

## 3.10. Quantification of DNA copy number during germination

To assess effects of the altered pyrimidine nucleotide concentrations in germinating seeds, genomic DNA of the *UMK* mutant lines was quantified. The abundance of genomic nuclear (ncDNA) and mitochondrial DNA (mtDNA) was determined by quantitative PCR (qPCR) using primers amplifying a specific fragment

from the nuclear *UBC21* gene or the mitochondrial *COX1* gene. The ratios of mtDNA to ncDNA were calculated by comparing the number of PCR cycles at which a fixed Ct value was surpassed. This experiment was performed together with Markus Niehaus from the Institute of Plant Nutrition, Leibniz University Hanover (LUH). Samples were taken from dry seeds or seeds that were germinated as explained before and incubated under long-day conditions for 24 or 48 hours. A lower mtDNA copy number does not automatically mean that there is a lower number of mitochondria present as not all mitochondria possess a genome (Preuten et al., 2010) and they are in a constant exchange via fusion and fission (Arimura et al., 2004). However, it may reduce the overall activity of mitochondria, as mitochondria without a genome are unable to synthesize mRNA and may have difficulty maintaining an intact proteome.

The results of the ratio mtDNA/ncDNA are displayed in **Figure 28**. In dry seeds, with the exception of the triple mutant line, no significant differences are detectable compared to the wild type. The dry seeds of the triple mutant already have significantly fewer mtDNA copies, indicating that the mother plant was unable to maintain the wild type level of mtDNA. After imbibition and 24 hours under long-day growth conditions, replication of mtDNA and ncDNA is beginning. At this time point, a reduction of mtDNA copy number can be detected for all lines, except *umk1-1*. The strongest reductions are visible in the triple mutant and the *umk2 umk3$_{ΔE76}$* line. The fact that the *UMK2* single mutant has less mtDNA copies indicates an impairment of mtDNA replication, although the global dCTP concentration in *umk2-1* is on the same level as in the wild type (**Figure 27**). This suggests that the mitochondria of the *umk2* line likely have a local dCTP deficiency, caused by the missing dCMP phosphorylation catalyzed by UMK2. Thus, UMK2 activity contributes to normal mitochondrial DNA replication during germination. The same is true for a mutation of *UMK3*, which also reduces mtDNA copy number. Consequently, the mtDNA replication is reliant on dCTP supply generated in the mitochondria and the cytosol. The *umk2 umk3$_{ΔE76}$* line has an even lower amount of mtDNA copies, which strengthens this theory. A significant contribution of UMK1 activity towards the mtDNA replication appears unlikely, as there is no effect on mtDNA copy number observable in *umk1-1* and no additional reduction of mtDNA copy number in *umk1 umk3$_{ΔE76}$* compared to *umk3$_{ΔE76}$*. The quantification of mtDNA can also answer the question from the previous chapter on why there is no difference in the global dCTP content in the *umk1 umk3$_{ΔE76}$* line compared to the triple mutant 24 hours after the seeds were moved to light (**Figure 27**). Since the *umk1 umk3$_{ΔE76}$* line has more mtDNA copies than the triple mutant, it is likely that the dCTP generated by dCMP phosphorylation in *umk1 umk3$_{ΔE76}$* is consumed for mitochondrial DNA replication and therefore no difference in dCTP content is detectable. After 48 hours, the reduction in mtDNA copy number can still be observed, indicating that mtDNA replication is still impaired. Interestingly, this impairment does not seem to affect seedling development

under the employed growth conditions, as the seven-day-old mutant seedlings grew like the wild type plants (**Figure 18**).

The measurements of mtDNA copy numbers suggest that there are intracellular differences of nucleotide concentrations, which are not detectable in whole cell extracts. Although the global dCTP content in *umk2* is unaltered compared to the wild type, the mtDNA replication is impaired in this line. It also becomes apparent, that UMK3 contributes to a normal mtDNA synthesis, which implies that mitochondria rely on the import of dCDP or dCTP early on, although they possess their own dCMP phosphorylation capacity.
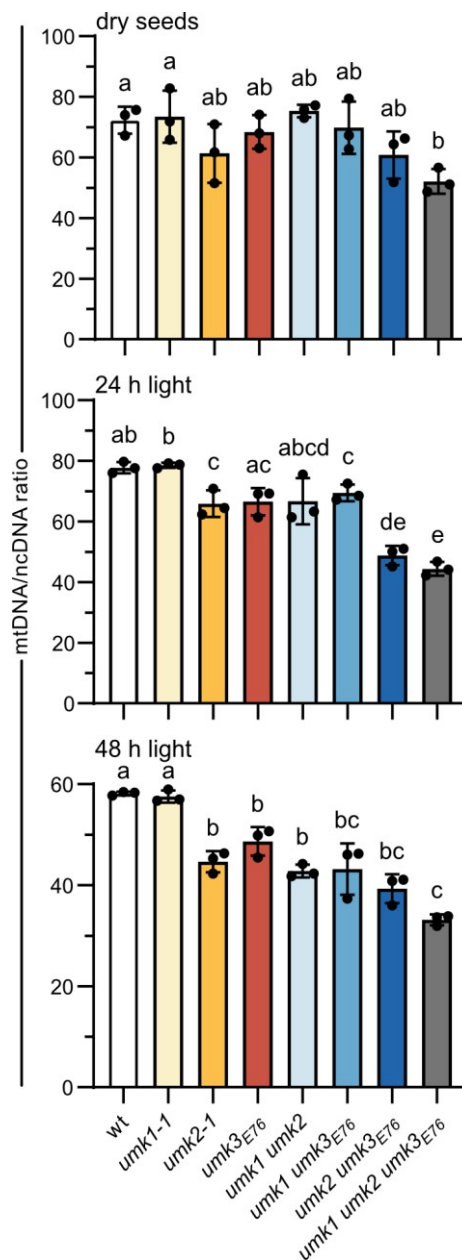
**Figure 28. Ratio between mitochondrial and nuclear genome copy number in *UMK* mutant seeds during germination.**
Seeds were imbibed at 4°C for 48 hours and then moved to long-day growth conditions. Nuclear and mitochondrial DNA was quantified by qPCR with primers amplifying part of the nuclear *UBC21* gene and the mitochondrial *COX1* gene. The same threshold Ct-value was chosen for all samples to ensure comparability. Three biological replicates were analyzed per genotype and each data point of a biological replicate represents the mean of three technical replicates. Statistical analysis was performed using two-sided Tukey's pairwise comparisons. Different letters indicate differences at P < 0.05. Error bars are SD.

## 3.11. Proteome analysis of *UMK2* mutants

A lower dCTP concentration in the mitochondria is probably the cause of the observed impairment of mtDNA replication. It is similarly possible that the *umk2* line also has a lower UTP and/or CTP concentration in the mitochondria causing transcriptional and subsequently translational impairments. To investigate this, a shotgun proteomic experiment was performed by Nils Rugen from the Plant Proteomics Department, LUH. In this experiment the proteome of dry and germinating (24 hour time point) wild type and *umk2-1* seeds was compared for alterations with a focus on mitochondria encoded proteins. Three replicates of 20 mg seeds were used per treatment and genotype. Proteins were extracted from the samples (Klusch et al., 2023) and digested with trypsin. After purification, the samples were analyzed with a timsTOF Pro mass spectrometer and evaluated in MaxQuant 2.2.0.0. The proteomic data was further analyzed in Perseus and a principal component analysis performed as well as a volcano plot generated to visualize the results (**Figure 29**). The principal component analysis showed that the samples of the same genotype and treatment grouped together (**Figure 29 A**). However, there were no significant differences in protein abundance, when evaluating the volcano plot for the comparison of germinating wild type and *umk2-1* seeds (**Figure 29 B**). In the plot, a significant change in protein abundance would be visible by data points being located outside the black lines, but this is not the case. Also when focusing on mitochondria encoded proteins, which are highlighted in orange, no trend towards a lower abundance of mitochondria encoded proteins can be observed. A power analysis by Frank Schaarschmidt of the Institute of Cell Biology and Biophysics, LUH showed, that a 20 % reduction in the abundance of a particular protein (20 % reflecting the reduction in mtDNA copies), would not be detectable with this experimental setup.
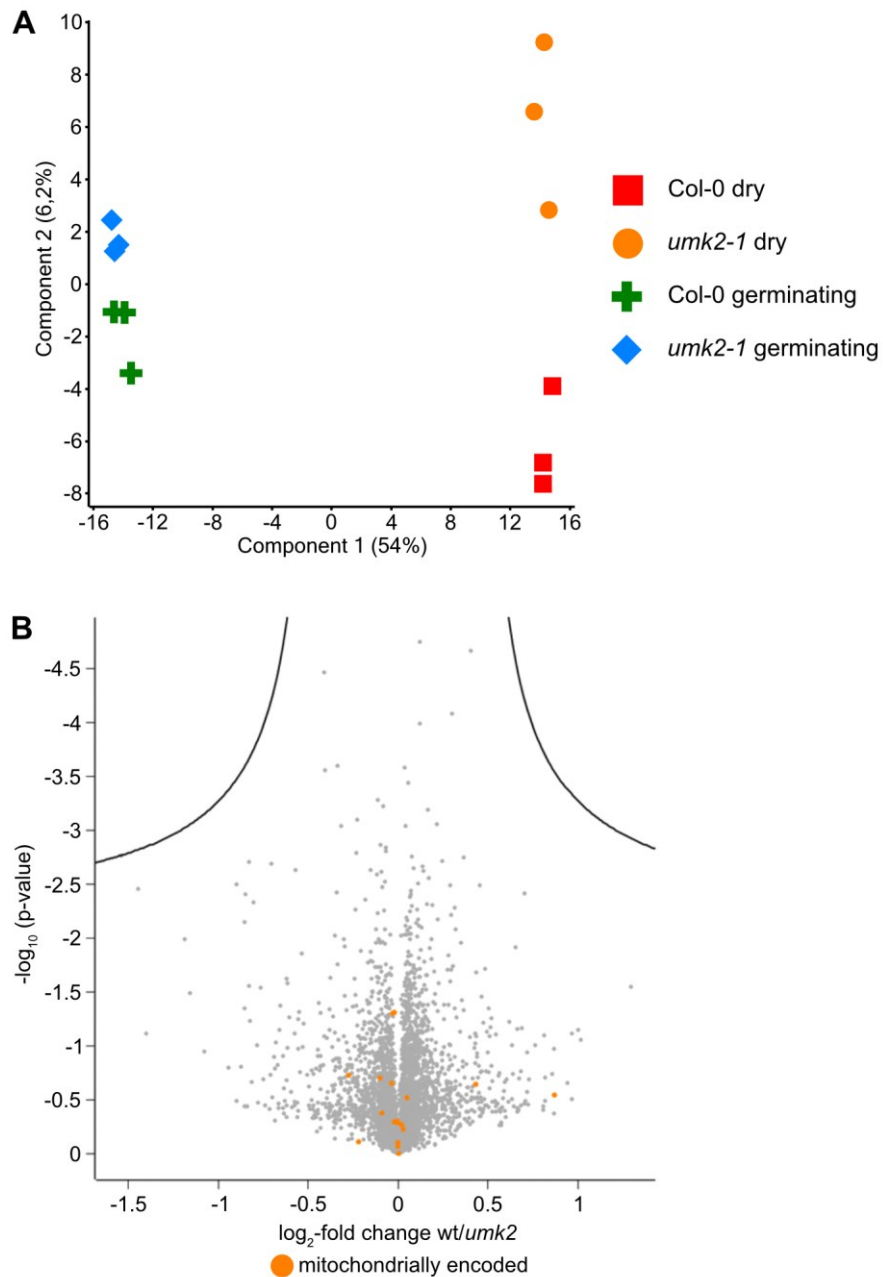
**Figure 29. Proteome analysis of dry and germinating (24 h) wild type and *umk2-1* seeds.**
Three biological replicates of dry seeds from both genotypes were analyzed as well as seeds that had been imbibed for 48 hours in the dark at 4°C followed by 24 hours under long-day conditions in the light. **A)** Principal component analysis for all 12 proteomics experiments based on filtered, log2-transformed and Z-score normalized iBAQ values. **B)** Volcano plot showing differentially expressed proteins between germinating wild-type and *umk2* seeds. For each protein, significance expressed as log-transformed *p*-value was graphed in function of difference between samples (log2 fold change). The black curve indicates truly significant changes in protein expression as determined by the "Significance Analysis of Microarrays" method (Tusher et al., 2001; Tyanova et al., 2016b). Only proteins above the black curve are considered to show a significant change in protein expression between wild-type and *umk2-1* seeds. Proteins encoded in the mitochondrial genome are labeled in orange.

# 4. Discussion

The aim of this thesis was to characterize the AMK-like UMK family of *Arabidopsis thaliana*. Based on information from the literature (Lange et al., 2008) and a phylogenetic analysis by Prof. Claus-Peter Witte, three candidate genes were identified: *UMK1* (At3g60180), *UMK2* (At4g25280) and *UMK3* (At5g26667). During this thesis, it was confirmed that all three genes encode functional UMKs with different subcellular localizations, which are able to utilize UMP, CMP and dCMP as substrates with varying catalytic efficiencies. The quantification of nucleotides in *UMK* mutant plants allowed insights into the organization of the pyrimidine nucleotide metabolism of Arabidopsis and helped to determine the specific roles of the three UMKs. The following discussion will start with examining the roles of the three UMKs in pyrimidine nucleotide metabolism in order of apparent importance, starting with UMK3, followed by UMK2 and lastly UMK1. In the following sections, the results will be discussed in the context of Arabidopsis pyrimidine nucleotide metabolism. Finally, an outlook identifies open questions and possible experiments for the future.

## 4.1. Characterization of UMK3

### 4.1.1. UMK3 is essential for seed establishment and the central UMK in pyrimidine metabolism

To gain insights into the pyrimidine metabolism of *Arabidopsis thaliana UMK* mutant lines were generated using the CRISPR/Cas9 system. While homozygous null-mutant lines of *UMK1* and *UMK2* could be obtained, the generation of a *UMK3* knockout line turned out to be impossible. Heterozygous plants carrying one frameshift allele could be generated, but never a homozygous plant carrying two frameshift alleles. It was also observed that the frameshift alleles could not be transferred to the subsequent plant generation. However, it was possible to generate a 'weak' *UMK3* allele with a three-base-pair deletion, that did not result in a frameshift. This allele, which encoded a UMK3 protein missing the glutamic acid on position 76, could be transferred to the subsequent plant generation and a homozygous line could be obtained. For essential plant genes it can often be challenging or impossible to analyze homozygous null-mutants, as plants with a null-mutation are not able to survive or require specific growth conditions. For example, Arabidopsis plants with a homozygous mutation in the gene encoding URATE OXIDASE (UOX) require sucrose during germination for seedling establishment (Hauck et al., 2014). In the case of *UOX*, plants carrying a heterozygous knockout allele segregated normally to allow the rescue of plants carrying a homozygous null-mutation. This was not possible for *UMK3*, as the frameshift alleles did not segregate

as expected. To explain this phenomenon, it has to be understood how the frameshift mutations were generated in the first place. The *Cas9* gene encoded on the vectors used for transformation is expressed by an egg cell-specific promoter. Consequently, mutations should only be induced in the egg cell during the reproductive phase. However, due to protein and mRNA stability, the Cas9 protein or mRNA can persist longer in the cell and still be active after fertilization (Wang et al., 2015). If the CRISPR/Cas9 system induces a frameshift mutation on one allele after the fertilized cell is in a diploid stage, the second allele, which codes for an intact UMK3, can compensate for this. When a plant with a heterozygous frameshift mutation in *UMK3* enters the reproductive phase, neither the haploid egg cell nor the pollen are viable if they carry the frameshift allele. The non-viable egg cells do not develop into seeds and are aborted early (**Figure 14 A B**). If pollen carrying a frameshift mutation was able to fertilize egg cells, at least some heterozygous plants carrying that frameshift mutation should be observed in the progeny, which was not the case. The importance of UMK3 during reproduction is supported by the fact that the UMK3 protein is highly abundant in pollen, almost two orders of magnitude more than UMK1 or UMK2 (Mergner et al., 2020). It can be stated, that an intact UMK3 is essential during the reproduction phase of Arabidopsis.

Some plants obtained from transformation with H982, the CRISPR/Cas9 construct targeting *UMK3* with the highly efficient sgRNA that generated the *UMK3$_{\Delta E76}$* allele, developed two morphologically distinct types of seeds: darker and wrinkled seeds as well as regularly colored and shaped seeds (**Figure 13**). The wrinkled seeds still glowed faintly green, indicating metabolic activity and transgenicity, but did not germinate anymore. This phenotype is likely caused by an embryo-lethal homozygous frameshift mutation in *UMK3*, which was induced after fertilization. The line H982 #08, which was biallelic for the three-base-pair deletion and a one-base-pair insertion, only developed normal looking seeds, while the line H982 #12, which was heterozygous with a one-base-pair insertion and a wild type allele, showed some wrinkled seeds (**Figure 14 D E**). Both lines were still transgenic and the only difference between them was that H982 #08 had a three-base-pair deletion, whereas H982 #12 still had an intact wild type allele. When grown into the next generation, the allele with the one-base-pair insertion is lost already in the male and female gametophytes for both lines. The fertilized egg cell of H982 #12 can consequently only have two wild type alleles, which are a target for the induction of novel mutations by the CRISPR/Cas9 system. If homozygous frameshift mutations are induced there, they likely cause the abortion of the embryo development. This leads to the observed seed phenotype (dark and wrinkled), which is typical for embryo-lethality (Meinke and Sussex, 1979). By contrast, in H982 #08 the fertilized egg cell will only have two copies of the three-base-pair deletion allele. This allele is protected from CRISPR/Cas9 editing, as the three-base-pair mutation prevents recognition by the sgRNA. Thus, no novel frameshift mutations can be induced and all seeds look

normal, because the three-base-pair deletion allele encodes the sufficiently functional UMK3$_{\Delta E76}$. This indicates, that UMK3 is not only required for male and female gametophyte development but also for correct embryo development.

As already mentioned, the three-base-pair deletion produced a 'weak' *UMK3* allele, which could be transferred into subsequent plant generations and the homozygous line *umk3$_{\Delta E76}$* could be generated. The protein encoded by that allele, UMK3$_{\Delta E76}$, showed reduced activity with UMP and CMP and almost no activity with dCMP (**Figure 16**). This impairment had a strong influence on the metabolome of *umk3$_{\Delta E76}$* plants in different growth stages. A lower content of UTP could be detected in *umk3$_{\Delta E76}$* during seed imbibition and germination, as well as in 18-day-old and two-month-old plants compared to the wild type (**Figure 20, Figure 25, Figure 26**). In contrast, a full knockout of *UMK1* or *UMK2*, or both, did not result in a lower UTP content in any of the analyzed developmental stages. Similarly, there is no additional reduction of UTP content, when *UMK1* or *UMK2*, or both, are mutated in *umk3$_{\Delta E76}$* background. This indicates, that the *in vivo* UMP phosphorylation is exclusively performed by UMK3. The biochemical *in vitro* characterization of the three UMKs supports this, as UMK3 possesses the lowest K$_M$ and highest k$_{cat}$ for UMP. Thus, UMK3 is a significantly better kinase for the phosphorylation of UMP than UMK1 or UMK2 (**Table 1-Table 3**). In contrast to CMP and dCMP, which are only generated from nucleoside salvaging reactions or directly from breakdown of RNA and DNA, UMP is also synthesized *de novo*. Consequently, more UMP is generated in the cell than CMP or dCMP, which makes the catalytic efficiency for the phosphorylation of UMP the most important feature of a UMK. As the last step of UMP *de novo* synthesis takes place in the cytosol (Witz et al., 2012), it appears plausible that also UMK3 is located there. It is evident, that UMK3 has an essential function in the pyrimidine nucleotide metabolism of Arabidopsis as a kinase for UMP and that this function cannot be fulfilled by any of the other UMKs.

### 4.1.2. UMK3 has a higher catalytic velocity than reported and shows substrate inhibition

The only characterization of an Arabidopsis UMK from the AMK-like UMK family was performed by Zhou et al., 1998 for UMK3. In the publication, the enzyme was biochemically characterized, but some discrepancies to the results presented in this thesis are apparent. Zhou et al., 1998 expressed *UMK3* cDNA fused to a glutathione-S-transferase gene in *Escherichia coli* cells and purified the proteins via gluthatione-Sepharose beads. UMK3 was then cleaved from the fused glutathione-S-transferase protein, purified by HPLC, dialyzed and concentrated. Characterization of the purified protein revealed similar K$_M$ values for UMP and CMP as found in this thesis. However, the k$_{cat}$ values were reported to be 33 times lower for UMP

and 23 times lower for CMP. There are several possible explanations for the lower catalytic velocity that was observed by Zhou et al., 1998: (1) Heterologous expression in bacteria can lead to incorrect folding of the protein or the formation of inclusion bodies consisting of inactive UMK3 aggregates. Although the protein isolates from Zhou et al., 1998 appear to be pure, the degree of protein integrity cannot be assessed. (2) Bacteria perform a relatively low number of posttranslational protein modifications and UMK3 possesses several potential sites for a posttranslational phosphorylations (**Figure A 1**). Missing modifications can influence the activity of heterologously expressed proteins. (3) The purification method employed by Zhou et al., 1998 included more steps than the StrepTactin chromatography used in this thesis. These additional steps, especially a HPLC purification with an acetonitrile gradient, could have led to a denaturation of UMK3. It appears likely that the lower $k_{cat}$ values reported by Zhou et al., 1998 are due to a lower integrity of their purified UMK3 proteins. Additionally, similar $k_{cat}$ values to the ones measured here were found for the UMKs from human and slime mold (Pasti et al., 2003).

Zhou et al., 1998 also detected minor activity of UMK3 with dCMP (without reporting any kinetic data), and postulated that this activity will not be sufficient for UMK3 to be a dCMP kinase *in vivo* and consequently a different enzyme must be responsible for the conversion of dCMP to dCDP. However, this conclusion is likely incorrect based on the metabolomic data shown in **Figure 21**, showing an accumulation of dCMP in the *umk3$_{\Delta E76}$* line, indicating that UMK3 is significantly contributing towards the phosphorylation of dCMP *in vivo*. In fact, with an intact UMK3, no dCMP accumulation can be observed for any of the mutant lines at the different growth stages, except for *umk2* during germination (**Figure 27**). This demonstrates that UMK3 is responsible for most of the dCMP phosphorylation *in vivo*. The rather low catalytic efficiency of UMK3 with dCMP (**Table 3**) is probably compensated by the strong expression (**Figure 24**). Additionally, the amount of dCMP generated in the cell is rather small compared to UMP or CMP. Therefore, less total activity is required to phosphorylate the produced dCMP.

Another phenomenon that was detected here and not reported by Zhou et al., 1998 is the substrate inhibition that was observed for UMK3 with UMP or CMP as substrate (**Figure 5**). The authors do not state what concentrations of UMP or CMP they used for determining the $K_M$ and $k_{cat}$ of UMK3. It is possible that they did not reach the concentrations, where substrate inhibition becomes apparent, although this would have been required to record a complete Michaelis-Menten curve. Another possibility is, that substrate inhibition requires *in planta* posttranslational modifications of the enzyme. While Zhou et al., 1998 did not report the inhibition, Pasti et al., 2003 detected a substrate inhibition with the human UMK for UMP and CMP, but not dCMP. Also for prokaryotic UMKs a substrate inhibition of varying intensities is reported (Evrin et al., 2007). The substrate inhibition may be caused by a nonproductive binding of the NMP to the

NTP binding site, which is not as strict in its binding specificity (Yan and Tsai, 1999). A similar regulative mechanism is observed for the deoxynucleoside kinase of Drosophila, where ATP can nonproductively bind to the NMP binding site of the enzyme (Johansson et al., 2001). Overall, the results of Pasti et al., 2003 with regard to the kinetic properties of the human enzyme are very similar to the results obtained here for UMK3. Thus, it is possible that the observed substrate inhibition is indeed a genuine function of AMK-like UMKs although it is still unclear if these kinetic properties are of any relevance *in vivo*, as it occurs at concentrations of UMP that likely exceed the cytosolic concentration. Based on some assumptions, the cytosolic UMP concentration can be estimated from the generated metabolite data. Studies on the intracellular distribution of UMP show that it is preferentially localized in the cytosol and not in the vacuoles (Dancer et al., 1990; Oikawa et al., 2011). This is supported by the fact, that the last step of the *de novo* synthesis of UMP takes place in the cytosol (Witz et al., 2012). Based on the formula for the calculation of intracellular nucleotide concentrations from Straube et al., 2021, the estimation of the cytosolic volume fraction of the total cell volume from Koffler et al., 2013 and the here measured concentration of UMP (**Figure A 5**), the cytosolic UMP concentration can be estimated to be approximately 300 µM (**Calculation A 1**). This estimate does not take into account that the UMP pools measured here were falsely enlarged by a low degree of degradation of the abundant UDP-sugars, and it is assumed that all cellular UMP is located in the cytosol. In reality, therefore, the concentration is probably lower. The estimate is in the same range as UMP quantifications of various mammalian tissues, where the average UMP concentration was about 180 µM (Traut, 1994). The inhibitory effect at 1 mM UMP, which corresponds to at least three times the estimated amount of cytosolic UMP, is only about 30 % *in vitro* (**Figure 5**). Thus, it appears questionable whether the substrate inhibition is relevant *in vivo*. However, the *in vivo* situation is different to the assay conditions, as the environment of the plant cell is way more complex. It is possible that other factors influence the activity of UMK3 in the cell and contribute to an *in vivo* regulation. In a stress situation where the plant is unable to maintain the phosphorylation of *de novo* synthesized UMP and an accumulation occurs, substrate inhibition of UMK3 could serve as a mechanism to further enhance this accumulation. If this is the case, the substrate inhibition could be involved in the feedback regulation of UMP *de novo* synthesis, which is mediated by UMP in the plastids (Bellin et al., 2021a). Such a regulatory mechanism could be beneficial in situations where the plant is exposed to biotic or abiotic stress factors that require a redirection of resources to cope with the stress. A negative regulation of UMP *de novo* synthesis governed by substrate inhibition of UMK3 could prevent the plant from investing resources into growth in the short term and free those resources to overcome the stress.

## 4.1.3. UMK3 is vital for synthesis of UDP-sugars and normal plant growth

The lower UTP concentration observed in plants of the $umk3_{\Delta E76}$ line also causes a reduction in the content of UDP-sugars (**Figure 22**). In their function as precursors for cellulose and hemicellulose synthesis, UDP-sugars are significantly involved in plant cell wall synthesis and thus, biomass production. Mutation of different genes involved in UDP-sugar metabolism result in reduced plant growth and other morphological abnormalities (Rösti et al., 2007; Park et al., 2010; Zhao et al., 2018). While mutation of *UMK3* does not directly interfere with UDP-sugar synthesis, it does lead to a lower availability of precursors and subsequently, to a lower content of UDP-sugars. This lower UDP-sugar content, especially of UDP-glucose, is likely the reason for the reduced leaf area associated with the *UMK3* mutation (**Figure 19**). As observed for UTP, the mutation of *UMK1* and/or *UMK2* does not seem to have any influence on the UDP-sugar content of 18-day-old plants either.

Mutation of *UMK3* has varying effects on the different UDP-sugars. Some UDP-sugar contents in the $umk3_{\Delta E76}$ line are reduced by as much as 80 %, while others show a reduction of only 30 % (**Figure 22, Figure A 7**). The various UDP-sugars can either be synthesized directly from UTP and the respective sugar-1-phosphate or in a series of enzymatic steps from UDP-glucose. The fact that the observed reductions seem to not follow a pattern based on their position in the synthesis pathway, makes it appear more likely that the varying reductions are caused by the direct synthesis of the respective UDP-sugar and not by conversions. However, this is just an assumption and cannot be said with certainty, as the UDP-sugar metabolism is complex. Arabidopsis possesses two UDP GLUCOSE PYROPHOSPHORYLASEs (UGPs) that specifically synthesize UDP-glucose (Meng et al., 2008). The UGP activity ensures a steady supply of UDP-glucose, which could be the explanation why the reduction of UDP-glucose in $umk3_{\Delta E76}$ is not as prominent compared to some other UDP-sugars. The other UDP-sugars are synthesized by UDP SUGAR PYROPHOSPHORYLASE (USP) which has varying catalytic efficiencies for the formation of the different UDP-sugars (Kotake et al., 2007). It could be assumed that in a UTP deficiency situation, the UDP-sugars, for which the catalytic properties of USP are better, are preferentially synthesized. However, this cannot be observed. USP shows the highest catalytic efficiency with glucuronate-1-phosphate, but UDP-glucoronate shows a strong reduction of 80 % in $umk3_{\Delta E76}$. In contrast, only a small reduction is visible for UDP-xylose, while xylose-1-phosphate is one of the poorer substrates of USP (Kotake et al., 2007). In contrast, *USP* knockdown lines were observed to accumulate arabinose, indicating less synthesis of UDP-arabinose (Geserick and Tenhaken, 2013). This observation is in agreement with the strong reduction of UDP-arabinose observed in the $umk3_{\Delta E76}$ line. The varying reductions in the content of the different UDP-

sugars indicates that there is a regulatory mechanism in UDP-sugar synthesis that governs their formation in a UTP deficiency situation.

Although there are no differences in UDP-sugar contents detected between $umk3_{\Delta E76}$ and the triple mutant in the 18-day-old plants, there is a significant growth reduction in the 35-day-old triple mutant plants. This indicates that during the later stages of the growth phase, the other UMKs may become more important for UTP synthesis. In fact, when looking at the expression data in **Figure 24**, the expression of *UMK3* is lower in mature leaves, indicating a decreased role of UMK3 in fully developed plant leaves. This suggests that formation of UTP and subsequent synthesis of UDP-sugars can partially be taken over by UMK1 and UMK2 during later growth stages. The strong growth reduction only appears in the triple mutant, but not *umk1 umk3*$_{\Delta E76}$ or *umk2 umk3*$_{\Delta E76}$. This implies that the double mutants still possess enough UMP phosphorylation capacities to ensure sufficient supply of UDP-sugars to prevent this severe phenotype.

### 4.1.4. UMK3 supports mitochondrial DNA replication during germination

Nucleotide quantification from germinating seeds of the $umk3_{\Delta E76}$ line revealed similar alterations in pyrimidine nucleotide contents as in the 18-day-old plants (**Figure 26, Figure 27**). Compared to the wild type, the $umk3_{\Delta E76}$ line has a lower content of UTP and CTP 24 hours after imbibition and a lower dCTP content 24 hours after transfer to the light. The reduced pyrimidine NTP contents can also be observed at all later time points. Although *UMK3* is expressed comparably low during imbibition, the impairment of UMK3$_{\Delta E76}$ is negatively affecting the pyrimidine NTP amounts there. This shows that UMK3 already has its central role in pyrimidine metabolism during imbibition, although *UMK1* and *UMK2* are more strongly expressed. At least for UMK2 a higher protein abundance than for UMK3 can be confirmed in imbibed seeds (Mergner et al., 2020). The central role of UMK3 can likely be attributed to its better catalytic efficiency in the phosphorylation of rNMPs. After the seeds are transferred to light, also the dCTP content is reduced in the $umk3_{\Delta E76}$ line. An explanation for this delayed reduction compared to the rNTPs could be that nuclear and organellar DNA replication only start after the seeds are moved into light. Consequently, dCTP is not consumed before. After the DNA replication starts, the dCTP content in $umk3_{\Delta E76}$ decreases and supply cannot be sustained. This decrease of dCTP also negatively affects the mitochondrial DNA replication (**Figure 28**), which is interesting because mitochondria have their own UMK (UMK2) to generate dCTP to supply precursors for DNA replication. This supply seems to be insufficient however, and import of pyrimidine nucleotides generated in the cytosol by UMK3 is necessary for normal mitochondrial DNA replication during germination.

### 4.1.5. Summarizing the function of UMK3 and its role in pyrimidine metabolism of Arabidopsis

The results of this work show that UMK3 is universally involved in pyrimidine metabolism in Arabidopsis at all analyzed growth stages. This conclusion greatly benefited from the coincidental generation of a 'weak' *UMK3* allele (*UMK3$_{\Delta E76}$*), which allowed insights into the pyrimidine metabolism that would otherwise not have been possible. These are the main findings regarding UMK3 made in this project:

- UMK3 is the main UMK in the pyrimidine metabolism of Arabidopsis
- Knockout of UMK3 is lethal for the egg cell, the pollen and likely the embryo
- UMK3 localizes to the cytosol, where it is catalyzing most of the UMP phosphorylation
- The enzyme has a higher catalytic velocity than described in the literature and exhibits a substrate inhibition
- It is also the main dCMP kinase *in vivo*
- UDP-sugar and biomass production are reliant on UMK3
- UMK3 supports the mitochondrial DNA replication during germination showing that mitochondria can import dCDP or dCTP from the cytosol.

## 4.2. Characterization of UMK2

### 4.2.1. UMK2 localizes to mitochondria and is involved in deoxycytidine salvage

UMK2 is located in mitochondria (**Figure 7**). Mitochondria have their own genome and transcription machinery, which require NTPs. Consequently, it is plausible that plants possess a mitochondrial enzyme that can synthesize these compounds directly where they are required. Other NMP kinases, like GMP kinase (Sugimoto et al., 2007) or TMP kinase (Ronceret et al., 2008) are shown to also have a mitochondrial localization. Interestingly, these enzymes dual-localize to mitochondria and plastids, which has not been observed for UMK2. Mitochondrial localization makes it unlikely that UMK2 is overly involved in the phosphorylation of *de novo* synthesized UMP. As discussed earlier, UMK3 is the main UMP kinase *in vivo*. UMK2 would thus only be responsible for recycling NMPs that originate from the breakdown of DNA and RNA or from salvage pathways inside mitochondria. As mitochondrial nucleotide pools are small in comparison to the cytosolic pools, alterations of the mitochondrial nucleotide content may not always be detectable in whole-cell extracts (Niehaus et al., 2022).

Deoxynucleosides are salvaged in Arabidopsis by three different DEOXYNUCLEOSIDE KINASEs (dNKs). Two of those (TK1a and TK1b) are specific for thymidine, while a broad substrate dNK is responsible for the phosphorylation of deoxyadenosine, deoxycytidine and deoxyguanosine (Clausen et al., 2012). It has been postulated that the broad substrate dNK localizes to mitochondria (Clausen et al., 2014), but also a plastidic localization has been observed (our unpublished data). Consequently, dCMP originating from deoxycytidine salvage is generated inside mitochondria, where UMK2 could further phosphorylate it to make it available for DNA synthesis. Indeed, there are hints in the metabolic data showing an involvement of UMK2 in dCMP phosphorylation. The 18-day-old *umk2 umk3$_{\Delta E76}$* double mutant accumulates more dCMP than the plants of the *umk3$_{\Delta E76}$* line (**Figure 21**) and the *umk2* single mutant is accumulating dCMP compared to the wild type during germination (**Figure 26**). Especially during germination, where dCMP is released from DNA damage repair (Waterworth et al., 2015), UMK2 activity seems to be important for the recycling of this dCMP. This is also supported by the kinetic and expression data. The catalytic efficiency with dCMP is higher for UMK2 than for UMK3 (**Table 2, Table 3**) and *UMK2* is stronger expressed during seed imbibition. During the vegetative growth phase, where probably less dCMP is generated in the cell, the loss of UMK2 hardly influences the global dCMP concentration. Here, *UMK3* is stronger expressed and UMK3 is able to catalyze most dCMP phosphorylation. The disappearance of dCMP accumulation from germinating seeds to 18-day-old plants suggests that the mitochondria are probably able to export dCMP to the cytosol to make it available for UMK3. However, it is also possible that mitochondrial dCMP accumulation is not detectable during later growth stages.

## 4.2.2. Loss of UMK2 does not affect plant growth as mitochondria can import nucleotides

The *UMK2* gene is highly conserved in vascular plants (**Figure 3**). Consequently, it was surprising that a null-mutation of the gene did not cause any phenotypical abnormalities and only small alterations in metabolite content compared to wild type plants (**Figure 17, Figure 18, Figure 19, Figure 26, Figure 27**). Mutation of gene encoding the mitochondrial and plastidic TMP KINASE is embryo-lethal and a knockout of the gene encoding mitochondrial and plastidic GMP KINASE leads to pale leaves in rice (Sugimoto et al., 2007; Ronceret et al., 2008). However, these phenotypes are likely caused by the absence of the enzymes from the plastid. Nevertheless, the fact that the *umk2* plants are able to grow normally without their mitochondria being able to generate pyrimidine NTPs autonomously implies, that mitochondria are able to meet their demands by importing NDPs or NTPs. While the outer membrane of mitochondria is permeable for smaller molecules, the inner membrane is more tightly packed and metabolite exchange is only possible via carrier proteins. Arabidopsis possesses 58 potential mitochondrial carrier proteins

belonging to different families (Picault et al., 2004). One of these carrier families are the adenylate carriers, which exchange mitochondrial ATP for cytosolic ADP (Da Fonseca-Pereira et al., 2018). The adenylate carriers are the most abundant mitochondrial inner-membrane carriers (Fuchs et al., 2020) and it is possible that they are able to channel small amounts of other nucleotides, as it has been demonstrated for the ADENINE NUCLEOTIDE TRANSPORTER 1 (Palmieri et al., 2008). It is also possible that a yet undescribed mitochondrial pyrimidine nucleotide carrier exists in Arabidopsis. Mitochondrial pyrimidine carriers have been identified in yeast and human (Marobbio et al., 2006; Floyd et al., 2007). The metabolic data supports the existence of pyrimidine nucleotide transport across the inner mitochondrial membrane in both directions. There were no differences in the dCMP or CMP contents of 18-day-old *umk2* plants compared to the wild type (**Figure 21**). Consequently, the pyrimidine NMPs, which are constantly generated from RNA and DNA breakdown or deoxycytidine salvage must be able to leave the mitochondria. Alternatively, they are dephosphorylated to nucleosides by a yet unknown UMP/CMP phosphatase (Witte and Herde, 2020). Although this is a possibility, it does not seem plausible, as the presence of such a phosphatase in the mitochondria would counteract the dNK salvaging reaction. Thus, an export mechanism for pyrimidine NMPs to the cytosol probably exists. There are also hints, that mitochondria are able to export pyrimidine NDPs or NTPs. The *UMK* triple mutant shows a significantly reduced growth compared to the *umk1 umk3*$_{\Delta E76}$ double mutant. As established earlier, this is likely caused by lower UTP production and subsequently less UDP-sugar synthesis. In the *umk1 umk3*$_{\Delta E76}$ line, UMK2 is able to support UDP-sugar synthesis, which implies that the mitochondria are able to import UMP and export UDP or UTP to the cytosol.

### 4.2.3. Loss of UMK2 causes a nucleotide deficiency in mitochondria affecting mtDNA replication

The metabolic data shows, that UMK2 activity has no influence on global nucleotide pools in Arabidopsis plants during the vegetative growth phase (**Figure 20, Figure 21**). This can be interpreted to mean that there is no metabolic effect caused by the *UMK2* mutation in those plants. However, it is possible that local nucleotide deficiencies occur inside the mitochondria, which are not detectable in whole cell extracts. This has been previously observed (Niehaus et al., 2022) and is supported by the quantification of mitochondrial genome copies during germination (**Figure 28**). The *UMK2* single mutant has a lower mtDNA copy number than wild type plants, although the global dCTP content is on the same level as in the wild type (**Figure 20**). This impairment of mitochondrial DNA replication is likely caused by a local dCTP deficiency inside the mitochondria, which is not detectable in the global pool. The DNA replication is even more hindered in the *umk2 umk3*$_{\Delta E76}$ line, showing that an additional reduction in the global dCTP content

is even more detrimental. This indicates that the early mitochondrial DNA replication is relying on dCTP generated inside the mitochondria, as well as dCDP or dCTP imported from the cytosol. Interestingly, a reduction of the global CTP content is visible in the *umk2* line during imbibition. This indicates that UMK2 is also involved in CMP phosphorylation in this early developmental stage, where *UMK3* is not yet strongly expressed. A lack of CTP here could also negatively influence the mRNA synthesis required for production of the DNA replication machinery and the production of tRNAs, which both could contribute to the observed reduction in mtDNA copy number.

With an impaired mitochondrial DNA replication and potentially less RNA building blocks inside mitochondria, one could assume that this negatively influences early plant development or mitochondrial protein expression. The plant development was assessed by analyzing the germination rate of mutant seeds (**Figure 17**) and leaf area quantification of seven-day-old seedlings (**Figure 18**). No deviations from the wild type could be observed for the *umk2* lines. However, seeds of the *umk2 umk3*$_{ΔE76}$ line showed a lower germination rate compared to wild type seeds, which might be related to the impaired mtDNA replication in this line (**Figure 28**). To assess whether the *UMK2* mutation has an effect on protein expression, the protein content in germinating *umk2* seeds was analyzed by shotgun proteomic MS analysis (**Figure 29**). No significant changes were detected in the proteome of the *umk2* line in comparison to the wild type. As mentioned earlier, a power analysis revealed that changes of protein expression in the same range as the observed reduction in mtDNA copy number would have not been detectable by the employed method, but there is also no tendency towards a lower abundance of mitochondria-encoded proteins observable. It appears that the mitochondria and the plant in general are unaffected by a lower genomic mtDNA copy number. On average, the ratio of mitochondria to full mtDNA copies inside the cell of a wild type Arabidopsis plant is three to one, meaning that there are more mitochondria than full genomic mtDNA copies (Preuten et al., 2010). The mitochondria are able to compensate this by fusion and fission, a mechanism to exchange and homogenize their contents. This mechanism may be the reason why a lower mtDNA copy number does not result in further impairments for the plant under the employed growth conditions. However, the growth conditions in the laboratory represent an environment with a low level of biotic and abiotic stress factors. The evolutionary conservation of *UMK2* in all vascular plant species indicates that the gene confers a selective advantage. Thus, it is possible that negative effects of a *UMK2* mutation would become visible when the plants are exposed to harsher growth conditions as they occur in nature.

**4.2.4. Summarizing the function of UMK2 and its role in pyrimidine metabolism of Arabidopsis**

The results leave open questions about the function of *UMK2* and why it is universally conserved in vascular plants. Nevertheless, some hints about the role of UMK2 in the pyrimidine metabolism of Arabidopsis were found:

- UMK2 is located in mitochondria

- A null-mutation of *UMK2* does not cause any visible phenotypes

- Mutation of *UMK2* causes changes in pyrimidine nucleotide content during germination, but not in older plants

- UMK2 is mainly involved in dCMP phosphorylation *in vivo* and likely operates downstream of the mitochondrial dNK, which salvages deoxycytidine

- UMK2 has a more pronounced role during germination, where it supplies dCTP for the mitochondrial DNA replication

## 4.3. Characterization of UMK1

**4.3.1. UMK1 is a less-active isozyme of UMK3 that is not universally conserved**

UMK1 is closely related to UMK3, as both proteins have a sequence identity of 75 % (**Figure A 1**). While the gene encoding UMK3 is universally conserved in all analyzed vascular plants, not all of them also have a gene encoding UMK1 (**Figure 3**). A clear distinction between *UMK1* and *UMK3* based on their sequence is possible, but the genes that were annotated as *UMK4* could also be grouped as *UMK1*, which would extend the number of plant families having a *UMK1* gene. Still, *UMK1* is absent in various species, showing that in these plants its function is not required and may be taken over by *UMK2* or *UMK3*.

The results of the phylogenetic analysis indicate a subordinate role for UMK1 in plant pyrimidine nucleotide metabolism. This assumption is consistent with the results obtained in this work for the *UMK1* mutant lines. No phenotypic or metabolomic changes were observed in any of the developmental stages analyzed in comparison with wild type plants. Some metabolomic phenotypes were detected in the *umk1 umk3_{ΔE76}* line compared to the *umk3_{ΔE76}* single mutant, which indicates that UMK1 is able to compensate to a certain extent for a partially compromised UMK3. This partial compensation is possible, because they share the cytosolic localization (**Figure 7**) and both are active with UMP, CMP and dCMP (**Table 1, Table 3**). However, the catalytic efficiencies of UMK1 for the phosphorylation of UMP and CMP are lower than

those of UMK3. Additionally, *UMK3* is stronger expressed during most developmental stages of the plant. Thus, this only allows for a partial compensation of the lower UMK3$_{\Delta E76}$ activity by UMK1. But what is the purpose of UMK1, if it is only relevant when UMK3 activity is reduced? It is possible, that UMK1 becomes important under growth conditions, which deviate from the laboratory conditions that were used here. As discussed above, UMK3 may be subject to a stress-based regulatory mechanism that leads to lower UMK3 activity, as suggested by substrate inhibition. This mechanism is not present to such an extend for UMK1 and UMK2. While utilizing resources for the phosphorylation of *de novo* synthesized UMP must probably be avoided in a stress situation, the rNMPs generated from RNA breakdown and salvaging still occur inside the cell and need to be recycled. For these recycling processes, a UMK with a lower catalytic efficiency such as UMK1 may be sufficient. The 'weak' *UMK3$_{\Delta E76}$* allele encodes a less-active version of UMK3, which may mimic an *in vivo* situation, where UMK3 activity is reduced based on a regulative mechanism. Thus, metabolomic phenotypes are visible in *umk1 umk3$_{\Delta E76}$* compared to *umk3$_{\Delta E76}$* and a function of UMK1 is unmasked in *umk3$_{\Delta E76}$* background.

## 4.3.2. UMK1 is mainly functioning as a CMP kinase and may be involved in nucleotide balancing

The metabolomic phenotypes observed in the *umk1 umk3$_{\Delta E76}$* line compared to the *umk3$_{\Delta E76}$* line only affect the cytidylates and deoxycytidylates. In 18-day-old plants, the *umk1 umk3$_{\Delta E76}$* line accumulates CMP compared to the wild type, which cannot be observed for *umk3$_{\Delta E76}$*. Additionally, *umk1 umk3$_{\Delta E76}$* has a lower CTP content than *umk3$_{\Delta E76}$*. This indicates, that UMK1 is able to compensate the lower CMP phosphorylation capacity of UMK3$_{\Delta E76}$ during the vegetative growth phase. A similar effect cannot be detected for UTP, which is in line with the better catalytic efficiency of UMK1 with CMP than UMP (**Table 1**). Thus, during the growth phase, UMK1 is mainly functioning as a CMP kinase. A stronger accumulation of dCMP in *umk1 umk3$_{\Delta E76}$* cannot be observed in comparison to *umk3$_{\Delta E76}$*, suggesting that UMK1 is not involved in dCMP phosphorylation during the vegetative growth phase. However, during germination, a stronger accumulation of dCMP can be observed in *umk1 umk3$_{\Delta E76}$* compared to *umk3$_{\Delta E76}$*, indicating an involvement of UMK1 in dCMP phosphorylation during this developmental stage.

The function of UMK1 as a CMP kinase may be important for nucleotide balancing inside the cell. This becomes apparent, when looking at the UTP/CTP ratios in the 18-day-old mutant plants (**Figure 20 B**). In all mutant lines, a UTP:CTP ratio of about 2:1 is maintained, except for the lines where *UMK1* is knocked out in the *umk3$_{\Delta E76}$* background. Here the UTP:CTP ratio is shifted towards UTP (3:1). Thus, UMK1 may function in balancing the UTP/CTP pools, in particular when UMK3 activity is low, which might be mimicked

in *umk3~ΔE76~* as mentioned above. An imbalance in the nucleotide pools can negatively influence cell proliferation (Diehl et al., 2022). However, no growth phenotypes were observed for *umk1 umk3~ΔE76~* in comparison to *umk3~ΔE76~*. Possibly because the imbalance observed here is too small to cause any growth defects under the employed growth conditions. The involvement of UMK1 in CMP phosphorylation becomes even more apparent during germination, where the *umk1 umk3~ΔE76~* line strongly accumulates CMP, while the *umk3~ΔE76~* single mutant only shows a slight accumulation compared to the wild type (**Figure 27**). More CMP is released from the breakdown of parental RNA during germination compared to the growth phase (Li et al., 2006). UMK1 supports the recycling of this CMP, which may become even more important, if UMK3 activity is lowered during germination.

### 4.3.3. Summarizing the function of UMK1 and its role in pyrimidine metabolism of Arabidopsis

The results regarding UMK1 leave room for speculations about its dispensability for Arabidopsis. No clear physiological role can be deduced from the data generated in this thesis, but it was possible to theorize, in which situations UMK1 could be important. Here, the main findings about UMK1 are summarized:

- UMK1 is a less-active isozyme of UMK3, that also localizes to the cytosol
- A knockout of *UMK1* does not cause any developmental or metabolic phenotypes
- UMK1 is mainly functioning as a CMP kinase *in vivo*
- During germination, UMK1 is also involved in dCMP phosphorylation
- Its role in pyrimidine metabolism is likely limited to conditions, where UMK3 is less active

### 4.4. Pyrimidine metabolism of Arabidopsis during the vegetative growth phase

During vegetative growth, plants have a constant demand of pyrimidine nucleotides to sustain biomass accumulation. UDP-sugars are required for cell wall synthesis and nucleotide triphosphates are incorporated into DNA and RNA to maintain cellular processes. In the introduction, a basic model of the pyrimidine nucleotide metabolism was shown (

**Figure 1**). This model was revised based on the findings of this thesis (**Figure 30**).
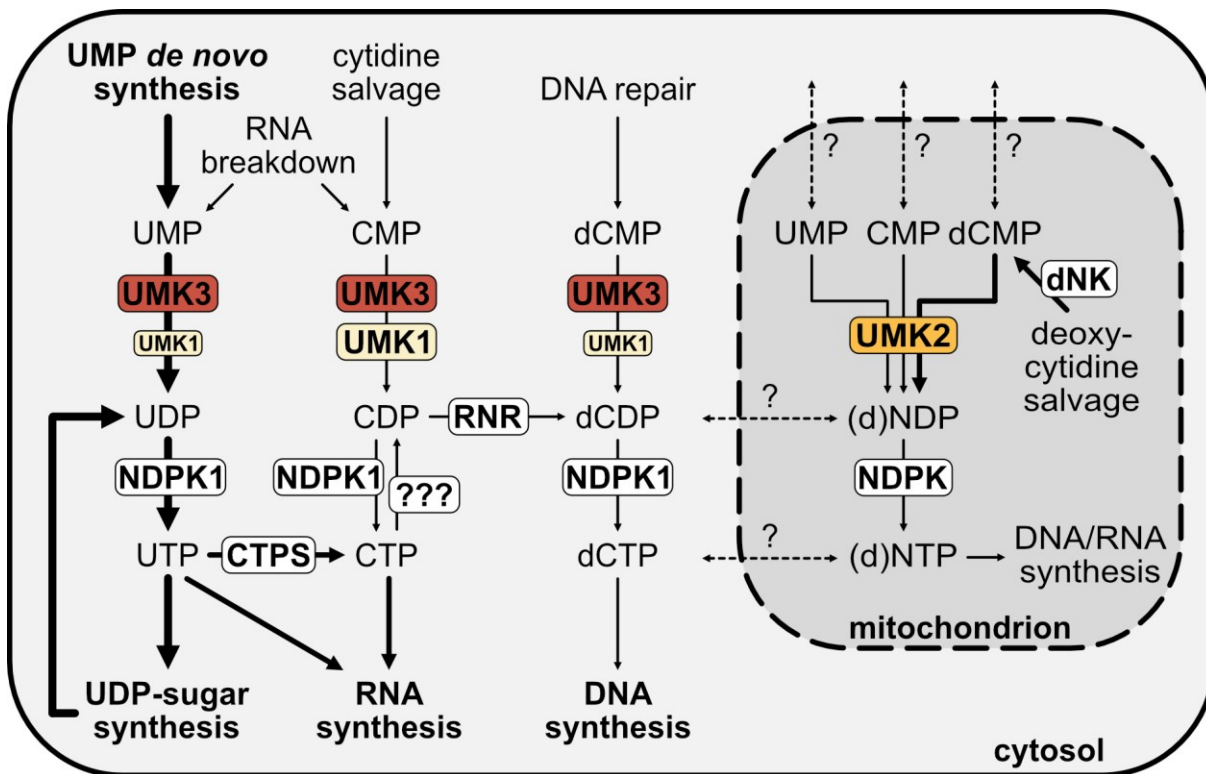
**Figure 30. Revised model of pyrimidine nucleotide metabolism.**
The presumed metabolic flux through the pathway is indicated by the thickness of the arrows. Thicker arrows indicate a higher flux. It should be noted that the arrows are not drawn to scale, as the actual quantities are not known. The thickness is only used to illustrate the assumptions made in the discussion. Similarly, the size of the boxes around the various UMKs is intended to indicate their involvement in the reaction where they are positioned. UMK, UMP KINASE; NDPK, NUCLEOSIDE DIPHOSPHATE KINASE; CTPS, CTP SYNTHASE; RNR, RIBONUCLEOTIDE REDUCTASE; dNK, DEOXYNUCLEOSIDE KINASE.

## 4.4.1. The main route of pyrimidine nucleotide synthesis in the cytosol

UMP is the entry molecule of pyrimidine nucleotide metabolism as it represents the precursor for the synthesis of all other pyrimidine nucleotides. In contrast to CMP and dCMP, UMP can be synthesized *de novo*. All three pyrimidine NMPs are also constantly generated in the cell from DNA- and RNA-turnover or from (deoxy)nucleoside salvage pathways. Recycling of these NMPs via the UMKs can support the NTP production, but the majority of UMK activity is likely devoted to the phosphorylation of *de novo* synthesized UMP. The plant cell has a high demand for UTP, which is used for the synthesis of CTP, phospholipids derived from intermediates activated by reaction with CTP, UDP-sugars or directly incorporated into RNA. Thus, it is plausible that the UMP phosphorylation is performed by an enzyme with a high abundance and catalytic velocity like UMK3. UMK3 is able to catalyze up to 250 reactions per second. In comparison, for Arabidopsis GMP KINASE a $k_{cat}$ value of 4 reactions per second was reported

(Kumar et al., 2000). It is possible however, that this reported value is too low for the same reasons that were listed earlier for the UMK characterization by Zhou et al., 1998. Cytosolic GMP KINASE from rice is able to catalyze 70 reactions per second (Nomura et al., 2014). Thus, UMK3 appears to have a rather high catalytic velocity in comparison to similar enzymes. The high catalytic rate is essential for maintaining a high metabolic flux for UTP production. It is also important that the UMP pools are subjected to a high metabolic flux to prevent an accumulation resulting in feedback inhibition of *de novo* synthesis (Bellin et al., 2021a), which would delay plant development.

The further phosphorylation of UDP to UTP is catalyzed by enzymes of the NDPK family. NDPKs accept all NTPs and all NDPs as substrates, but likely mainly utilize ATP as phosphate donor *in vivo* as it is the most abundant NTP in the cell. NDPKs also catalyze the phosphate transfer with a high velocity to keep up the flux towards the synthesis of UTP. Turnover rates of up to 2000 reactions per second are reported (Johansson et al., 2008). UTP synthesis is relying mainly on the cytosolic NDPK1, which is responsible for the majority of NDPK activity *in vivo* and present in all tissues. An especially high expression can be detected in meristematic cells (Dorion et al., 2006). NDPK1 thus has an important role for plant growth as it is synthesizing the UTP, which can subsequently be used for production of UDP-sugars, directly in the developing tissue where it is required. Due to the same subcellular localization, it can confidently be assumed that NDPK1 is acting downstream of UMK3 and is responsible for the phosphorylation of UDP generated from *de novo* synthesized UMP.

The UTP can then be used for the synthesis of UDP-sugars (mainly UDP-glucose) for biomass accumulation, the synthesis of the RNA building block CTP or as a building block of RNA itself. It can be assumed that the demand of the cell for CTP and UTP as a building block of RNA is similar. The synthesis of UDP-glucose from UTP is catalyzed by UGP and CTP is generated from UTP by CTPS. Comparing the enzymatic activities and abundances of the enzymes indicates that the flux towards the production of UDP-glucose is likely way higher than the flux towards CTP. UGP has much better catalytic properties than CTPS (Meng et al., 2008; Daumann et al., 2018) and it is one order of magnitude more abundant in most tissues (Mergner et al., 2020). CTPS is also inhibited by its product CTP (Noree et al., 2014; Daumann et al., 2018). It has to be noted though, that Arabidopsis possesses five CTPS isoforms, of which only one has been biochemically characterized so far. After the glucose of UDP-glucose is incorporated into the cell wall, UDP is released and can again be phosphorylated to UTP. Thus, there should be no net loss of uridinylates by UDP-glucose synthesis. Apart from nucleotide catabolism, the only way to remove uridylates from the cell is via the CTPS reaction or incorporation of UTP into RNA. Both of these processes should only metabolize small

amounts of uridylates. Still, a high catalytic efficiency of the enzymes involved in the pathway up to the production of UTP is plausible, as the accumulation of biomass increases the demand for more UDP-sugars.

The enzyme complex RNR is responsible for the production of deoxynucleotides and requires rNDPs as substrates. This means that the CTP obtained from the CTPS reaction has to be dephosphorylated, before it can serve as a substrate of the RNR. It is speculated that a dedicated CTP phosphatase might exist (Witte and Herde, 2020). However, this may not be necessary as there are other CDP sources in the cell. CMP can be obtained from cytidine salvage or RNA turnover. The CMP can be phosphorylated to CDP in the cytosol by either UMK1 or UMK3. Both enzymes appear to be equally able to handle the amount of CMP generated during the vegetative growth phase, evident by the fact that CMP only accumulated in the $umk1\ umk3_{\Delta E76}$ line (**Figure 21**). The metabolomic data indicate that both of these processes produce enough CMP to have an influence on the global CTP concentration, evident by the lower CTP content in $umk1\ umk3_{\Delta E76}$ compared to $umk3_{\Delta E76}$ (**Figure 20**). This implies that RNA breakdown and/or cytidine salvage produce non-negligible amounts of cytidylates, which may be sufficient to cover the demand of CDP required for the RNR reaction. Alternatively, the reaction catalyzed by the NDPK1 may also release CDP. As the NDPK enzymes are unspecific in their substrates, it is likely that they will also utilize a certain amount of CTP. In fact, it was shown that the NDPK1 of rice has its best catalytic efficiency with CTP as phosphate donor (Kihara et al., 2011). CDP is also released from the phospholipid precursors CDP-DAG and CDP-Etn (Nakamura, 2017). Additionally, the cell has a lower demand for dNTPs than rNTPs, evident by the fact that the dNTP concentrations are more than three orders of magnitude lower than their rNTP counterparts (**Figure A 5** and **Figure A 6**). These facts make it questionable, whether a dedicated CTP phosphatase is even necessary to fuel the RNR reaction *in vivo*.

### 4.4.2. Organellar pyrimidine metabolism during the growth phase

Mitochondria and plastids also possess the necessary enzymes for the synthesis of NTPs. In mitochondria, UMK2 catalyzes the phosphorylation of UMP, CMP and dCMP. Like in the cytosol, NMPs are also constantly generated in the mitochondria from DNA repair and RNA breakdown. As there is no *de novo* synthesis of UMP in mitochondria, the total amount of NMPs generated in mitochondria is much lower compared to the cytosol. Consequently, not that much UMK activity is required to catalyze their phosphorylation, which explains the lower catalytic efficiency of UMK2 with UMP and CMP in comparison to UMK3 (**Table 2, Table 3**) and the lower abundance of UMK2 protein (Mergner et al., 2020). The catalytic efficiency of UMK2 with

dCMP is better compared to UMK3, as the phosphorylation of dCMP, obtained from deoxycytidine salvage, appears to be the main function of UMK2 *in vivo*.

As already mentioned, mitochondria are probably able to import and export pyrimidine nucleotides of different phosphorylation levels. As *umk2* plants are able to survive without any obvious growth deficits, the mitochondria must be able to import nucleotides as either NDPs or NTPs. Studies on the localization of the mitochondrial NDPK suggest, that it is located in the mitochondrial intermembrane space or anchored to the inner mitochondrial membrane, where it is shown to interact with an adenine nucleotide translocator (Sweetlove et al., 2001; Knorpp et al., 2003). With that in mind, it appears unlikely that the nucleotides are imported as NDPs, as they are likely phosphorylated by the NDPK while in the intermembrane space. This mechanism could also facilitate ATP export from the mitochondrial matrix by lowering the ATP concentration in the intermembrane space and the generated ADP could be directly re-imported (Knorpp et al., 2003). It is not clear whether the mitochondrial NDPK is also present in the mitochondrial matrix or if the matrix anchored protein is able to exert its activity inside mitochondria. If not, mitochondria would have to export their pyrimidines as NDPs to then re-import them as NTPs. This import and export mechanism would work in the opposite direction as the adenylate carriers, which export ATP and import ADP. If this theory is true, a yet uncharacterized mitochondrial pyrimidine carrier likely exists.

Like mitochondria, plastids also possess their own genome and transcriptional machinery and require NTPs. Plastids possess NMP kinases for all nucleotides and NDPKs that phosphorylate NDPs. While the plastid NMP kinases for GMP and TMP are encoded by the same gene as the mitochondrial versions (Sugimoto et al., 2007; Ronceret et al., 2008), the UMK (PUMPKIN) is encoded by a different gene belonging to the family of eubacterial UMKs (Schmid et al., 2019). PUMPKIN shares a high sequence similarity to the UMK of *Escherichia coli*, which is specific for UMP and does not have activity with CMP or dCMP (Serina et al., 1995). Consequently, it is questionable, whether PUMPKIN is active with these substrates and whether plastids are able to generate CDP and dCDP. PUMPKIN has been shown to be active with UMP, but the specific activity is almost two orders of magnitude lower than the specific activity of UMK2 (Schmid et al., 2019). In the light of this low activity and the uncertain activity with cytidylates, plastids may rely on the import of pyrimidine nucleotides. The low activity of PUMPKIN also makes it unlikely that it is significantly contributing to the global pyrimidine NTP pools *in vivo*.

## 4.5. Pyrimidine metabolism during germination

The pyrimidine metabolism of Arabidopsis during germination is organized differently than during the growth phase. This is due to a high amount of rNMPs, which are generated upon seed imbibition from breakdown of parental RNA (Li et al., 2006). Consequently, nucleotide pools during imbibition are preferentially fueled by the recycling of these rNMPs, rather than the UMP *de novo* synthesis. Expression of the gene encoding the first enzyme of UMP *de novo* synthesis, ATC, is low during early imbibition and induced upon transfer to growth conditions (Narsai et al., 2011). After transfer to growth conditions the *de novo* synthesis of UMP is becoming more and more important, as rNMPs obtained from recycling processes are consumed. It has been shown, that the pyrimidine nucleotide reserves of Arabidopsis seeds last for five days, before UMP *de novo* synthesis is mandatory to maintain growth (Chen and Slocum, 2008). Nevertheless, *de novo* synthesis is initiated earlier, to prevent growth arrest. Associated with UMP *de novo* synthesis is the correlation of the UTP pool with the downstream CTP and dCTP pools, which has been observed during the vegetative growth phase (**Figure 20**). This correlation is not as evident during seed imbibition (**Figure 26**), again suggesting that early metabolic activity of the seed is likely relying more on rNMPs obtained from the recycling of parental RNA than on UMP derived from *de novo* synthesis. A correlation between the three pools can be observed in the *umk3$_{\Delta E76}$* line 24 hours after the seeds were moved to the light, indicating that the production of pyrimidine NTPs is already relying on UMP *de novo* synthesis at this time point.

# 5.    Summary and outlook

In this work, the cytosolic UMK3 was identified as the central UMK of pyrimidine nucleotide metabolism in Arabidopsis. The phosphorylation of *de novo* synthesized UMP is performed by UMK3, and a reduction in UMK3 activity has a negative effect on vegetative plant growth. UMK3 is also essential during the reproductive phase of the plant, as egg cells or pollen carrying a *UMK3* null mutation are not viable. Lowering the activity of UMK3 influences the pyrimidine nucleotide content of Arabidopsis plants at all developmental stages. Additionally, clues were found for a regulatory mechanism for UMK3 by means of substrate inhibition. Mitochondrial UMK2 is evolutionarily conserved, but appears to be less important for pyrimidine nucleotide metabolism than UMK3. Plants carrying a homozygous UMK2 null-mutation show no developmental or metabolomic changes during the vegetative growth phase. However, during germination, UMK2 appears to play an increased role as dCMP kinase, and its activity is important for the production of precursors for mitochondrial DNA replication. No clear role could be determined for UMK1. The data suggest a function as a CMP kinase, but this was only observed when UMK3 activity was reduced. It has been theorized that UMK1 is only important when UMK3 activity is downregulated *in vivo*.

The characterization of the three AMK-like UMKs of Arabidopsis revealed distinct roles of the three enzymes in pyrimidine nucleotide metabolism, but some open questions remain. It was not possible to generate *UMK3* knockout lines with the CRISPR/Cas9 system, because the function of UMK3 is essential during the reproductive phase and for embryonic development. However, this does not mean that its function is also irreplaceable during other growth stages, where activity of UMK1 and UMK2 may be sufficient to keep the plant alive. It may therefore be interesting to explore options for a deactivation of UMK3 during later growth stages, like an inducible RNAi system or removal of UMK3 proteins by targeting them for degradation (Ludwicki et al., 2019). By switching off UMK3 and monitoring metabolomic or developmental changes, further insights into the organization of pyrimidine nucleotide metabolism at different growth stages could be obtained. Another topic worth investigating is the substrate inhibition of UMK3, which was observed *in vitro*. It is not clear whether this mechanism has any relevance *in vivo* or is related to a regulation of UMK3. To further explore this theory, it may be beneficial to analyze the metabolome of wild type and *umk3$_{\Delta E76}$* plants subjected to different growth limiting stress conditions. If clues towards an *in vivo* regulation are found, also the analysis of *UMK1* mutant lines under such conditions may be interesting to elucidate a potential role of UMK1 in the context of a downregulated UMK3. The results of this thesis strongly suggest, that mitochondria are able to exchange pyrimidine nucleotides with the cytosol in both directions. To facilitate this exchange *in vivo*, a dedicated pyrimidine transporter may

exist, that has not yet been described in the literature. The search for this transporter could form the basis for a completely new research project. In the context of mitochondrial pyrimidine metabolism, there are still open questions about the importance of *UMK2*. While the gene is universally conserved in vascular plants, a knockout seemed to have no negative influence on plant performance and only caused small metabolic alterations during germination. Also the reduction in mtDNA copy number appears to be negligible for the seedling under the employed growth conditions. Since the laboratory conditions resemble an ideal environment for the plant with a low amount of biotic and abiotic stress, harsher conditions could help identify a function of UMK2.

# 6.  Material and methods

## 6.1.  Cloning

Primers used in this thesis can be found in **Table A 2**. Constructs for protein expression, subcellular localization, targeting of *UMK* genes with the CRISPR/Cas9 system, an *in vitro* cleavage assay and *UMK3* complementation were created as part of this work. Additionally, constructs generated by Nieves Medina-Escobar were used for co-expression of an *mRuby* subcellular marker gene localizing to the cytosol (V238) or mitochondria (V241, unpublished).

The coding regions with introns of the *UMKs* were amplified with added restriction sites from genomic *Arabidopsis thaliana* Col-0 DNA. For *UMK1*, the primers P1379 and P1380 introducing *Eco*RI- and *Xma*I-sites were used. With the same added restriction sites, UMK2 was amplified with P1381 and P1382. P1383 and P1384 introducing *Cla*I- and *Xma*I-sites were used for amplification of *UMK3*. Likewise, *UMK3$_{ΔE76}$* was amplified with P1383 and P1384 from genomic *umk3$_{ΔE76}$* DNA. Fragments were cloned into pJet1.2 (K1231; Thermo Fisher) to generate the constructs H520 (*UMK1*), H521 (*UMK2*), H522 (*UMK3*) and H1326 (*UMK3$_{ΔE76}$*).

For biochemical characterization of the UMKs, constructs for transient expression of C-terminal hemagglutinin (HA) and Strep-tagged UMK variants in *Nicotiana benthamiana* were generated. For this, *UMK1* and *UMK2* were cloned using the *Eco*RI/*Xma*I sites from H520 and H521 into pXCScpmv-HAStrep (V69, Myrach et al., 2017) creating the constructs H551 and H552. Likewise, *UMK3* and *UMK3$_{ΔE76}$* were cloned into V69 using the *Cla*I/*Xma*I sites from H522 and H1326 creating the constructs H554 and H1327.

The subcellular localization of the UMKs was determined using constructs for transient expression of C-terminal mNeonGreen-tagged variants in *Nicotiana benthamiana* (Shaner et al., 2013). A synthetic *mNeonGreen* gene with an intron (sequence can be found in **Figure A 8**) was cloned into pJet 1.2 creating H176. It was then amplified with P1099 and P1100 adding *Xma*I and *Xba*I restriction sites and again cloned into pJet 1.2 to create construct H449. Via the *Xma*I/*Xba*I sites, *mNeonGreen* from H449 was moved into pXCS-YFP (V36, Dahncke and Witte, 2013) replacing the *YFP* and generating pXCS-mNeongreen (V165). The three *UMKs* were inserted into V165 via *Eco*RI/*Xma*I sites (*UMK1* from H520 and *UMK2* from H521) or *Cla*I/*Xma*I sites (*UMK3* from H522) generating constructs H543 (*UMK1*), H544 (*UMK2*) and H546 (*UMK3*).

The vectors for the *UMK2 in vitro* cleavage assay were constructed using a modified pEn-Chimera vector (H858, Fauser et al., 2014), where a *lacZ* operon was inserted behind the T7-promoter to facilitate

selection of correctly assembled constructs (Jana Streubel, unpublished). sgRNA candidates were ordered as complementary primers with overhangs to be inserted into H858. The primers were annealed and inserted by *Bbs*I cut/ligation. The following primer combinations were used: sgRNA1, P2451 + P2452; sgRNA2, P2453 + P2454; sgRNA3, P2455 + P2456; sgRNA4, P2457 + P2458; sgRNA5, P2459 + P2460; sgRNA6, P2461 + P2462. The finished constructs (pIVCUMK2sg1 – pIVCUMK2sg6) were used for *in vitro* transcription using T7 RNA polymerase.

The constructs used to generate the CRISPR/Cas9 mutant lines were cloned as described in Niehaus et al., 2022. The final constructs consisted of four expression cassettes, which were combined in the MoClo level 2 recipient vector pAGM4723 by *Bsa*I cut/ligation. The four expression cassettes were (1) a glufosinate resistance gene expressed by the *Nopaline Synthase* promoter (V183), (2) a *Cas9* gene expressed by the egg-cell specific *EC1.2* gene promoter (V182, Wang et al., 2015), (3) a *Green Fluorescent Protein (GFP)* gene expressed by the seed-specific *Brassica napus NAPIN-A* promoter (V181) and (4) a specific guideRNA (sgRNA) coding sequence under control of the *ATU6-26* promoter. The sgRNA cassettes were constructed by annealing complementary primers or in two steps according to Xie et al., 2015. The vectors encoding the sgRNAs targeting *UMK1* were generated by annealing the complementary primers P1831 and P1832 or P1833 and P1834 creating suitable overhangs to *Bbs*I-cut sgRNA shuttle vectors (Ordon et al., 2017) with MoClo compatible sites (Jana Streubel, unpublished). The resulting level_0 constructs H769 and H770 were used in a *Bsa*I cut/ligation together with the level_1 vector pICH47751 to create the constructs H771 and H772. These were used for assembly in pAGM4723 via *Bbs*I cut/ligation together with the other three expression cassettes generating the constructs H773 and H774. The level_0 vectors with sequences encoding sgRNAs targeting *UMK2* and *UMK3* were constructed by first amplifying the sgRNA coding sequences in two parts from the pGTR vector (Xie et al., 2015) followed by a second reaction with the two parts fused by *Bsa*I cut/ligation as template. The following primer combinations were used for the initial reactions: *UMK2* sgRNA1, P2471 + P272 and P2472 + P293; *UMK2* sgRNA2, P2473 + P272 and P2474 + P293; *UMK2* sgRNA3, P2475 + P272 and P2476 + P293; *UMK2* sgRNA4, P2477 + P272 and P2478 + P293; *UMK2* sgRNA5, P2479 + P272 and P2480 + P293; *UMK2* sgRNA6, P2481 + P272 and P2482 + P293; *UMK3* gRNA1, P1653 + P272 and P1654 + P293; *UMK3* sgRNA2, P1655 + P272 and P1656 + P293. For the second reaction, the non-gene-specific flanking primers P274 and P294 were used and the amplicons inserted into the sgRNA shuttle vector by *Bbs*I cut/ligation generating the constructs H1103, H1104, H1105, H1106, H1107, H1108, H976 and H977. Like before the sgRNA cassettes were then inserted into the level_1 vector pICH47751 via *Bsa*I cut/ligation generating the constructs H1124, H1125, H1126, H1127, H1128, H1129, H978 and H979. The final assembly in pAGM4723 was performed with the three other cassettes by *Bbs*I

cut/ligation generating the constructs H1130, H1131, H1132, H1133, H1134, H1135, H982 and H983 for plant transformation.

Also a construct for complementation of the $umk3_{\Delta E76}$ line was generated by amplifying a 1024 base pair fragment of the native *UMK3* promoter with P1874 and P1875 and cloning the fragment into pJet1.2 creating the construct H794. The promoter sequence was then inserted into construct H554 (see above) via the *Asc*I/*Eco*RI sites replacing the 35S promoter generating construct H1178.

## 6.2. Protein expression, purification and characterization

The *UMK* genes were transiently expressed in *Nicotiana benthamiana*, by infiltrating the plants with liquid cultures of *Agrobacterium tumefaciens* carrying the constructs H551 (*UMK1*), H552 (*UMK2*), H554 (*UMK3*) or H1327 (*UMK3$_{\Delta E76}$*). UMK proteins were purified by StrepTactin affinity chromatography as described in Werner et al., 2008. The purified proteins were quantified with Bovine Serum Albumin standards by sodium dodecyl sulfate (SDS) gel electrophoresis and Coomassie Brilliant Blue staining. Strep-tagged proteins were also detected by immunoblot with 3000-fold diluted monoclonal Anti-Strep Tag 7G8 antibody (bsbs300780, Antibody Facility Peine-Ost) and 3000-fold diluted Anti-Mouse IgG (A3562, Sigma). The purified proteins were used in a substrate screen by incubating them in a coupled assay containing 7500-fold diluted Pyruvate Kinase (600-1000 U mL$^{-1}$; PK), 7500-fold diluted Lactate Dehydrogenase (900-1400 U mL$^{-1}$; LDH; P0294, Sigma), 1mM NMP, 1 mM ATP, 40 mM Tris H$_2$SO$_4$ (pH 7.5), 20 mM KCl, 4 mM MgCl$_2$, 2 mM phosphoenolpyruvate and 0.17 mM nicotinamide adenine dinucleotide (NADH). Reactions were carried out in 96-well plates with a volume of 300 μL at 22°C and started by adding ATP. NADH consumption was monitored for 30 minutes in a photometer at a wavelength of 340 nm. Reactions were carried out in triplicates and specific activities were calculated based on the linear phase of the assay. Kinetic constants for CMP, UMP and dCMP were determined using concentrations of 0.025, 0.05, 0.1, 0.25, 0.5, 1, 2.5 and 10 mM with five replicates per concentration. The data were fitted in GraphPad Prism 4.01 using the Michaelis-Menten equation or, in case of substrate inhibition, equation 5.44 from Copeland, 2000: V=Vmax*S/(Km + S*(1+X/Ki)). The comparison of UMK3 and UMK3$_{\Delta E76}$ enzymatic activity was performed with 0.1 mM UMP or CMP and 0.5 mM dCMP with three replicates.

## 6.3. Subcellular localization

To determine the subcellular localization of the UMK-mNeonGreen fusion proteins, the respective gene fusions were transiently expressed in *Nicotiana benthamiana* by infiltrating liquid culture of Agrobacterium tumefaciens carrying the constructs H543 (*UMK1*), H544 (*UMK2*) or H546 (*UMK3*). Leaf discs of the plants were analyzed with a Leica SP8 confocal fluorescence microscope through a HC PL APO CS2 40 x 1.10 water immersion objective three days after infiltration. For UMK1 and UMK3, we co-infiltrated liquid cultures of *Agrobacterium tumefaciens* carrying the construct V238, which expresses a gene encoding a cytosolic variant of mRuby3 red fluorescence protein. Likewise, for UMK2 a mitochondrial mRuby3 fluorescence protein (V241) was chosen. Images were acquired using the sequential scan feature. For mNeonGreen detection an excitation wavelength of 488 nm and an emission wavelength window of 500-528 nm was used. The signal for mRuby3 was detected with an excitation wavelength of 552 nm and an emission window of 575-620 nm. Images were processed with the Leica Application Suite Advanced Fluorescence software. Pearson correlation coefficients and Van Steensel cross correlation functions were determined using the JACoP plugin in ImageJ (van Steensel et al., 1996; Bolte and Cordelières, 2006).

## 6.4. Plant material and cultivation

*Arabidopsis thaliana* and *Nicotiana benthamiana* plants were grown in a growth chamber equipped with Osram Fluora 36W/77 light tubes under long-day conditions (16 hours light of 85 µmol s$^{-1}$ m$^{-2}$, 22°C and 8 hours dark, 20°C, 60% relative humidity). Arabidopsis plants were grown from a uniform seed batch on soil (Steckmedium, Klasmann-Deilmann, Geeste, Germany) in randomized fashion for seven or 35 days for phenotype characterization and for 18 or 60 days for metabolome measurements. Before transfer to long-day growth conditions, the pots with the seeds were incubated for 48 hours at 4°C in the dark. Seedlings used in germination experiments were grown on filter paper soaked with modified half-strength Murashige and Skoog (MS) medium in Petri dishes as described in Niehaus et al., 2022. For each replicate, approximately 10 mg of seeds were spread on the soaked filter paper and imbibed for 48 h at 4°C in the dark. Afterwards, the Petri dishes were transferred to long-day growth conditions as described above. *Nicotiana benthamiana* plants were also grown on soil and three- to four-week-old plants were used for infiltration.

## 6.5. *In vitro* cleavage assay

For the *in vitro* cleavage assay, a PCR amplification with the primers P-1997 and P-1998 and the vectors pIVCUMK2sg1 – pIVCUMK2sg6 as template was performed. The resulting fragments were used in an *in vitro* transcription using T7 RNA polymerase (ThermoFisher, EP0111) according to the manufacturer's instructions. The template DNA was then removed from the reaction by adding 2 units of DNase (Promega, M199A) and the transcribed RNA was diluted to a final concentration of 330 ng µL$^{-1}$. In parallel, the pJet vector containing the UMK2 gene H522 was linearized by *Bsa*I restriction digest and diluted to 100 ng µL$^{-1}$. Before the assay, EnGen *Spy*Cas9 NLS (New England Biolabs, M0646T) was diluted 1:20 in NEBuffer 3.1 (New England Biolabs, B7203). For each sample, 1 µL of the diluted Cas9 was then mixed with 3 µL of the respective transcribed RNA in a total volume of 27 µL containing 1:10 diluted NEBuffer 3.1. After 10 minutes, 3 µL of linearized H522 was added as DNA target for the assembled sgRNA/Cas9 complexes. The reaction was incubated for 15 minutes at 37°C and then stopped by adding Proteinase K (AppliChem, A3830). Samples were then loaded on an agarose gel to visualize the results.

## 6.6. Fragment length analysis

To screen plants for potential editing events induced by the CRISPR/Cas9 system, genomic DNA was extracted from leaf material according to Edwards et al., 1991. Fragment length analysis via capillary gel electrophoresis was carried out as described in Rinne et al., 2021. Briefly, genomic DNA from a wild type or potentially edited plants was used in a PCR setup containing three primers (**Table 7**). A genomic fragment of 200 to 400 base pair spanning the sgRNA target site was amplified. Different dye-labelled primers were used for the reaction containing wild type or potentially edited genomic DNA, respectively. The chosen dyes were JOE (green peaks) for wild type fragments and 6-FAM (blue peaks) for potentially edited fragments (Ju et al., 1995). The following forward and reverse primers were used: *UMK1*, P1942 + P1943; *UMK2*, P2565 + P2566 for sgRNA4, 5 and 6; P2567 + P2568 for sgRNA2; P2569 + P2570 for sgRNA1 and 3; *UMK3*, P1946 + P1947. After amplification, 1 µL of wild type reaction was mixed with 1 µL of the reaction from a potentially edited plant and 0.25 µL Orange-500 DNA Size Standard (NimaGen, DSMO-100) in 10 µL Hi-Di Formamide (Applied Biosystems). The samples were then heated to 95°C in a thermoblock for 5 minutes and cooled down on ice. Analysis was performed with an ABI 310 capillary sequencer. Chromatograms were displayed using GeneMapper ID 3.2.

**Table 7: Composition of a PCR reaction for fragment length analysis with a total volume of 20 µL.**

| Component | Concentration |
| --- | --- |
| PCRBio Hifi Buffer 5x | 1x |
| forward primer with M13-tail | 50 nM |
| reverse primer | 250 nM |
| M13 dye-labelled primer | 125 nM |
| genomic DNA | 1-10 ng µL$^{-1}$ |
| PCRBio Hifi Polymerase | 0.01 U µL$^{-1}$ |

## 6.7. Leaf area quantification

The leaf area was quantified as explained in Niehaus et al., 2022. In short, pictures of 35-day-old plants were taken with a Samsung Galaxy S20 or S23 from the same distance. Image analysis was carried out using the image processing package Fiji (Schindelin et al., 2012) in ImageJ 1.53c. The following Color Thresholds for analysis were selected: Hue 46-100; Saturation: 0-255; Brightness: automatic. The leaf area was calculated for each individual plant from the detected green pixels. A leaf area of 1 cm$^2$ corresponded to 222x222 pixels. Images of seven-day-old seedlings were taken with a Nikon DS-RI2 binocular and analyzed the same way. A leaf area of 1 mm$^2$ corresponded to 341x341 pixels.

## 6.8. Structural protein modeling

The structures of UMK3 and UMK3$_{\Delta E76}$ were predicted using AlphaFold (Jumper et al., 2021) on the ColabFold v1.5.2 webserver (Mirdita et al., 2022) with standard settings. For both proteins the structure with the highest average predicted local distance difference test (pLDDT) score was chosen for comparison. Alignment and visualization of the two structures was done with PyMol v2.5.5.

## 6.9. Metabolite analysis

Metabolites were extracted from germinating seeds, 18-day-old or 60-day-old plants as described in Straube et al., 2021 with slight modifications (Straube et al., 2023). Starting material was either 10 mg of

germinating seeds, three whole 18-day-old plants without roots or leafs of 60-day-old plants equaling 50 to 100 mg. Plant material was immediately frozen in liquid nitrogen. Samples were ground with metal beads in a swing mill at 28 Hz for 2.5 min and quenched by adding 1 mL of 15% ice-cold trichloroacetic acid containing isotope standards of known concentrations. After a 10 minutes centrifugation step at 4°C with 40000$g$, the supernatant was transferred to a new tube containing 1 mL 60/40 dichloromethane/trioctylamine. Samples were vortexed and centrifuged at 4°C with 5000$g$. The supernatant was then transferred to a new tube containing 1 mL water with 0.0025% acetic acid. In parallel, a Strata X-AW column (Phenomenex) was sequentially equilibrated with 1 mL methanol, 1 mL 2/25/73 formic acid/methanol/water and 1 mL 10 mM ammonium acetate pH 4.5. Samples were loaded on and sucked through the column. After a washing step with 1 mL 1 mM ammonium acetate (pH 4.5) and 1 mL methanol, the metabolites were eluted with 1 mL 2/25/73 ammonia/methanol/water, dried in a speed vacuum centrifuge and re-suspended in 100 µL 5 mM ammonium acetate pH 9.5.

LC-MS analysis was performed with an Agilent HPLC 1290 system coupled to and Agilent 6470C series triple quadrupole mass spectrometer. Nucleotides and UDP-sugars were separated on a Hypercarb column (50 x 4.6 mm, particle size 5 µm; ThermoScientific). Nucleotides were measured in positive ion mode as described in Straube et al., 2021 and UDP-sugars in negative ion mode (Rautengarten et al., 2019). Precursor and product ions can be found in these two publications. For both methods, the flow rate was set to 0.6 mL min$^{-1}$ and the temperature of the column compartment to 25°C. Mobile phase A was 5 mM ammonium acetate pH 9.5 and mobile phase B acetonitrile. The run started with 100% mobile phase A and the amount of mobile phase B was gradually increased to 30% over 18 minutes. Subsequently, the column was washed with 100% mobile phase B for 4 min and equilibrated again with 100% mobile phase A for 8 min to prepare the next run. The source parameters were: gas temperature 250°C; gas flow 12 L min$^{-1}$; nebulizer 20 psi; sheath gas temperature 395°C; sheath gas flow 12 L min$^{-1}$; capillary voltage 3000 V. Data analysis was done with Agilent MassHunter Quantitative Analysis Version B.09.00 Build 9.0.647.0.

## 6.10. Quantitative PCR

Nucleic acid extraction for quantitative PCR (qPCR) was performed using a cetyltrimethylammonium bromide (CTAB)-based approach from dry or germinating seeds. Quantification of nuclear and mitochondrial genome copy number was performed via qPCR as described in Pedroza-García et al., 2019 and Niehaus et al., 2022. The seeds were incubated as described earlier. The qPCRs were carried out with a QuantStudio3 (Thermo Fisher) qPCR cycler and qPCRBIO SyGreen Mix (PCR Biosystems) according to the

manufacturer's instructions. Nuclear genome copy number was quantified by PCR with the primers P1577 and P1578 amplifying a fragment from the *UBC21* gene (At5g25760) as reference. For mitochondrial genome copy number *COX1* (AtMg01360) was amplified with P1581 and P1582. Each data point represents the mean of three technical replicates and three biological replicates were made per time point. For all measurements, a fixed threshold Ct value of 0.3 was used and the genome copy numbers were compared by applying the $2^{-\Delta Ct}$ method (Livak and Schmittgen, 2001).

## 6.11. Proteome analysis

Shotgun proteomics was performed by Nils Rugen of the Plant Proteomics Department, LUH as described in Klusch et al., 2023. Proteins were prepared for mass spectrometry analysis via the single-pot-solid-phase-enhanced sample preparation (SP3) protocol from Hughes et al., 2019 with slight modifications (Mikulášek et al., 2021). 20 mg of frozen seeds were ground with metal beads in a precooled swing mill at 28 Hz for 2.5 minutes, reconstituted in 500 µl of 1x SDT buffer (4% [w/v] SDS, 0.1 M DTT, 0.1% Tris-HCl, pH 7.6) and incubated on a thermoshaker for 1 hour at 60°C and 1000 rpm. After centrifugation for 10 minutes at 20000g, the supernatant was transferred into a new reaction tube and sonicated in a water bath for 10 minutes. Proteins were alkylated by incubation in 20 mM iodacetamid for 30 minutes at 600 rpm at room temperature in the dark. Alkylation was stopped by addition of 5 mM dithiothreitol. Carboxylate-modified hydrophilic beads (Sera-Mag, GE Life Sciences, 24152105050250) were combined 1:1 with corresponding hydrophobic beads (GE Life Sciences, 44152105050250) adding 600 µg beads to each sample. Proteins were precipitated by addition of 70 µL ethanol and subsequent incubation for 10 minutes at 1000 rpm and 24°C on a thermoshaker. Beads were pelleted on a magnetic rack for 2 minutes and proteins were washed three times with 140 µL 80% ethanol. After protein clean-up, beads were transferred in 80% ethanol into low protein-binding tubes (Low Binding Micro Tubes, Sarstedt) and ethanol removed. Proteins were digested with 2 µg of sequencing grade modified Trypsin (Promega, V5111) in 50 mM ammonium bicarbonate at 37°C and 1000 rpm overnight in a total reaction volume of 60 µL. On the next day, Trypsin activity was stopped by addition of 1% (v/v) formic acid (FA). The pH of each sample was controlled and adjusted to < 3 by further addition of 1% (v/v) FA if necessary. Tryptic peptides were further cleaned via solid-phase extraction on SepPak Vac 1cc (50 mg) tC18 cartridges (Waters). Cartridges were wetted with 1 mL 100% acetonitrile and 1 mL 0.1% (v/v) FA in 50% (v/v) acetonitrile. Cartridge equilibration was performed by adding 2 x 1 mL of 0.1% FA (v/v) in water. Acidified peptides (pH < 3) were loaded onto the cartridges and washed two times with 0.1% FA (v/v) in water and eluted twice in 200 µL of 0.1% FA (v/v)

in 50% (v/v) acetonitrile. Cleaned peptides were dried in a vacuum centrifuge and stored at -20°C. The final peptide concentration was determined with the PierceTM peptide quantification kit (Thermo Scientific) following the manufacturer's instructions.

A nanoElute HPLC (Bruker Daltonics) was coupled to a timsTOF Pro ion-mobility spectrometry quadrupole time of flight mass spectrometer (Bruker Daltonics). Peptides were reconstituted in 0.1% FA and 200 ng peptides per sample were directly transferred onto an "Aurora" reversed phase analytical column with integrated emitter tip (25cm x 75 µm inner diameter, IonOpticks). Peptides were separated on the analytical column at 50°C via a 70 minutes gradient (solvent A: 0.1% FA; solvent B: 0.1% FA in 100% acetonitrile) at a flow rate of 300 µL min$^{-1}$. A linear gradient from 2-37% B for the first 60 minutes was followed by a 10 minutes washing step at 95% B.

The timsTOF Pro mass spectrometer was operated in DDA PASEF mode and the pre-installed method "DDA PASEF-standard_1.1sec_cycletime" was used. Automatic recalibration of ion mobility before each sample run was activated. MS and MS/MS scan range was 100-1700 m/z, the ion mobility range (1/K0) was 0.6 – 1.6 V*s/cm$^2$. A polygon filtering was applied in the m/z and ion mobility area to exclude the low m/z of singly charged ions for PASEF precursor selection. Ramp and accumulation time was set to 100 ms to achieve a duty cycle close to 100%. The number of PASEF ramps was set to 10 with a charge maximum of 5. The quadrupole isolation width was set to 2 for m/z = 700 and 3 for m/z = 800. The collision energy was 20 eV for ion mobility (1/K0) 0.6 V*s/cm$^2$ and 59 eV for ion mobility (1/K0) 1.6 V*s/cm$^2$, respectively.

MaxQuant 2.2.0.0 (Tyanova et al., 2016a) was used to query acquired MS/MS spectra against a modified TAIR10 database including models of mitochondrial and plastid genes after RNA editing to improve sequence coverage of affected proteins (Fuchs et al., 2020). MaxQuant is pre-equipped with a database of common contaminants. Default parameters were used with the following exception: calculation of iBAQ values (Schwanhäusser et al., 2011) was activated, the options "log fit" and "charge normalization" were enabled. Identification transfer between individual runs via the "match between runs" feature was applied with the default parameters. Proteomic data were further analyzed with Perseus version 1.6.15.0 (Tyanova et al., 2016b). Protein groups labeled as "only identified by site", "potential contaminant" as well as "reverse" were initially filtered out to remove contaminants and false-positive hits from the results. iBAQ values were used as quantitative values for proteome comparison. Only protein groups with an iBAQ value >0 in at least 70% of all samples where considered for further analysis. Missing values of the remaining protein groups were replaced by random numbers from a normal distribution. Protein abundance within each sample was normalized via Z-scores. Principal component analysis as well as volcano plot were generated with default parameters.

# 7. References

**Abel, S., and Glund, K.** (1987). Ribonuclease in plant vacuoles: purification and molecular properties of the enzyme from cultured tomato cells. Planta **172** (1): 71–78.

**Arimura, S., Yamamoto, J., Aida, G.P., Nakazono, M., and Tsutsumi, N.** (2004). Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. Proceedings of the National Academy of Sciences of the United States of America **101** (20): 7805–7808.

**Bellin, L., Del Caño-Ochoa, F., Velázquez-Campoy, A., Möhlmann, T., and Ramón-Maiques, S.** (2021a). Mechanisms of feedback inhibition and sequential firing of active sites in plant aspartate transcarbamoylase. Nature communications **12** (1): 947.

**Bellin, L., Scherer, V., Dörfer, E., Lau, A., Vicente, A.M., Meurer, J., Hickl, D., and Möhlmann, T.** (2021b). Cytosolic CTP Production Limits the Establishment of Photosynthesis in Arabidopsis. Frontiers in plant science **12**: 789189.

**Bernard, C., Traub, M., Kunz, H.-H., Hach, S., Trentmann, O., and Möhlmann, T.** (2011). Equilibrative nucleoside transporter 1 (ENT1) is critical for pollen germination and vegetative growth in Arabidopsis. Journal of experimental botany **62** (13): 4627–4637.

**Bieniawska, Z., Paul Barratt, D.H., Garlick, A.P., Thole, V., Kruger, N.J., Martin, C., Zrenner, R., and Smith, A.M.** (2007). Analysis of the sucrose synthase gene family in Arabidopsis. The Plant Journal **49** (5): 810–828.

**Bolte, S., and Cordelières, F.P.** (2006). A guided tour into subcellular colocalization analysis in light microscopy. Journal of microscopy **224** (Pt 3): 213–232.

**Bölter, B., Sharma, R., and Soll, J.** (2007). Localisation of Arabidopsis NDPK2--revisited. Planta **226** (4): 1059–1065.

**Buck, S.** de, **Podevin, N., Nolf, J., Jacobs, A., and Depicker, A.** (2009). The T-DNA integration pattern in Arabidopsis transformants is highly determined by the transformed target cell. The Plant journal for cell and molecular biology **60** (1): 134–145.

**Carbonell, J., Felíu, J.E., Marco, R., and Sols, A.** (1973). Pyruvate kinase. Classes of regulatory isoenzymes in mammalian tissues. European journal of biochemistry **37** (1): 148–156.

**Chateigner-Boutin, A.-L., Des Francs-Small, C.C., Delannoy, E., Kahlau, S., Tanz, S.K., Longevialle, A.F.** de, **Fujii, S., and Small, I.** (2011). OTP70 is a pentatricopeptide repeat protein of the E subgroup involved in splicing of the plastid transcript rpoC1. The Plant journal for cell and molecular biology **65** (4): 532–542.

References

**Chen, C.T., and Slocum, R.D.** (2008). Expression and functional analysis of aspartate transcarbamoylase and role of de novo pyrimidine synthesis in regulation of growth and development in Arabidopsis. Plant physiology and biochemistry PPB **46** (2): 150–159.

**Chen, F., Dong, G., Ma, X., Wang, F., Zhang, Y., Xiong, E., Wu, J., Wang, H., Qian, Q., Wu, L., and Yu, Y.** (2018). UMP kinase activity is involved in proper chloroplast development in rice. Photosynthesis research **137** (1): 53–67.

**Chen, X., Kim, S.-H., Rhee, S., and Witte, C.-P.** (2023). A plastid nucleoside kinase is involved in inosine salvage and control of purine nucleotide biosynthesis. The Plant cell **35** (1): 510–528.

**Chiba, Y., and Green, P.J.** (2009). mRNA Degradation Machinery in Plants. J. Plant Biol. **52** (2): 114–124.

**Clausen, A.R., Girandon, L., Ali, A., Knecht, W., Rozpedowska, E., Sandrini, M.P.B., Andreasson, E., Munch-Petersen, B., and Piškur, J.** (2012). Two thymidine kinases and one multisubstrate deoxyribonucleoside kinase salvage DNA precursors in Arabidopsis thaliana. The FEBS journal **279** (20): 3889–3897.

**Clausen, A.R., Mutahir, Z., Munch-Petersen, B., and Piškur, J.** (2014). Plants salvage deoxyribonucleosides in mitochondria. Nucleosides, nucleotides & nucleic acids **33** (4-6): 291–295.

**Copeland, R.A.** (2000). Enzymes: A practical introduction to structure, me chanism, and data analysis, 2nd edition ([S.l.]: WILEY-BLACKWELL).

**Córdoba-Cañero, D., Morales-Ruiz, T., Roldán-Arjona, T., and Ariza, R.R.** (2009). Single-nucleotide and long-patch base excision repair of DNA damage in plants. The Plant journal for cell and molecular biology **60** (4): 716–728.

**Cornelius, S., Witz, S., Rolletschek, H., and Möhlmann, T.** (2011). Pyrimidine degradation influences germination seedling growth and production of Arabidopsis seeds. Journal of experimental botany **62** (15): 5623–5632.

**Da Fonseca-Pereira, P., Neri-Silva, R., Cavalcanti, J.H.F., Brito, D.S., Weber, A.P.M., Araújo, W.L., and Nunes-Nesi, A.** (2018). Data-Mining Bioinformatics: Connecting Adenylate Transport and Metabolic Responses to Stress. Trends in plant science **23** (11): 961–974.

**Dahncke, K., and Witte, C.-P.** (2013). Plant purine nucleoside catabolism employs a guanosine deaminase required for the generation of xanthosine in Arabidopsis. The Plant cell **25** (10): 4101–4109.

**Dancer, J., Neuhaus, H.E., and Stitt, M.** (1990). Subcellular compartmentation of uridine nucleotides and nucleoside-5' -diphosphate kinase in leaves. Plant physiology **92** (3): 637–641.

**Daumann, M., Hickl, D., Zimmer, D., DeTar, R.A., Kunz, H.-H., and Möhlmann, T.** (2018). Characterization of filament-forming CTP synthases from Arabidopsis thaliana. The Plant journal for cell and molecular biology **96** (2): 316–328.

References

**Diaz, C., Lemaître, T., Christ, A., Azzopardi, M., Kato, Y., Sato, F., Morot-Gaudry, J.-F., Le Dily, F., and Masclaux-Daubresse, C.** (2008). Nitrogen recycling and remobilization are differentially controlled by leaf senescence and development stage in Arabidopsis under low nitrogen nutrition. Plant physiology **147** (3): 1437–1449.

**Diehl, F.F., Miettinen, T.P., Elbashir, R., Nabel, C.S., Darnell, A.M., Do, B.T., Manalis, S.R., Lewis, C.A., and Vander Heiden, M.G.** (2022). Nucleotide imbalance decouples cell growth from cell proliferation. Nat Cell Biol **24** (8): 1252–1264.

**Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E.** (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nature biotechnology **32** (12): 1262–1267.

**Doremus, H.D., and Jagendorf, A.T.** (1985). Subcellular localization of the pathway of de novo pyrimidine nucleotide biosynthesis in pea leaves. Plant physiology **79** (3): 856–861.

**Dorion, S., Matton, D.P., and Rivoal, J.** (2006). Characterization of a cytosolic nucleoside diphosphate kinase associated with cell division and growth in potato. Planta **224** (1): 108–124.

**Dorion, S., and Rivoal, J.** (2015). Clues to the functions of plant NDPK isoforms. Naunyn-Schmiedeberg's archives of pharmacology **388** (2): 119–132.

**Edwards, K., Johnstone, C., and Thompson, C.** (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic acids research **19** (6): 1349.

**Evrin, C., Straut, M., Slavova-Azmanova, N., Bucurenci, N., Onu, A., Assairi, L., Ionescu, M., Palibroda, N., Bârzu, O., and Gilles, A.-M.** (2007). Regulatory mechanisms differ in UMP kinases from gram-negative and gram-positive bacteria. The Journal of biological chemistry **282** (10): 7242–7253.

**Fauser, F., Schiml, S., and Puchta, H.** (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. The Plant Journal **79** (2): 348–359.

**Feller, U., and Fischer, A.** (1994). Nitrogen Metabolism in Senescing Leaves. Critical Reviews in Plant Sciences **13** (3): 241–273.

**Floyd, B.E., Morriss, S.C., MacIntosh, G.C., and Bassham, D.C.** (2015). Evidence for autophagy-dependent pathways of rRNA turnover in Arabidopsis. Autophagy **11** (12): 2199–2212.

**Floyd, S., Favre, C., Lasorsa, F.M., Leahy, M., Trigiante, G., Stroebel, P., Marx, A., Loughran, G., O'Callaghan, K., Marobbio, C.M.T., Slotboom, D.J., Kunji, E.R.S., Palmieri, F., and O'Connor, R.** (2007). The insulin-like growth factor-I-mTOR signaling pathway induces the mitochondrial pyrimidine nucleotide carrier to promote cell growth. Molecular biology of the cell **18** (9): 3545–3555.

**Fuchs, P., Rugen, N., Carrie, C., Elsässer, M., Finkemeier, I., Giese, J., Hildebrandt, T.M., Kühn, K., Maurino, V.G., Ruberti, C., Schallenberg-Rüdinger, M., Steinbeck, J., Braun, H.-P., Eubel, H., Meyer,

References

**E.H., Müller-Schüssele, S.J., and Schwarzländer, M.** (2020). Single organelle function and organization as estimated from Arabidopsis mitochondrial proteomics. The Plant journal for cell and molecular biology **101** (2): 420–441.

**Geserick, C., and Tenhaken, R.** (2013). UDP-sugar pyrophosphorylase is essential for arabinose and xylose recycling, and is required during vegetative and reproductive growth in Arabidopsis. The Plant journal for cell and molecular biology **74** (2): 239–247.

**Girke, C., Daumann, M., Niopek-Witz, S., and Möhlmann, T.** (2014). Nucleobase and nucleoside transport and integration into plant metabolism. Frontiers in plant science **5**: 443.

**Gorelova, V., Lepeleire, J.** de, **van Daele, J., Pluim, D., Meï, C., Cuypers, A., Leroux, O., Rébeillé, F., Schellens, J.H.M., Blancquaert, D., Stove, C.P., and van der Straeten, D.** (2017). Dihydrofolate Reductase/Thymidylate Synthase Fine-Tunes the Folate Status and Controls Redox Homeostasis in Plants. The Plant cell **29** (11): 2831–2853.

**Hauck, O.K., Scharnberg, J., Escobar, N.M., Wanner, G., Giavalisco, P., and Witte, C.-P.** (2014). Uric acid accumulation in an Arabidopsis urate oxidase mutant impairs seedling establishment by blocking peroxisome maintenance. The Plant cell **26** (7): 3090–3100.

**Hickl, D., Scheuring, D., and Möhlmann, T.** (2021). CTP Synthase 2 From Arabidopsis thaliana Is Required for Complete Embryo Development. Frontiers in plant science **12**: 652434.

**Hughes, C.S., Moggridge, S., Müller, T., Sorensen, P.H., Morin, G.B., and Krijgsveld, J.** (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. Nature protocols **14** (1): 68–85.

**Johansson, K., Ramaswamy, S., Ljungcrantz, C., Knecht, W., Piskur, J., Munch-Petersen, B., Eriksson, S., and Eklund, H.** (2001). Structural basis for substrate specificities of cellular deoxyribonucleoside kinases. Nature structural biology **8** (7): 616–620.

**Johansson, M., Hammargren, J., Uppsäll, E., MacKenzie, A., and Knorpp, C.** (2008). The activities of nucleoside diphosphate kinase and adenylate kinase are influenced by their interaction. Plant Science **174** (2): 192–199.

**Ju, J., Ruan, C., Fuller, C.W., Glazer, A.N., and Mathies, R.A.** (1995). Fluorescence energy transfer dye-labeled primers for DNA sequencing and analysis. Proceedings of the National Academy of Sciences of the United States of America **92** (10): 4347–4351.

**Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W.,**

References

**Kavukcuoglu, K., Kohli, P., and Hassabis, D.** (2021). Highly accurate protein structure prediction with AlphaFold. Nature **596** (7873): 583–589.

**Kihara, A., Saburi, W., Wakuta, S., Kim, M.-H., Hamada, S., Ito, H., Imai, R., and Matsui, H.** (2011). Physiological and biochemical characterization of three nucleoside diphosphate kinase isozymes from rice (Oryza sativa L.). Bioscience, biotechnology, and biochemistry **75** (9): 1740–1745.

**Kim, S.-K., Lormand, J.D., Weiss, C.A., Eger, K.A., Turdiev, H., Turdiev, A., Winkler, W.C., Sondermann, H., and Lee, V.T.** (2019). A dedicated diribonucleotidase resolves a key bottleneck for the terminal step of RNA degradation. eLife **8**.

**Kleczkowski, L.A., Kunz, S., and Wilczynska, M.** (2010). Mechanisms of UDP-Glucose Synthesis in Plants. Critical Reviews in Plant Sciences **29** (4): 191–203.

**Klepikova, A.V., Kasianov, A.S., Gerasimov, E.S., Logacheva, M.D., and Penin, A.A.** (2016). A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. The Plant journal for cell and molecular biology **88** (6): 1058–1070.

**Klusch, N., Dreimann, M., Senkler, J., Rugen, N., Kühlbrandt, W., and Braun, H.-P.** (2023). Cryo-EM structure of the respiratory I + III2 supercomplex from Arabidopsis thaliana at 2 Å resolution. Nature plants **9** (1): 142–156.

**Knorpp, C., Johansson, M., and Baird, A.-M.** (2003). Plant mitochondrial nucleoside diphosphate kinase is attached to the membrane through interaction with the adenine nucleotide translocator. FEBS letters **555** (2): 363–366.

**Koffler, B.E., Bloem, E., Zellnig, G., and Zechmann, B.** (2013). High resolution imaging of subcellular glutathione concentrations by quantitative immunoelectron microscopy in different leaf areas of Arabidopsis. Micron (Oxford, England 1993) **45**: 119–128.

**Kotake, T., Hojo, S., Yamaguchi, D., Aohara, T., Konishi, T., and Tsumuraya, Y.** (2007). Properties and physiological functions of UDP-sugar pyrophosphorylase in Arabidopsis. Bioscience, biotechnology, and biochemistry **71** (3): 761–771.

**Kotake, T., Takata, R., Verma, R., Takaba, M., Yamaguchi, D., Orita, T., Kaneko, S., Matsuoka, K., Koyama, T., Reiter, W.-D., and Tsumuraya, Y.** (2009). Bifunctional cytosolic UDP-glucose 4-epimerases catalyse the interconversion between UDP-D-xylose and UDP-L-arabinose in plants. The Biochemical journal **424** (2): 169–177.

**Kumar, V., Spangenberg, O., and Konrad, M.** (2000). Cloning of the guanylate kinase homologues AGK-1 and AGK-2 from Arabidopsis thaliana and characterization of AGK-1. European journal of biochemistry **267** (2): 606–615.

References

**Lange, P.R., Geserick, C., Tischendorf, G., and Zrenner, R.** (2008). Functions of chloroplastic adenylate kinases in Arabidopsis. Plant physiology **146** (2): 492–504.

**Lazar, G., Zhang, H., and Goodman, H.M.** (1993). The origin of the bifunctional dihydrofolate reductasethymidylate synthase isogenes of Arabidopsis thaliana. The Plant Journal **3** (5): 657–668.

**Li, B., Wu, H., and Guo, H.** (2018). Plant mRNA decay: extended roles and potential determinants. Current opinion in plant biology **45** (Pt A): 178–184.

**Li, Q., Feng, J.-X., Han, P., and Zhu, Y.-X.** (2006). Parental RNA is Significantly Degraded During Arabidopsis Seed Germination. J Integrative Plant Biology **48** (1): 114–120.

**Lincker, F., Philipps, G., and Chabouté, M.-E.** (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 research **32** (4): 1430–1438.

**Livak, K.J., and Schmittgen, T.D.** (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif.) **25** (4): 402–408.

**Ludwicki, M.B., Li, J., Stephens, E.A., Roberts, R.W., Koide, S., Hammond, P.T., and DeLisa, M.P.** (2019). Broad-Spectrum Proteome Editing with an Engineered Bacterial Ubiquitin Ligase Mimic. ACS central science **5** (5): 852–866.

**Mainguet, S.E., Gakière, B., Majira, A., Pelletier, S., Bringel, F., Guérard, F., Caboche, M., Berthomé, R., and Renou, J.P.** (2009). Uracil salvage is necessary for early Arabidopsis development. The Plant journal for cell and molecular biology **60** (2): 280–291.

**Marobbio, C.M.T., Di Noia, M.A., and Palmieri, F.** (2006). Identification of a mitochondrial transporter for pyrimidine nucleotides in Saccharomyces cerevisiae: bacterial expression, reconstitution and functional characterization. The Biochemical journal **393** (Pt 2): 441–446.

**Meinke, D.W., and Sussex, I.M.** (1979). Embryo-lethal mutants of Arabidopsis thaliana. A model system for genetic analysis of plant embryo development. Developmental biology **72** (1): 50–61.

**Meng, M., Wilczynska, M., and Kleczkowski, L.A.** (2008). Molecular and kinetic characterization of two UDP-glucose pyrophosphorylases, products of distinct genes, from Arabidopsis. Biochimica et biophysica acta **1784** (6): 967–972.

**Mergner, J., Frejno, M., List, M., Papacek, M., Chen, X., Chaudhary, A., Samaras, P., Richter, S., Shikata, H., Messerer, M., Lang, D., Altmann, S., Cyprys, P., Zolg, D.P., Mathieson, T., Bantscheff, M., Hazarika, R.R., Schmidt, T., Dawid, C., Dunkel, A., Hofmann, T., Sprunck, S., Falter-Braun, P., Johannes, F., Mayer, K.F.X., Jürgens, G., Wilhelm, M., Baumbach, J., Grill, E., Schneitz, K., Schwechheimer, C., and Kuster, B.** (2020). Mass-spectrometry-based draft of the Arabidopsis proteome. Nature **579** (7799): 409–414.

References

**Mikulášek, K., Konečná, H., Potěšil, D., Holánková, R., Havliš, J., and Zdráhal, Z.** (2021). SP3 Protocol for Proteomic Plant Sample Preparation Prior LC-MS/MS. Frontiers in plant science **12**: 635550.

**Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M.** (2022). ColabFold: making protein folding accessible to all. Nature methods **19** (6): 679–682.

**Müller-Dieckmann, H.J., and Schulz, G.E.** (1995). Substrate specificity and assembly of the catalytic center derived from two structures of ligated uridylate kinase. Journal of molecular biology **246** (4): 522–530.

**Myrach, T., Zhu, A., and Witte, C.-P.** (2017). 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. The Journal of biological chemistry **292** (35): 14556–14565.

**Nakamura, Y.** (2017). Plant Phospholipid Diversity: Emerging Functions in Metabolism and Protein-Lipid Interactions. Trends in plant science **22** (12): 1027–1040.

**Narsai, R., Law, S.R., Carrie, C., Xu, L., and 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 physiology **157** (3): 1342–1362.

**Niehaus, M., Straube, H., Specht, A., Baccolini, C., Witte, C.-P., and Herde, M.** (2022). The nucleotide metabolome of germinating Arabidopsis thaliana seeds reveals a central role for thymidine phosphorylation in chloroplast development. The Plant cell **34** (10): 3790–3813.

**Nomura, Y., Izumi, A., Fukunaga, Y., Kusumi, K., Iba, K., Watanabe, S., Nakahira, Y., Weber, A.P.M., Nozawa, A., and Tozawa, Y.** (2014). Diversity in guanosine 3',5'-bisdiphosphate (ppGpp) sensitivity among guanylate kinases of bacteria and plants. The Journal of biological chemistry **289** (22): 15631–15641.

**Noree, C., Monfort, E., Shiau, A.K., and Wilhelm, J.E.** (2014). Common regulatory control of CTP synthase enzyme activity and filament formation. Molecular biology of the cell **25** (15): 2282–2290.

**Ohler, L., Niopek-Witz, S., Mainguet, S.E., and Möhlmann, T.** (2019). Pyrimidine Salvage: Physiological Functions and Interaction with Chloroplast Biogenesis. Plant physiology **180** (4): 1816–1828.

**Oikawa, A., Matsuda, F., Kikuyama, M., Mimura, T., and Saito, K.** (2011). Metabolomics of a single vacuole reveals metabolic dynamism in an alga Chara australis. Plant physiology **157** (2): 544–551.

**Ordon, J., Gantner, J., Kemna, J., Schwalgun, L., Reschke, M., Streubel, J., Boch, J., and Stuttmann, J.** (2017). Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. The Plant Journal **89** (1): 155–168.

**Palmieri, L., Santoro, A., Carrari, F., Blanco, E., Nunes-Nesi, A., Arrigoni, R., Genchi, F., Fernie, A.R., and Palmieri, F.** (2008). Identification and characterization of ADNT1, a novel mitochondrial adenine nucleotide transporter from Arabidopsis. Plant physiology **148** (4): 1797–1808.

References

**Park, J.-I., Ishimizu, T., Suwabe, K., Sudo, K., Masuko, H., Hakozaki, H., Nou, I.-S., Suzuki, G., and Watanabe, M.** (2010). UDP-glucose pyrophosphorylase is rate limiting in vegetative and reproductive phases in Arabidopsis thaliana. Plant & cell physiology **51** (6): 981–996.

**Parks, R.E., and Aganwal, R.P.** (1973). 9 Nucleoside Diphosphokinases. In Group Transfer Part A: Nucleotidyl Transfer Nucleosidyl Transfer Acyl Transfer Phosphoryl Transfer (Elsevier), pp. 307–333.

**Pasti, C., Gallois-Montbrun, S., Munier-Lehmann, H., Veron, M., Gilles, A.-M., and Deville-Bonne, D.** (2003). Reaction of human UMP-CMP kinase with natural and analog substrates. European journal of biochemistry **270** (8): 1784–1790.

**Paszkiewicz, G., Gualberto, J.M., Benamar, A., Macherel, D., and Logan, D.C.** (2017). Arabidopsis Seed Mitochondria Are Bioenergetically Active Immediately upon Imbibition and Specialize via Biogenesis in Preparation for Autotrophic Growth. The Plant cell **29** (1): 109–128.

**Pedroza-García, J.-A., Nájera-Martínez, M., Mazubert, C., Aguilera-Alvarado, P., Drouin-Wahbi, J., Sánchez-Nieto, S., Gualberto, J.M., Raynaud, C., and Plasencia, J.** (2019). Role of pyrimidine salvage pathway in the maintenance of organellar and nuclear genome integrity. The Plant Journal **97** (3): 430–446.

**Picault, N., Hodges, M., Palmieri, L., and Palmieri, F.** (2004). The growing family of mitochondrial carriers in Arabidopsis. Trends in plant science **9** (3): 138–146.

**Preuten, T., Cincu, E., Fuchs, J., Zoschke, R., Liere, K., and Börner, T.** (2010). Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. The Plant journal for cell and molecular biology **64** (6): 948–959.

**Rautengarten, C., Heazlewood, J.L., and Ebert, B.** (2019). Profiling Cell Wall Monosaccharides and Nucleotide-Sugars from Plants. Current protocols in plant biology **4** (2): e20092.

**Reumann, S., Quan, S., Aung, K., Yang, P., Manandhar-Shrestha, K., Holbrook, D., Linka, N., Switzenberg, R., Wilkerson, C.G., Weber, A.P.M., Olsen, L.J., and Hu, J.** (2009). In-depth proteome analysis of Arabidopsis leaf peroxisomes combined with in vivo subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. Plant physiology **150** (1): 125–143.

**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.

**Ronceret, A., Gadea-Vacas, J., Guilleminot, J., Lincker, F., Delorme, V., Lahmy, S., Pelletier, G., Chabouté, M.-E., and Devic, M.** (2008). The first zygotic division in Arabidopsis requires de novo transcription of thymidylate kinase. The Plant journal for cell and molecular biology **53** (5): 776–789.

References

**Rösti, J., Barton, C.J., Albrecht, S., Dupree, P., Pauly, M., Findlay, K., Roberts, K., and Seifert, G.J.** (2007). UDP-glucose 4-epimerase isoforms UGE2 and UGE4 cooperate in providing UDP-galactose for cell wall biosynthesis and growth of Arabidopsis thaliana. The Plant cell **19** (5): 1565–1579.

**Sakamoto, W., and Takami, T.** (2014). Nucleases in higher plants and their possible involvement in DNA degradation during leaf senescence. Journal of experimental botany **65** (14): 3835–3843.

**Scheffzek, K., Kliche, W., Wiesmüller, L., and Reinstein, J.** (1996). Crystal structure of the complex of UMP/CMP kinase from Dictyostelium discoideum and the bisubstrate inhibitor P1-(5'-adenosyl) P5-(5'-uridyl) pentaphosphate (UP5A) and Mg2+ at 2.2 A: implications for water-mediated specificity. Biochemistry **35** (30): 9716–9727.

**Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A.** (2012). Fiji: an open-source platform for biological-image analysis. Nature methods **9** (7): 676–682.

**Schlichting, I., and Reinstein, J.** (1997). Structures of active conformations of UMP kinase from Dictyostelium discoideum suggest phosphoryl transfer is associative. Biochemistry **36** (31): 9290–9296.

**Schmid, L.-M., Ohler, L., Möhlmann, T., Brachmann, A., Muiño, J.M., Leister, D., Meurer, J., and Manavski, N.** (2019). PUMPKIN, the Sole Plastid UMP Kinase, Associates with Group II Introns and Alters Their Metabolism. Plant physiology **179** (1): 248–264.

**Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M.** (2011). Global quantification of mammalian gene expression control. Nature **473** (7347): 337–342.

**Serina, L., Blondin, C., Krin, E., Sismeiro, O., Danchin, A., Sakamoto, H., Gilles, A.M., and Bârzu, O.** (1995). Escherichia coli UMP-kinase, a member of the aspartokinase family, is a hexamer regulated by guanine nucleotides and UTP. Biochemistry **34** (15): 5066–5074.

**Shaner, N.C., Lambert, G.G., Chammas, A., Ni, Y., Cranfill, P.J., Baird, M.A., Sell, B.R., Allen, J.R., Day, R.N., Israelsson, M., Davidson, M.W., and Wang, J.** (2013). A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. Nature methods **10** (5): 407–409.

**Soltabayeva, A., Srivastava, S., Kurmanbayeva, A., Bekturova, A., Fluhr, R., and Sagi, M.** (2018). Early Senescence in Older Leaves of Low Nitrate-Grown Atxdh1 Uncovers a Role for Purine Catabolism in N Supply. Plant physiology **178** (3): 1027–1044.

**Spetea, C., Hundal, T., Lundin, B., Heddad, M., Adamska, I., and Andersson, B.** (2004). Multiple evidence for nucleotide metabolism in the chloroplast thylakoid lumen. Proceedings of the National Academy of Sciences of the United States of America **101** (5): 1409–1414.

References

**Straube, H., Niehaus, M., Zwittian, S., Witte, C.-P., and Herde, M.** (2021). Enhanced nucleotide analysis enables the quantification of deoxynucleotides in plants and algae revealing connections between nucleoside and deoxynucleoside metabolism. The Plant cell **33** (2): 270–289.

**Straube, H., Straube, J., Rinne, J., Fischer, L., Niehaus, M., Witte, C.-P., and Herde, M.** (2023). An inosine triphosphate pyrophosphatase safeguards plant nucleic acids from aberrant purine nucleotides. The New phytologist **237** (5): 1759–1775.

**Sugimoto, H., Kusumi, K., Noguchi, K., Yano, M., Yoshimura, A., and Iba, K.** (2007). The rice nuclear gene, VIRESCENT 2, is essential for chloroplast development and encodes a novel type of guanylate kinase targeted to plastids and mitochondria. The Plant journal for cell and molecular biology **52** (3): 512–527.

**Sweetlove, L.J., Mowday, B., Hebestreit, H.F., Leaver, C.J., and Millar, A.H.** (2001). Nucleoside diphosphate kinase III is localized to the inter-membrane space in plant mitochondria. FEBS letters **508** (2): 272–276.

**Thompson, J.E., Froese, C.D., Madey, E., Smith, M.D., and Hong, Y.** (1998). Lipid metabolism during plant senescence. Progress in lipid research **37** (2-3): 119–141.

**Traut, T.W.** (1994). Physiological concentrations of purines and pyrimidines. Molecular and cellular biochemistry **140** (1): 1–22.

**Tusher, V.G., Tibshirani, R., and Chu, G.** (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences of the United States of America **98** (9): 5116–5121.

**Tyanova, S., Temu, T., and Cox, J.** (2016a). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nature protocols **11** (12): 2301–2319.

**Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J.** (2016b). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature methods **13** (9): 731–740.

**van Rompay, A.R., Johansson, M., and Karlsson, A.** (1999). Phosphorylation of deoxycytidine analog monophosphates by UMP-CMP kinase: molecular characterization of the human enzyme. Molecular pharmacology **56** (3): 562–569.

**van Steensel, B., van Binnendijk, E.P., Hornsby, C.D., van der Voort, H.T., Krozowski, Z.S., Kloet, E.R.** de, and **van Driel, R.** (1996). Partial colocalization of glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons. Journal of cell science **109 (Pt 4)**: 787–792.

# References

**Wang, C., and Liu, Z.** (2006). Arabidopsis ribonucleotide reductases are critical for cell cycle progression, DNA damage repair, and plant development. The Plant cell **18** (2): 350–365.

**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** (1): 144.

**Waterworth, W.M., Bray, C.M., and West, C.E.** (2015). The importance of safeguarding genome integrity in germination and seed longevity. Journal of experimental botany **66** (12): 3549–3558.

**Werner, A.K., Sparkes, I.A., Romeis, T., and Witte, C.-P.** (2008). Identification, biochemical characterization, and subcellular localization of allantoate amidohydrolases from Arabidopsis and soybean. Plant physiology **146** (2): 418–430.

**Witte, C.-P., and Herde, M.** (2020). Nucleotide Metabolism in Plants. Plant physiology **182** (1): 63–78.

**Witz, S., Jung, B., Fürst, S., and Möhlmann, T.** (2012). De novo pyrimidine nucleotide synthesis mainly occurs outside of plastids, but a previously undiscovered nucleobase importer provides substrates for the essential salvage pathway in Arabidopsis. The Plant cell **24** (4): 1549–1559.

**Wood, T.** (1968). The inhibition of pyruvate kinase by ATP. Biochemical and biophysical research communications **31** (5): 779–785.

**Xie, K., Minkenberg, B., and Yang, Y.** (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proceedings of the National Academy of Sciences of the United States of America **112** (11): 3570–3575.

**Xu, J., Yang, J.-Y., Niu, Q.-W., and Chua, N.-H.** (2006). Arabidopsis DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. The Plant cell **18** (12): 3386–3398.

**Yan, H., and Tsai, M.D.** (1999). Nucleoside monophosphate kinases: structure, mechanism, and substrate specificity. Advances in enzymology and related areas of molecular biology **73**: 103-34, x.

**Yan, Y.-B.** (2014). Deadenylation: enzymes, regulation, and functional implications. Wiley interdisciplinary reviews. RNA **5** (3): 421–443.

**Yang, S.-Y., Huang, T.-K., Kuo, H.-F., and Chiou, T.-J.** (2017). Role of vacuoles in phosphorus storage and remobilization. Journal of experimental botany **68** (12): 3045–3055.

**Zhao, X., Liu, N., Shang, N., Zeng, W., Ebert, B., Rautengarten, C., Zeng, Q.-Y., Li, H., Chen, X., Beahan, C., Bacic, A., Heazlewood, J.L., and Wu, A.-M.** (2018). Three UDP-xylose transporters participate in xylan biosynthesis by conveying cytosolic UDP-xylose into the Golgi lumen in Arabidopsis. Journal of experimental botany **69** (5): 1125–1134.

**Zhou, L., Lacroute, F., and Thornburg, R.** (1998). Cloning, expression in Escherichia coli, and characterization of Arabidopsis thaliana UMP/CMP kinase. Plant physiology **117** (1): 245–254.

# 8. List of abbreviations

| | |
|---|---|
| A | Adenosine |
| AMK | ADENOSINE MONOPHOSPHATE KINASE |
| ATC | ASPARTATE CARBAMOYLTRANSFERASE |
| Bp | Base pair |
| C | Cytidine |
| Cas9 | CRISPR-associated 9 |
| CDP-DAG | Cytidine diphosphate diacylglycerol |
| CDP-Etn | Cytidine diphosphate diphosphoethanolamine |
| COX1 | CYTOCHROM OXIDASE 1 |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| Ct | Cycle threshold |
| CTPS | CTP SYNTHASE |
| dNDP | Deoxynucleoside diphosphate |
| dNK | DEOXYNUCLEOSIDE KINASE |
| dNMP | Deoxynucleoside monophosphate |
| dNTP | Deoxynucleoside triphosphate |
| ENT1 | EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 1 |
| FA | Formic acid |
| FAM | 6-Carboxyfluorescein |
| G | Guanosine |
| GFP | Green fluorescent protein |
| HA | Hemagglutinine |
| HPLC | High pressure liquid chromatography |
| iBAQ | Intensity based absolute quantification |
| InDels | Insertions or deletions |
| JOE | 4-5-Dichlorocarboxyfluorescein |
| lacZ | β-Galactosidase |
| LDH | Lactate Dehydrogenase |
| LUH | Leibniz University Hanover |

# List of abbreviations

| | |
|---|---|
| mRNA | Messenger RNA |
| MS | Mass spectrometry |
| mtDNA | mitochondrial DNA |
| NADH | Nicotinamide adenine dinucleotide |
| ncDNA | nuclear DNA |
| NDP | Nucleoside diphosphate |
| NDPK | NUCLEOSIDE DIPHOSPHATE KINASE |
| NMP | Nucleoside monophosphate |
| NSH1 | NUCLEOSIDE HYDROLASE 1 |
| NTP | Nucleoside triphosphate |
| PAM | Protospacer adjacent motif |
| PCR | Polymerase chain reaction |
| PEP | Phosphoenolpyruvate |
| PK | Pyruvate Kinase |
| pLDDT | Predicted local distance difference test |
| PLUTO | PLASTIDIC URACIL TRANSPORTER |
| PNK1 | PLASTIDIC NUCLEOSIDE KINASE 1 |
| PUMPKIN | PLASTID UMP KINASE |
| qPCR | Quantitative polymerase chain reaction |
| rNDP | Ribonucleoside diphosphate |
| rNMP | Ribonucleoside monophosphate |
| RNR | RIBONUCLEOTIDE REDUCTASE |
| rNTP | Ribonucleoside triphosphate |
| rRNA | Ribosomal RNA |
| s | Substrate concentration |
| SD | Standard deviation |
| SDS | Sodium dodecyl sulfate |
| sgRNA | Single guide RNA |
| SUSY | SUCROSE SYNTHASE |
| T | Thymidine |
| T-DNA | Transfer-DNA |

# List of abbreviations

| | |
|---|---|
| TK1a | THYMIDINE KINASE 1a |
| TK1b | THYMIDINE KINASE 1b |
| U | Uridine |
| UBC21 | UBIQUTIN-CONJUGATING ENZYME 21 |
| UCK | URIDINE CYTIDINE KINASE |
| UDP-Ara | Uridine diphosphate arabinose |
| UDP-Gal | Uridine diphosphate galactose |
| UDP-Glc | Uridine diphosphate glucose |
| UGP | UDP GLUCOSE PYROPHOSPHORYLASE |
| UMK | UMP KINASE |
| UPP | URACIL PHOSPHORIBOSYLTRANSFERASE |
| UOX | URATE OXIDASE |
| USP | UDP SUGAR PYROPHOSPHORYLASE |
| V | Reaction velocity |
| XRN | EXORIBONUCLEASE |
| YFP | Yellow fluorescent protein |

# 9. List of figures

# List of figures

# 10. List of tables

# 11. Appendix

```
AcmUMK2     1 ----------------------------------------------------------------
OsaUMK2     1 ----------------------------------------------------------------
BdiUMK2     1 ----------------------------------------------------------------
SitUMK2     1 ----------------------------------------------------------------
SbiUMK2     1 ----------------------------------------------------------------
ZmaUMK2     1 ----------------------------------------------------------------
StuUMK2     1 ----------------------------------------------------------------
SlyUMK2     1 ----------------------------------------------------------------
PvuUMK2     1 ----------------------------------------------------------------
MtrUMK2a    1 ----------------------------------------------------------------
MtrUMK2b    1 ----------------------------------------------------------------
MguUMK2     1 ----------------------------------------------------------------
CsaUMK2     1 ----------------------------------------------------------------
PtrUMK2     1 ----------------------------------------------------------------
VviUMK2     1 ----------------------------------------------------------------
AcoUMK2     1 ----------------------------------------------------------------
EsaUMK2     1 ----------------------------------------------------------------
AthUMK2     1 ----------------------------------------------------------------
BstUMK2     1 ----------------------------------------------------------------
CruUMK2     1 ----------------------------------------------------------------
RcoUMK2     1 MTILHPLKPLCNHQNFAAVALAAVIVAVALAVAVVVALAAVGFPLAGHQVRTGVQLHFVV
CclUMK2     1 ----------------------------------------------------------------
TcaUMK2     1 ----------------------------------------------------------------
MesUMK2     1 ----------------------------------------------------------------
PtrUMK4     1 ----------------------------------------------------------------
MesUMK4     1 ----------------------------------------------------------------
ZmaUMK4     1 ----------------------------------------------------------------
BdiUMK4     1 ----------------------------------------------------------------
OsaUMK4     1 ----------------------------------------------------------------
SbiUMK4     1 ----------------------------------------------------------------
SitUMK4     1 ----------------------------------------------------------------
StuUMK4     1 ----------------------------------------------------------------
SlyUMK4     1 ----------------------------------------------------------------
AcmUMK4     1 ----------------------------------------------------------------
EsaUMK1     1 ----------------------------------------------------------------
AthUMK1     1 ----------------------------------------------------------------
BstUMK1     1 ----------------------------------------------------------------
CruUMK1     1 ----------------------------------------------------------------
OsaUMK3c    1 ----------------------------------------------------------------
CsaUMK3b    1 ----------------------------------------------------------------
AcmUMK3     1 ----------------------------------------------------------------
BdiUMK3     1 ----------------------------------------------------------------
OsaUMK3a    1 ----------------------------------------------------------------
SbiUMK3a    1 ----------------------------------------------------------------
SitUMK3b    1 ----------------------------------------------------------------
OsaUMK3b    1 ----------------------------------------------------------------
ZmaUMK3b    1 ----------------------------------------------------------------
ZmaUMK3a    1 ----------------------------------------------------------------
SbiUMK3b    1 ----------------------------------------------------------------
SitUMK3a    1 ----------------------------------------------------------------
CsaUMK3a    1 ----------------------------------------------------------------
MguUMK3     1 ----------------------------------------------------------------
StuUMK3     1 ----------------------------------------------------------------
SlyUMK3     1 ----------------------------------------------------------------
VviUMK3     1 ----------------------------------------------------------------
TcaUMK3b    1 ----------------------------------------------------------------
AcoUMK3     1 ----------------------------------------------------------------
MesUMK3     1 ----------------------------------------------------------------
EsaUMK3     1 ----------------------------------------------------------------
CruUMK3     1 ----------------------------------------------------------------
BstUMK3     1 ----------------------------------------------------------------
AthUMK3     1 ----------------------------------------------------------------
MtrUMK3     1 ----------------------------------------------------------------
PvuUMK3     1 ----------------------------------------------------------------
TcaUMK3a    1 ----------------------------------------------------------------
RcoUMK3     1 ----------------------------------------------------------------
CclUMK3     1 ----------------------------------------------------------------
PtrUMK3a    1 ----------------------------------------------------------------
PtrUMK3b    1 ----------------------------------------------------------------
```

```
AcmUMK2      1  ---MWRRL-----GFLSPAISSSKRQNRIPESLKL------------------VMLFATD
OsaUMK2      1  ---MWRRQV----GALLLRHRSTPSSTLRHHLPLP------VPDQSPPLASNLLLRLFTS
BdiUMK2      1  ---MWRRRLGALLLRSPSSSTPSSSSSYQQHHHHL--------IPTPNEKPLALSLLRLF
SitUMK2      1  ---MWRRRLGVLLLRSPPSSSSTAASSCQSRRHHL-------LPSEEPLALNRRLARLFTS
SbiUMK2      1  ---MWRRRVG---ALLLQSPLASSSSSSTAASSCQRLRHHLLPSEEPLALNRRLARLFTS
ZmaUMK2      1  ---MWRRRVGALLRS--------PRTSSSASSACQSRRHHLLPSEEPLALNRRLARLFTS
StuUMK2      1  ---MWR-RF----TSLPLFF-----SHIQQVRRAD-------ELKICQ-AFCTEIVKPP
SlyUMK2      1  ---MWRRRF----TSLPLFF-----SHIQQVRRAD-------ELKICQ-AFCTETVKPP
PvuUMK2      1  ---MWR-RAAKS-SSFLPLLLQ---LPKHDASLPQ--------RFTT---GFPFHAPFQE
MtrUMK2a     1  ---MWKRATSSLKSLL--SFHITQEGSKLNNAFHC------------HRFISGSPLHFQ
MtrUMK2b     1  ---MWK------------CATSSFKSLLS--------------------LHITQ
MguUMK2      1  ---MWR-RV----SSLSPLF-SSSKSCPRN-QVAY--------GFNAWQ-MFTTQVLNPA
CsaUMK2      1  ---MWR-RA----VSVSHF--TFAHKSLAHNKDVC--------KLKFWE-TFTTETPMKE
PtrUMK2      1  ---MWR-RV----TSLSPVM-STSKPTSLD-QAAS--------GFKIWRSSFSTETPTLV
VviUMK2      1  ---MWR-RV----TSLSPFI-SSSKPSIRN-QASY--------GGKIWE-LFTTEILTPA
AcoUMK2      1  ---MWW-RV----ASLSRFISTTSSRSSSVNQAAY--------RLNFWQ-VYTTEILTRA
EsaUMK2      1  ---MWR-RV----ASLSPMI-SSSSRSISLNQAAS--------GLKVGE-SFATEIINPD
AthUMK2      1  ---MWR-RV----ALLSPMI-SSSSRSIKLSQAAS--------GLKVGE-SFATDIISQE
BstUMK2      1  ---MWR-RV----ALLSPMI-SSSSRSIALNQAAS--------GLKVGQ-SFATDITNPE
CruUMK2      1  ---MWR-RV----ALLSPMISSSSSRSITLNQAAF--------GLKVGE-SFATDAVNPE
RcoUMK2     61  YKIYWL-RN----GVLMMFM-ISSMLL-----QAAK-------RVKSWR-SLSTEISTLD
CclUMK2      1  ---MWR-RA----ASLSPFI-SSSGSLIRNHQTAY--------NLKIWE-SFTTEIPTQV
TcaUMK2      1  ---MWR-RV----ASLSSLI-SSSNSSFHG-QAAC--------RLTIWE-SLTTGIAQQA
MesUMK2      1  ---MWK-RM----ASLSPLV-SSSKSTLLN-QAAY--------GFNIRK-SLSTGISTPV
PtrUMK4      1  ---MYII----------------GFSIYN--------------------SMATSEFPAQ
MesUMK4      1  ---MKFNKILATALGRKPKIAANKGTNVTDKKIDQLLIRDMEEILESWELPSLLGSEEEE
ZmaUMK4      1  ----------------------------------------------------------M
BdiUMK4      1  -----------------------MADASK-----------------------------
OsaUMK4      1  -----------------------MAHANK----------------------------N
SbiUMK4      1  -----------------------MADAIK-----------------------------
SitUMK4      1  -----------------------MADAIK----------------------------T
StuUMK4      1  -----------------------MDLHKE-----------------------------
SlyUMK4      1  -----------------------MDLHKE-----------------------------
AcmUMK4      1  -----------------------MGTKVD----------------------------KDK
EsaUMK1      1  -----------------------METLVD----------------------------API
AthUMK1      1  -----------------------METPID----------------------------APN
BstUMK1      1  -----------------------METPVD----------------------------ALI
CruUMK1      1  -----------------------METHVE----------------------------AST
OsaUMK3c     1  --------------------MRGGLVASARLL----------PRPLVRWFLQRRAQQ
CsaUMK3b     1  -----------------------METDVG----------------------------VSTS
AcmUMK3      1  -----------------------MGHVEAAPVVT-----------------------NEEK
BdiUMK3      1  -----------------------MGTVVDAPAAV-----------------------TQKE
OsaUMK3a     1  -----------------------MGSVVDAPTVV-----------------------AGQE
SbiUMK3a     1  -----------------------MGTAVDASAAV-----------------------AE
SitUMK3b     1  -----------------------MGTSVDAPAVV-----------------------AE
OsaUMK3b     1  -----------------------MGSVVD----------------------APVVVE
ZmaUMK3b     1  -----------------------MGTVVD----------------------AAPAVVA
ZmaUMK3a     1  -----------------------MGTVVDAPAVV-----------------------AE
SbiUMK3b     1  -----------------------MGTVVDAAPAV-----------------------VAE
SitUMK3a     1  -----------------------MGTVVDAPAVV-----------------------AE
CsaUMK3a     1  -----------------------MGTVVD----------------------AAPI
MguUMK3      1  -----------------------MGTVVDAPKTV-----------------------Q
StuUMK3      1  -----------------------MGTVVE----------------------SAN
SlyUMK3      1  -----------------------MGTVVE----------------------SAN
VviUMK3      1  ----------------------MGSAPE----------------------SVIK
TcaUMK3b     1  -----------------------MGTADA----------------------AS
AcoUMK3      1  -----------------------MGTVVD----------------------AANK
MesUMK3      1  -----------------------MGTSVD----------------------AAI
EsaUMK3      1  -----------------------MGSVVD----------------------AAN
CruUMK3      1  -----------------------MGSIVD-----------------------------
BstUMK3      1  -----------------------MGSIVD-----------------------------
AthUMK3      1  -----------------------MGSVDA-----------------------------
MtrUMK3      1  -----------------------MGTVIE----------------------AAN
PvuUMK3      1  -----------------------MGSVES----------------------AK
TcaUMK3a     1  -----------------------MGTVDA----------------------AN
RcoUMK3      1  ----------------------METSVVG----------------------VAR
CclUMK3      1  -----------------------MGTVVE----------------------TPV
PtrUMK3a     1  -----------------------MGTAGD-----------------------------
PtrUMK3b     1  -----------------------MGSVSD-----------------------------
                                      ▲3(P)
                                      ▲3(P)
```

```
AcmUMK2     35 AVKDDITASR-----------------------------------------------------
OsaUMK2     48 QSGEGGDGAT-----------------------------------------------------
BdiUMK2     50 ASQAGSDGG------------------------------------------------------
SitUMK2     51 QAGSDGGHSR-----------------------------------------------------
SbiUMK2     55 QDGSDGGDTQ-----------------------------------------------------
ZmaUMK2     49 QAGSDGGDTQ-----------------------------------------------------
StuUMK2     39 VEGESNSRRN-----------------------------------------------------
SlyUMK2     40 VEGESNSGRN-----------------------------------------------------
PvuUMK2     42 KGGI-SPKQV-----------------------------------------------------
MtrUMK2a    43 EKDVVSPKHM-----------------------------------------------------
MtrUMK2b    20 EKDVVSPKHM-----------------------------------------------------
MguUMK2     42 ESGV-SDRRR-----------------------------------------------------
CsaUMK2     42 KGT--FQRDK-----------------------------------------------------
PtrUMK2     43 KDGTTSFKER-----------------------------------------------------
VviUMK2     42 KAGI-SSDEK-----------------------------------------------------
AcoUMK2     44 KDGR-PSTVK-----------------------------------------------------
EsaUMK2     43 ERVL-ASKEK-----------------------------------------------------
AthUMK2     43 ERVS-PPKEK-----------------------------------------------------
BstUMK2     43 ERDL-PAKEK-----------------------------------------------------
CruUMK2     44 ERDL-PPKEK-----------------------------------------------------
RcoUMK2    101 TYGTSSLGEK-----------------------------------------------------
CclUMK2     43 KGG-----EK-----------------------------------------------------
TcaUMK2     42 K--G-VSKEK-----------------------------------------------------
MesUMK2     42 KHGT-MLKEK-----------------------------------------------------
PtrUMK4     21 KDGSKAVLQE-----------------------------------------------------
MesUMK4     58 EEEEEEEEEEEEEEEDNVGGGSENLRCIHVKDEDSWQILVSKSSASEQSHMFHEVLSFWRK
ZmaUMK4      2 SQAGAFPPGK-----------------------------------------------------
BdiUMK4      7 DYVGSFPPGK-----------------------------------------------------
OsaUMK4      8 HIESFPPPGK-----------------------------------------------------
SbiUMK4      7 SQVGSFPPGK-----------------------------------------------------
SitUMK4      8 EAGSF-PPGK-----------------------------------------------------
StuUMK4      7 -GDRGSAKQK-----------------------------------------------------
SlyUMK4      7 -GDTGSAKQK-----------------------------------------------------
AcmUMK4     10 NSDGSLLRDK-----------------------------------------------------
EsaUMK1     10 KDDHESPRWQ-----------------------------------------------------
AthUMK1     10 KDEHECPRWK-----------------------------------------------------
BstUMK1     10 KDEHEIPRWK-----------------------------------------------------
CruUMK1     10 KDERESPRWK-----------------------------------------------------
OsaUMK3c    28 DINENMLGGK-----------------------------------------------------
CsaUMK3b    11 EEANGSTVQK-----------------------------------------------------
AcmUMK3     16 DVSESLLGDK-----------------------------------------------------
BdiUMK3     16 EVAENMLGNK-----------------------------------------------------
OsaUMK3a    16 EVTDNMLGDK-----------------------------------------------------
SbiUMK3a    14 EVTENMLGGK-----------------------------------------------------
SitUMK3b    14 EVTDNMLGGK-----------------------------------------------------
OsaUMK3b    13 GVAENMLGDK-----------------------------------------------------
ZmaUMK3b    14 EVTENMLGGK-----------------------------------------------------
ZmaUMK3a    14 EVTENMLGGR-----------------------------------------------------
SbiUMK3b    15 EVTENMLGGK-----------------------------------------------------
SitUMK3a    14 EVTENMLGGK-----------------------------------------------------
CsaUMK3a    11 KETNGSLAEK-----------------------------------------------------
MguUMK3     13 VGNGDAHTNK-----------------------------------------------------
StuUMK3     10 QGAVSLPTNK-----------------------------------------------------
SlyUMK3     10 QGAVSLPTNK-----------------------------------------------------
VviUMK3     11 ETNENLLAEK-----------------------------------------------------
TcaUMK3b     9 KDVTVRLAER-----------------------------------------------------
AcoUMK3     11 GVNGSLPNEK-----------------------------------------------------
MesUMK3     10 KEENGNIAEK-----------------------------------------------------
EsaUMK3     10 KEVNGTASGK-----------------------------------------------------
CruUMK3      7 -AANGTASGK-----------------------------------------------------
BstUMK3      7 -AANGTASGK-----------------------------------------------------
AthUMK3      7 ----ANGSGK-----------------------------------------------------
MtrUMK3     10 KDTNGSVLNK-----------------------------------------------------
PvuUMK3      9 KDANGGLLEK-----------------------------------------------------
TcaUMK3a     9 KDVNVSLTDE-----------------------------------------------------
RcoUMK3     11 KEENGDVVEK-----------------------------------------------------
CclUMK3     10 KEADATVTVK-----------------------------------------------------
PtrUMK3a     7 -------SEK-----------------------------------------------------
PtrUMK3b     7 -------AEK-----------------------------------------------------
               ▲10(P)
```

```
AcmUMK2     45  ------------------------------------------------ RPFITFVLGGPGSGK
OsaUMK2     58  ------------------------------------------------ KPFIAFVLGGPGSGK
BdiUMK2     59  ------------------------------------------------ RPFIAFILGGPGSGK
SitUMK2     61  ------------------------------------------------ KPFVAFVLGGPGSGK
SbiUMK2     65  ------------------------------------------------ KPFIAFVLGGPGSGK
ZmaUMK2     59  ------------------------------------------------ KPFIAFVLGGPGSGK
StuUMK2     49  ------------------------------------------------ IPFVAFVLGGPGSGK
SlyUMK2     50  ------------------------------------------------ SPFVAFVLGGPGSGK
PvuUMK2     51  ------------------------------------------------ APLITFVLGGPGSGK
MtrUMK2a    53  ------------------------------------------------ DSVITFVLGGPGSGK
MtrUMK2b    30  ------------------------------------------------ DSVITFVLGGPGSGK
MguUMK2     51  ------------------------------------------------ IPFVTFVLGGPGSGK
CsaUMK2     50  ------------------------------------------------ TPFITFVLGGPGSGK
PtrUMK2     53  ------------------------------------------------ NPFITFVLGGPGSGK
VviUMK2     51  ------------------------------------------------ TPFITFVMGGPGSGK
AcoUMK2     53  ------------------------------------------------ TPFVAFVLGGPGSGK
EsaUMK2     52  ------------------------------------------------ APFITFVLGGPGSGK
AthUMK2     52  ------------------------------------------------ APFITFVLGGPGSGK
BstUMK2     52  ------------------------------------------------ APFITFVLGGPGSGK
CruUMK2     53  ------------------------------------------------ APFITFVLGGPGSGK
RcoUMK2    111  ------------------------------------------------ TPFMTFVLGGPGSGK
CclUMK2     47  ------------------------------------------------ GPFICFVLGGPGSGK
TcaUMK2     49  ------------------------------------------------ TPFITFVLGGPGSGK
MesUMK2     51  ------------------------------------------------ TPFITFVLGGPGSGK
PtrUMK4     31  ------------------------------------------------ TTVVVFVLGGPGSGK
MesUMK4    118  KEKEESELKSKEESELKSKENGRFSSSMASSELPTQKVGSENSTN EAVVVFVLGGPGSGK
ZmaUMK4     12  ------------------------------------------------ KITVVFVIGGPGSGK
BdiUMK4     17  ------------------------------------------------ KITVVFVIGGPGSGK
OsaUMK4     18  ------------------------------------------------ KITIVFVIGGPGSGK
SbiUMK4     17  ------------------------------------------------ KITVVFVLGGPGSGK
SitUMK4     17  ------------------------------------------------ KITVVFVIGGPGSGK
StuUMK4     16  ------------------------------------------------ KVKIVFVIGGPGSGK
SlyUMK4     16  ------------------------------------------------ KVKIVFVIGGPGSGK
AcmUMK4     20  ------------------------------------------------ KITIVFVLGGPGCGK
EsaUMK1     20  ------------------------------------------------ KSTVVFVLGGPGSGK
AthUMK1     20  ------------------------------------------------ KSTVVFVLGGPGSGK
BstUMK1     20  ------------------------------------------------ KSTVVFVLGGPGSGK
CruUMK1     20  ------------------------------------------------ KSTVVFVLGGPGSGK
OsaUMK3c    38  ------------------------------------------------ KVKIIFVLGGPGSGK
CsaUMK3b    21  ------------------------------------------------ KPTVVFVLGGPGSGK
AcmUMK3     26  ------------------------------------------------ KVTVVFVLGGPGSGK
BdiUMK3     26  ------------------------------------------------ KVTVVFVLGGPGSGK
OsaUMK3a    26  ------------------------------------------------ KVTVVFVLGGPGSGK
SbiUMK3a    24  ------------------------------------------------ KVTVVFVLGGPGSGK
SitUMK3b    24  ------------------------------------------------ KVTVVFVLGGPGSGK
OsaUMK3b    23  ------------------------------------------------ KVTVVFVLGGPGSGK
ZmaUMK3b    24  ------------------------------------------------ KVTVVFVLGGPGSGK
ZmaUMK3a    24  ------------------------------------------------ KVTAVFVLGGPGSGK
SbiUMK3b    25  ------------------------------------------------ KVTVVFVLGGPGSGK
SitUMK3a    24  ------------------------------------------------ KVTVVFVLGGPGSGK
CsaUMK3a    21  ------------------------------------------------ KPTVVFVLGGPGSGK
MguUMK3     23  ------------------------------------------------ KVTVVFVLGGPGSGK
StuUMK3     20  ------------------------------------------------ KVTVIFVLGGPGSGK
SlyUMK3     20  ------------------------------------------------ KVTVIFVLGGPGSGK
VviUMK3     21  ------------------------------------------------ KVKVVFVLGGPGSGK
TcaUMK3b    19  ------------------------------------------------ KPIVVFVLGGPGSGK
AcoUMK3     21  ------------------------------------------------ KFKVVFVLGGPGSGK
MesUMK3     20  ------------------------------------------------ KPTVVFVLGGPGSGK
EsaUMK3     20  ------------------------------------------------ KPTVIFVLGGPGSGK
CruUMK3     16  ------------------------------------------------ KPTVIFVLGGPGSGK
BstUMK3     16  ------------------------------------------------ KPTVIFVLGGPGSGK
AthUMK3     13  ------------------------------------------------ KPTVIFVLGGPGSGK
MtrUMK3     20  ------------------------------------------------ NPTVVFVLGGPGSGK
PvuUMK3     19  ------------------------------------------------ NPTVVFVLGGPGSGK
TcaUMK3a    19  ------------------------------------------------ KPIVVFVLGGPGSGK
RcoUMK3     21  ------------------------------------------------ KPTVVFVLGGPGSGK
CclUMK3     20  ------------------------------------------------ KPTVVFVLGGPGSGK
PtrUMK3a    10  ------------------------------------------------ KPTVVFVLGGPGSGK
PtrUMK3b    10  ------------------------------------------------ KPAVIFVLGGPGSGK
```

▲27

# Appendix

| | | |
|---|---|---|
| AcmUMK2 | 60 | GTQCARIASAFGFAHLSAGDLLREEMSCDSKKGKLIRDLIMEGKIVPSEITVHLLKEAIM |
| OsaUMK2 | 73 | GTQCVRIASDFGFAHLSAGDLLRSEISTGSEKGELILNIIKEGKIVPSEITVELIRKAME |
| BdiUMK2 | 73 | GTQCTRIASDFGFAHVSAGDLLRNEISSGTDKGELILEIIKEGRIVPSEITVELIRKAIE |
| SitUMK2 | 76 | GTQCSKIASDFGFAHLSAGDLLRHEIASGSEKGQLILDIIKEGRIVSSEITVELIRKAME |
| SbiUMK2 | 80 | GTQCTKIASDFGFAHLSAGDLLRHEIASGSEKGELILEIIKEGRIVPSEITVELIRKAME |
| ZmaUMK2 | 74 | GTQCTKIASDFGFAHLSAGDLLRHEIASGSEKGELILEIIKEGRIVPSEITVELIRKAME |
| StuUMK2 | 64 | GTQCLKIAETFGFDHIGAGDLLRKEIHSDSENGAMIQKLMKEGSIAPSEVTVKLLKKAIE |
| SlyUMK2 | 65 | GTQCLKIAETFGFDHIGAGDLLRKEMHSDSENGAMIQKLMKEGSIAPSEVTVKLLKKAIE |
| PvuUMK2 | 66 | GTQCAKIVETFGFKHLSAGDLLRREMISDSEYGSSILNTISEGKIVPSEVTVKLLREME |
| MtrUMK2a | 68 | GTQCARIVETFGFKHLSAGDLLRKEMVSDSEYGAMILETIREGKIVPSAVTVKLLREMQ |
| MtrUMK2b | 45 | GTQCARIVETFGFKHLSAGDLLRKAMVSDSEYGAMILETIREGRIVPSAVTVKLLREMQ |
| MguUMK2 | 66 | GTQCTRIVENFGFTHLSAGDLLRHEISSNIENGSMILNTIKEGKIVPSEVTVKLLQKALE |
| CsaUMK2 | 65 | GTQCMKIVENFGFTHLSAGDLLRREIASNSADGTMILNTIKEGKIVPSEITVKLLQKEME |
| PtrUMK2 | 68 | GTQCQKIVETFGFKHLSAGELLRREIESNSEHWSQMLNTIKEGRIVPSEVTVKLLQQEME |
| VviUMK2 | 66 | GTQCAKIVETFGFTHLSAGELLRREISCNSEHGSMILDSIREGKIVPSEVTVKLLEKEME |
| AcoUMK2 | 68 | GTQCTKIVESFGFTHLSAGDLLRHAISSNYENSAMILDTIKEGKIVPSEVTVKLLQRALE |
| EsaUMK2 | 67 | GTQCEKIVETFGLTHLSAGDLLRREISLNTKNGAMILNLIKIGKIVPSEVTVKLLQKELE |
| AthUMK2 | 67 | GTQCEKIVETFGLQHLSAGDLLRREIAMHTENGAMILNIIKDGKIVPSEVTVKLLQKELE |
| BstUMK2 | 67 | GTQCEKIVETFGLQHLSAGDLLRREIAMNTKNGAMILNLIKDGKIVPSEVTVKLLQKELE |
| CruUMK2 | 68 | GTQCEKIVETFGLQHLSAGDLLRREIAMNTKNGDMILNLIKDGKIVPSEVTVKLLQRELE |
| RcoUMK2 | 126 | GTQCLKIAKTFGFKHLSAGDLLRREILSNSDDGAMILNTIKEGRIVPSEVTVKLLKKEME |
| CclUMK2 | 62 | GTQCAKIVKNYGLTHLSAGELLRREIASNSEYGTTILNTIKEGKIVPSEVTVSLLQKEME |
| TcaUMK2 | 64 | GTQCIKIVETFGFTHLSAGDLLRQEITSNSADGAMILNTIKEGRIVPSEVTVKLLQKEME |
| MesUMK2 | 66 | GTQCIKIAQTFGFTHLSAGDLLRREILSNSEYGTMILDTIKEGIIVPSEVTVKLLKKTIE |
| PtrUMK4 | 46 | GTQCPKIVEHFGFRHLCAGDLLQAEIESESENGTMIQNFKKEGKIVPSEITVKLLQQAMQ |
| MesUMK4 | 178 | GTQCPKIVEHFGFTNICAGELLQAEVESGSENGEMILKFREEGKIVPSEITVKLLQQAMQ |
| ZmaUMK4 | 27 | GTQCSKIVRHFGFTHLSAGDLLRQQVQSDIEHGAMIKNLVHEGKIVPSDIIVLLLTAML |
| BdiUMK4 | 32 | GTQCAKIVNQFGFTHLSAGELLREEVKSDIEGTMIKNLMQEGKIVPSDIIVKLILKAML |
| OsaUMK4 | 33 | GTQCAKIVKQFGFTHLSAGDLLREEAKYDIEQGTMIKNLVNEGKIVSSDLIVKLLFKAMR |
| SbiUMK4 | 32 | GTQCSNIVKQFGFTHLSAGDLLREEAKSDIEQGTMVKNLMHEGKIVPSELIIKLLLKAML |
| SitUMK4 | 32 | GTQCSKIVKHFGFTHLSAGDLLREEAKYDIEQGTMIKNLMHEGKIVPSELIVKLLLKAML |
| StuUMK4 | 31 | GTQCKRIAQQFGYTHLSVGEILRQEISSGSETGSMIQKIIMKEGKIVPSDVTVKLLQQAMQ |
| SlyUMK4 | 31 | GTQCKRIAQQFGYTHLSVGEILRQETSSGSETGHMVQKIIMKEGKIVPSDVTVKLLQQAMQ |
| AcmUMK4 | 35 | GTQCEKIAKHFGFTHLSVGDLLRAETKIGSAYGAMIQTMMKEGEIVPSEVVVKLLQQAML |
| EsaUMK1 | 35 | GTQCANIVEHFGYTHLSAGDLLRAEISSGSEFGAMIQSMIAEGRIVPSEITVKLLCEAMK |
| AthUMK1 | 35 | GTQCANIVKHFSYTHFSAGDLLRAEIKSGSEFGAMIQSMIAEGRIVPSEITVKLLCKAME |
| BstUMK1 | 35 | GTQCANIVKHFSYTHFSAGDLLRAEIKSGSEFGAMIQSMIAEGRIVPSEITVKLLCKAME |
| CruUMK1 | 35 | GTQCANIVKHFSYTHLSAGDLLRAEIKSGSEFGAMIQSMIAEGRIVPSEITVKLLCKAME |
| OsaUMK3c | 53 | GTQCSNIVEHFGFIHLSAGELLRAEINSGSENGTMIDTIITEGKIVPSEITIKLLQEAII |
| CsaUMK3b | 36 | GTQCACIVEHFGFTHFSAGDLLRAEIKSGSENGLMIKSMIGEGKIVPSEVTVKLLQKAMQ |
| AcmUMK3 | 41 | GTQCANIVQNFGFTHLSAGDLLRAEIKSGSENGTMIENMIKEGKIVPSEVTIKLLQRAML |
| BdiUMK3 | 41 | GTQCSNIVEHFGFTHLSAGDLLRAEIKSGSENGTMIENMIKEGKIVPSEVTIKLLQQAMI |
| OsaUMK3a | 41 | GTQCANIVEHFGFIHLSAGDLLRAEIKSGSENGTMIENMIKEGKIVPSEVTIKLLQEAMI |
| SbiUMK3a | 39 | GTQCTNIVEHFGFTHLSAGDLLRAEIKSGSENGTMIDNMIKEGKIVPSEVTIKLLKEAMI |
| SitUMK3b | 39 | GTQCTNIVEHFGFTHLSAGDLLRAEIKSGSENGTMIETMIKEGKIVPSEVTIKLLQEAMI |
| OsaUMK3b | 38 | GTQCANIVEHFGFTHLSAGDLLRAEIKSGSENGTMIENMIKEGKIVPSEVTIKLLQDAMI |
| ZmaUMK3b | 39 | GTQCANIVEHFGFTHLSAGDLLRAEIKSGSENGTMIENMIKEGKIVPSEVTIKLLQEAMI |
| ZmaUMK3a | 39 | GTQCANIVEHFGFTHLSAGDLLRAEIKSGSENGTMIENMIKEGKIVPSEVTIKLLQEAMI |
| SbiUMK3b | 40 | GTQCANIVEHFGFTHLSAGDLLRAEIKSGSENGTMIENMIKEGKIVPSEVTIKLLQEAMI |
| SitUMK3a | 39 | GTQCANIVEHFGFTHLSAGDLLRAEIKSGSENGTMIENMIKEGKIVPSEVTIKLLQQAMI |
| CsaUMK3a | 36 | GTQCANIVQHFGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQRAIE |
| MguUMK3 | 38 | GTQCANIVEHYGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQRAIE |
| StuUMK3 | 35 | GTQCTNIVEHFGYTHLSAGDLLRAEIKSGSENGTMISNMIKEGKIVPSEVTIKLLQRAIQ |
| SlyUMK3 | 35 | GTQCTNIVEHFGYTHLSAGDLLRAEIKSGSENGTMISNMIKEGKIVPSEVTIKLLQRAIQ |
| VviUMK3 | 36 | GTQCANIVKHFGYTHLSAGDLLRAEIKSGSENGNMIQSMIKEGKIVPSEVTIKLLQRAIL |
| TcaUMK3b | 34 | GTQCANIVQHFGYTHLSAGDLLRAECNSGSENGTMIQNMVNEGKIVPSEVTIELLQKAML |
| AcoUMK3 | 36 | GTQCANIVEHFGFTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTVKLLQRAMQ |
| MesUMK3 | 35 | GTQCTNIVQHFGYTHLSAGDLLREEIKSGSENGTMIQDMIKEGKIVPSEVTIKLLQKAMQ |
| EsaUMK3 | 35 | GTQCAYIVEHFGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAIQ |
| CruUMK3 | 31 | GTQCAYIVEHFGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAIQ |
| BstUMK3 | 31 | GTQCAYIVEHFGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAIQ |
| AthUMK3 | 28 | GTQCAYIVGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAIQ |
| MtrUMK3 | 35 | GTQCANVVEHFGFTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIRLLQQAIK |
| PvuUMK3 | 34 | GTQCANIVENFGFTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAMQ |
| TcaUMK3a | 34 | GTQCANIVQHFLYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAML |
| RcoUMK3 | 36 | GTQCANIVEHFGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAMQ |
| CclUMK3 | 35 | GTQCANIVEHFGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAME |
| PtrUMK3a | 25 | GTQCANIVEHFGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAMQ |
| PtrUMK3b | 25 | GTQCANIVEHFGYTHLSAGDLLRAEIKSGSENGTMIQNMIKEGKIVPSEVTIKLLQKAMQ |

▲29            ▲49        ▲57(P)              73▲ ▲75▲78(alt. P)

>-start of At3g60961--

X

```
AcmUMK2    120 SSSAD.FLIDGFPRSEENRIAFETIIG VEPNLVLFFDCPEEEMVKRVLGRNQGRVDDNIE
OsaUMK2    133 SSDAKRVLIDGFPRCEENRIAFERITGTEPDLVTFFDCPE EMVKRLLGRNQGRVDDNIE
BdiUMK2    133 STTAKRVLIDGFPRCEENRIAFEKITGTEPDLVIFFDCPE EMVKRLLGRNQGRVDDNIE
SitUMK2    136 TSNADKVLIDGFPRCEENRITFERIVGTEPDIVVFFDCPE DEMVKRVLGRVDDNIE
SbiUMK2    140 TKNAKRVLIDGFPRCEENRIAFEKIVGTEPDIVIFFDCPE EMVKRLLGRNQGRVDDNIE
ZmaUMK2    134 MNNAKRVLIDGFPRCEENRIAFERIVGTEPDIVIFFDCPE EMVKRLLGRNQGRVDDNIE
StuUMK2    124 SAENRKFLIDGFPRSEENRVAYERIIGAEPNFVLFFDCPEEVVKRVLNRNEGRVDDNEH
SlyUMK2    125 SAENRKFLIDGFPRSEENRVAYERIIGAEPNFVLFFDCPEEVVKRVLNRNEGRVDDNEH
PvuUMK2    126 ASDNQKFLIDGFPRSEENRAAFEQIVGAEPRFVLFFDCPEEEMVKRVLSRNEGRIDDNID
MtrUMK2a   128 YGDNRKFLIDGFPRSEENRIAFEHITGTEPDFVLFFDCPEEEMVKRVLSRNQGRIDDNID
MtrUMK2b   105 YGDNRKFLIDGFPRSEENRIAFEHITGTEPDFVLYFDCPEEEMVKRVLSRNQGRIDDNID
MguUMK2    126 SSENC.FLIDGFPRTEENRIAVERVTGSEPDIVVERVTGSEPDIVVFFDCPEEEMVKRVLNRNDDNLD
CsaUMK2    125 SSDNYKFLIDGFPRSEENRIAFEQIMGVEPDVVLFFDCPE DEMVKRVLNRNQGRVDDNIV
PtrUMK2    128 SSDSNKFLIDGFPRTEENRIAFEQLIGLEPNVVLFFDCPEEEMVKRVLNRNQGRVDDNID
VviUMK2    126 SSKNNKFLIDGFPRTEENRIAFERVIGAEPNFVLFFHCPEEEMVKRLLSRNEGRVDDNID
AcoUMK2    128 SSDNYKFLIDGFPRSEENRIAFEKIIGAEPNIVLFFDCPEEEMVKRVLSRNEGRVDDNID
EsaUMK2    127 SSDSCKFLIDGFPRSEENRVAFERIIGADPNVVLFFDCPEEEMVKRVLNRNQGRVDDNIT
AthUMK2    127 SSDNRKFLIDGFPRTEENRVAFERIIRADPDVVLFFDCPEEEMVKRVLNRNQGRVDDNIT
BstUMK2    127 SSDSRKFLIDGFPRTEENRVAFERIIRADPDVVLFFDCPEEEMVKRVLNRNQGRVDDNIT
CruUMK2    128 SSDSRKFLIDGFPRTEENRVAFERIIRADPDVVLFFDCPEEEVTRVLNRNQGRVDDNIT
RcoUMK2    186 LSDNSKFLIDGFPRTEENRIAFEHIIGAEPNIVLFFDCPQEEMVKRVLNRNEGRVDDNID
CclUMK2    122 SSDSKKFLIDGFPRSEENRAAFERIMGAEPDIVLFFDCPEEEMVNRVLNRNEGRVDDNID
TcaUMK2    124 SNDNHKFLIDGFPRSEENRIAFERIIGAEPNIVLFFDCPEEEMVKRVLNRNQGRVDDNID
MesUMK2    126 SSDNYKFLIDGFPRSEENRIAFEHIIGVEPNVVLFFDCPEEEMVKRVLNRNEGRVDDNID
PtrUMK4    106 QSDNKRFIIDGFSRNEENRAAFENIVRIKPEFVLFFDCPEEQITKRILNRNQGRVDDNIE
MesUMK4    238 QSEKKKFLIDGFPRNEENRTMKIEPDLVLFFDCPSEVITKRLLSRNQGRVDDNIY
ZmaUMK4     87 QSGNDKFLVDGFPRNEENRRAYESVIGIEPELVLFIDCPREELERRILHRDQGRDDDNVD
BdiUMK4     92 ESGNDKFLIDGFPRNEENRQAYENIVNIEPEFVLFIDCSLEEMERRILNRNQGRDDDNVT
OsaUMK4     93 ESGNDKFLVDGFPRNEENRHAYENIIHIEPEFILFIDCSKEEMERRILNRNQGRDDDNID
SbiUMK4     92 QSGNDKFLVDGFPRNEENRQAYESVIGIEPEFILFIDCPKEEIERRILHRNQGRDDDNID
SitUMK4     92 QSGNDKFLVDGFPRNEENRQAYDNIIGIEPEFVLFIDCSKEEMERRILNRNQGRDDDNID
StuUMK4     91 GINSDKFLIDGFPRNEENVKAFEDLTKVEPEFVLYLDCPQDEMEKRLLSRNEGREDDNIE
SlyUMK4     91 GIDNDKFLIDGFPRDEENVKAFEDLTKVEPEFVLYLDCPQDEMEKRLLSRNEGREDDNIE
AcmUMK4     95 RSGNNKFLIDGFPRNEENRLTYESVMKIEPAFILFLDCPQEEMERRVLNRNQGRDDDNIE
EsaUMK1     95 ESGNDKFLIDGFPRNQENRIVFENVAKIEPAFVLFFDCPEEEIERRIMNRNQGREDDNIE
AthUMK1     95 ESGNDKFLIDGFPRNEENRNVFENVARIEPAFVLFFDCPEEEIERRIMSRNQGREDDNIE
BstUMK1     95 ESDNDNFLIDGFPRNEENRIAFENVAKIEPAFVLFFDCPEEEIEKRIMNRNQGREDDNIE
CruUMK1     95 ESGNDKFLIDGFPRNEENRIVFENVAKIEPAFVLFFDCPEEEIERRIMNRNQGREDDNIE
OsaUMK3c   113 KGGNDKYIIDGFPRNEENRVVFESVISISPEFVLFFDCSEEEMERRILLGRNQGRSDDNIE
CsaUMK3b    96 ESGNDKFLIDGFPRNQENRAAFEAVTGIEPAFVLFFDCPEEEMERRILHRNQGRDDDNIE
AcmUMK3    101 ESGNDKFLIDGFPRNEENRAAFENVTKITPEFVLFFDCSEEEMERRILLNRNQGRDDNIE
BdiUMK3    101 NNENDKFLIDGFPRNEENRAAFENVTKISPAFVLFFNCSEEEMERRLLGRNEGRVDDNIE
OsaUMK3a   101 KSGNDKFLIDGFPRNEENRAAFENVTKITPAFVLFFDCSEEEMERRLLGRNQGRVDDNIE
SbiUMK3a    99 KSENDKFLIDGFPRNEENRAAFENVTKISPAFVLFFDCAEEEMERRLLGRNQGRVDDNIE
SitUMK3b    99 KSENDKFLIDGFPRNEENRSAFENVTKISPAFVLFFDCTEEEMEKRLLGRNQGRVDDNIE
OsaUMK3b    98 KNENDKFLIDGFPRNEENRAAFENVTKISPAFVLFFDCSEEEMERRLLGRNQGRVDDNIE
ZmaUMK3b    99 KSENDKFLIDGFPRNEENRAAFENVTKISPAFVLFFDCSEDDMEKRLLGRNQGRVDDNIE
ZmaUMK3a    99 KNENDKFLIDGFPRNEENRAAFENVTKISPAFVLFFDCSEKEMEKRLLGRNQGRVDDNIE
SbiUMK3b   100 KNENDKFLIDGFPRNEENRAAFENVTKISPAFVLFFDCSEEEMERRLLGRNQGRVDDNIE
SitUMK3a    99 KSENDKFLIDGFPRNEENRAAFENVTKISPAFVLFFDCSEEEMERRLLGRNQGRVDDNIE
CsaUMK3a    96 ETGNDKFLIDGFPRNEENRAAFEVVTGIEPSIVLFFDCPEEEMEKRLLSRNEGRVDDNIE
MguUMK3     98 ENGNDKFLIDGFPRNEENRAAFESVTGIEPEFVLFFDCSEEEMERRLLSRNQGREDDNIE
StuUMK3     95 ENGNDKFLIDGFPRNEENRAAFELVTGIEPEFVLFFDCPEAEMEKRLLGRNQGREDDNIE
SlyUMK3     95 ENGNDKFLIDGFPRNEENRAAFELVTGIEPEFVLFFDCPEAEMEKRLLGRNQGREDDNIE
VviUMK3     96 EDSNDKFLIDGFPRNEENRAAFEAVTKIEPEFVLFFDCSEEEMERRILNRNQGREDDNVE
TcaUMK3b    94 ESGNNKFLIDGFPRNEENRAAFEAVTKIEPEFVLFFNCPQEEMERRILNRNQGRDDDNIE
AcoUMK3     96 ESDNDKFLIDGFPRNEENRAAFENVTGIVPEFVLFFDCSEEEMEKRLLSRNQGRDDDNIE
MesUMK3     95 ENENDKFLIDGFPRNEENRAAFESITKIEPQFVLFFDCSEEEMERRILNRNQGREDDNIE
EsaUMK3     95 DNGNDKFLIDGFPRNEENRAAFEKVTEIEPKFVLFFDCSEEEMEMERRLLGRNQGREDDNIE
CruUMK3     91 ENGNDKFLIDGFPRNEENRAAFEKVTEIEPKFVLFFDCSEEEMEKRLLGRNQGREDDNID
BstUMK3     91 ENGNDKFLIDGFPRNEENRAAFEKVTEIEPKFVLFFDCSEEEMEKRLLGRNQGREDDNIE
AthUMK3     88 ENGNDKFLIDGFPRNEENRAAFEKVTEIEPKFVLFFDCPEEEMEKRLLGRNQGREDDNIE
MtrUMK3     95 DNGNDKFLIDGFPRNEENRAAFERVTGIEPAFVLYFDCPEEEMERRVLSRNQGRVDDNIE
PvuUMK3     94 ENGNDKFLIDGFPRNEENRAAFEKVTGIEPAFVLYFDCPEEEMERRVLSRNQGRVDDNIE
TcaUMK3a    94 ESGNDKFLIDGFPRNEENRAAFEAVTKIEPDFVLFFKCPEEEMERRLLSRNQGREDDNIE
RcoUMK3     96 ESGNDKFLIDGFPRNEENRAAFESVTKIVPEFVLFFDCSEEEMERRLLSRNQGRVDDNIE
CclUMK3     95 ESGNDKFLIDGFPRNEENRAAFEAVTKIEPEFVLFFDCSEEEMERRILNRNQGREDDNVE
PtrUMK3a    85 DSGNDKFLIDGFPRNEENRAAFEAVTKIEPAFVLFFDCPEEEMEKRILSRNQGREDDNIE
PtrUMK3b    85 ESGNDKFLIDGFPRNEENRAAFEAVTKIEPAFVLFFNCPEEEMEKRILSRNQGREDDNIE
                              99▲  ▲101▲105      ▲113(P)           ▲133▲137
```

# Appendix

```
AcmUMK2    180  TIKKRLKVFEMLNLPVINYYSAKGKLYKINATGTVDEIFEK-VRHLFGSLS | SVEKQKSIS
OsaUMK2    193  TIKKRLKVFESLNIPVVDYYTSRGKVHKINATGTEEEIFGA-VHKIFSSLR | F--------
BdiUMK2    193  TIKKRLKVFESLNIPVVEYYSRGKAHKINATGTEDEIFEA-VRKLFSSLR | L--------
SitUMK2    196  TIKKRLKVFESLNIPVVSYYSSRGKVHKINATGTADEIFEA-VRRLFSSLR | F--------
SbiUMK2    200  TIKKRLKVFESLNIPVVDYYSSRGKVHKINATGTADEIFEA-VRRLFSSLR | L--------
ZmaUMK2    194  TIKKRLKVFESLNIPVVDYYSSRGKVHKINATGTADEIFEA-VRRLFSSLR | ---------
StuUMK2    184  TVKERLKVYKAITLPVANHYAKKGKLYKVDGTGTQEEIFER-VRPIFASLS | RLST-----
SlyUMK2    185  TVKERLKVYKAITLPVANHYAMKGKLYKVDGTGTQEEIFER-VRPIFASLR | LST------
PvuUMK2    186  TMRNRLKVFEALNLPVIYYSKKGKLFINAVGTVDEIFEQ-VRPVFEACE | QEAK-----
MtrUMK2a   188  TIKKRLKVFEALNLPVIHYARRGRLHRINAVGTEDEIFEQ-VRPVFAACE | QTAA-----
MtrUMK2b   165  TIKKRLKVFEALNLPVIHYARRGRLHRINAVGTEDEIFEQ-VRPVFAACE | QTAA-----
MguUMK2    186  TVKERLKVTKLNLPVIDHYSKKGKLYKIDGTGSEDEIYER-VRPVFAALR | ---------
CsaUMK2    185  TIKKRLKVFDALNLPVKYYMEKGKLYKIRAVGSVDEIYKQ-VYPVFASLN | FEQQVRE--
PtrUMK2    188  TVKKRLKVFEILNLPVIYYSKRGKLCKINAVGTEDEIFEK-VRPIFSACA | GK-------
VviUMK2    186  TIKKRLEVFTALHLPVIKYYSEKGKLYKINAVGTVDEIFEQ-VRPVFAVCE | ATK------
AcoUMK2    188  TIKKRLKVFAALNLPVIKHYSERGKVFKINAVGTVDDIFEQ-VRPVFTAHE | KEIE-----
EsaUMK2    187  TMKKRLKIFEALNRPVIYYKNKGKLYTINAVGTVDDIFQQ-VLPIFTPFE | QLKQSRHVN
AthUMK2    187  TMKKRLKIFNALNRPVIYYKNKGKLYTINAVGTVDDIFQH-VLPIFNSFE | QLKESSHVN
BstUMK2    187  TMKKRLKIFNALNRPVIDYKNKGKLYTINAIGTVDDIFQH-VLPIFNSFE | QLKESRHVN
CruUMK2    188  TMKKRLKIFNGLNRPVIDYKNKGKLYTINAVGTVDDIFQQHVLPIFSSFE | QLKKESRHV
RcoUMK2    246  TIKKRLEVFSALNLPVIGYYSKKGKLHTINAVGTVDEIFEQ-VRPVFAACE | AMK------
CclUMK2    182  TVRKRLQVEKALNLPVINYYARRGKLYTINAVGTVDEIFEQ-VRAVFAALK | LVTE-----
TcaUMK2    184  TVRKRLKVFEALNLPVINYYSQRGKLYTINAVGTVDEIFEQ-VLPVFTASE | ---------
MesUMK2    186  TIKKRLKVFSSLNLPVNYYSKRGKLHTINAVGTVDEIFEQ-VRAVFSVCE | AMK------
PtrUMK4    166  TIGKRLKVYFESTLPVINYYNSKGKLVQKIDAQRSIEEVFED-VKSVFAKLR | PVARVGSTK
MesUMK4    298  TIQKRLEVIFESTLPVINYYSSKGKIKKIDAQRSIDEVFED-VKGVFSKLK | PKSVVGLKN
ZmaUMK4    147  TIRKRFQVFHDSILPVVLYYDRMGKVRRVDGAKSADAVFLD-VKAIFTQLL | TTQVHSLTH
BdiUMK4    152  TVRRRFGVFQESTLPVIQHYEKLGKLRRVDGDRQPDIVFED-VKAVFAQLN | IQTNQASHV
OsaUMK4    153  TIRRRFDVFQQQTLPVIQYYEKRGKLRRVDGNRQVDEVFED-VKAIFAQLN | NQKIHGGQQ
SbiUMK4    152  TIRKRFEVFQESTLPVVQYYEKRGKLIKIDGAKSPDAVFED-VKAIFSQLN | TTQENQGSS
SitUMK4    152  TIRKRFQVFQESTLPVVQYYEKRGKLRRVDGAKSADEVFED-VKAIFVQLN | TQVNQGSSV
StuUMK4    151  TIRKRFVMESTLPTIEYYESKGKIRKVDAGKSVDEVFES-IKVIFSQGK | DNKVPPSRH
SlyUMK4    151  TIRKRFVMESTLPTIQYYESKGKLRKVDAGKSIDEVFES-IKVIFSPGK | DNKMPPSKH
AcmUMK4    155  AVRKRFKVFSESTLPVIEYYERKGKVRKVDAAKPIDEVFND-VKAIFAPYK | ANFVSSSNF
EsaUMK1    155  TIKKRFKVFVESTLPIVSYYESKGKLRKISAAKPSAEVFEA-VRDIFASET | GEGEARDHL
AthUMK1    155  TIKKRFKVFVESTLPIISYYESKGKLRKINAAKSSEEVFEA-VRVLFASET | ---------
BstUMK1    155  TIKKRFKVFVESTLPIVSYYESKGKLRTINAAKPSEEVFEA-VKVLFASET | ---------
CruUMK1    155  TIKKRFKVYVESTLPIVSYYESKGKLRKINAAKPSEEVFEA-VKVLFTSET | ---------
OsaUMK3c   173  TIRKRLKVFVESSLPVIFYYESKGMVKKIDATKPAPEVFED-VKAIFTHMA | ---------
CsaUMK3b   156  TIRKRFKVFLESSLPVVQFYESIGKVHKIDAARPVEEVFES-VKAVFTSVN | EKDD-----
AcmUMK3    161  TIRKRFKVFVESSLPVIEYYESKGKVKKIDAAKPIPEVFED-VKAIFSSYG | AKKVNDVEA
BdiUMK3    161  TIRKRFKVFVESSLPVIYYDAKEKVKKIDAAKPIPEVFED-VKAIFAPYA | KAA------
OsaUMK3a   161  TIRKRFKVFVESSLPVIYYNAKDKVKKIDAAKPIPEVFED-VKAIFAPYA | PNALLSGVT
SbiUMK3a   159  TIRKRFKVFVESSLPVIYYSSKDKVKKIDAAKPIPEVFED-VKAIFAPYA | PKVE-----
SitUMK3b   159  TIRKRFKVFVESSLPVIFYYSSKDKVKKIDAAKPIPEVFED-VTAIFAPYA | PKV------
OsaUMK3b   158  TIRKRFKVFVESSLPVIDHYNAKNKVKKIDAAKPISEVFED-VKAIFAPYA | KVE------
ZmaUMK3b   159  TIKKRFKTFVDSILPVIEHYNSKDKVKKIDAAKPIPEVFED-VKAIFAPYS | LKAE-----
ZmaUMK3a   159  TIKKRFKTFVESSLPVIEHYNSRDKVKKIDAAKPIPEVFED-VKTIFAPYA | PKAE-----
SbiUMK3b   160  TIKKRFKTFVESSLPVIEHYNSRDKVKKIDAAKPIPEVFED-VKAIFAPFS | SKAE-----
SitUMK3a   159  TIKKRFKTFVESSLPVIEHYNSRDKVKKIDAAKPIPEVFED-VKAIFAPYA | PKAE-----
CsaUMK3a   156  TIRKRFKVFLESSIPVIQYYESKEKVRKIDAARPVEEVFES-VKAVFTPKS | AKAE-----
MguUMK3    158  TIRKRFKVYMESSLPVIYYNAKGKVRKVDAARPIEEVFEA-VKSVFTPPC | DKVAAA---
StuUMK3    155  TIRKRFNVYMESSLPVIYYNSKGKVRKIDAVKPVGEVFEA-VKAVFTPAN | EKVAA----
SlyUMK3    155  TIKKRFNVYMESSLPVIEHYNSKGKVRKIDAVKPVGEVFEA-VKAVFAPSN | EKVAA----
VviUMK3    156  TIRKRFKVFLESSLPVIQYYESKGKVRKIDAAQSIEEVFEA-VKAVFTPTN | EQVDA----
TcaUMK3b   154  TIRKRFKVFEESLPVIYYKARGKVREIDAKSIEEVFDA-LKVIFTTTV | GKVVTT---
AcoUMK3    156  TIKKRFNVFVESSLPVIYYGSKGKVQKIDACKPVEEVFAA-VKAIFTQVH | EKTA-----
MesUMK3    155  TIRKRFKVFLESSMPVIYYGSKGKVRKIDASKPVEEVFEA-VKAIFTQKD | EKAAV----
EsaUMK3    155  TIRKRFKVFLESSLPVIQYYEAKGKVRKIHAAKPIEDVFQE-VKAVFSPDA | EKVEA----
CruUMK3    151  TIRKRFKVFLESSLPVIKYYEAKGKVRKINAAKPIEAVFE-VKTIFSPEA | EKVEA----
BstUMK3    151  TIRKRFKVFLESSLPVIQYYEAKGKVRKINAAKPIEAVFEA-VKAVFSPEA | DKVEA----
AthUMK3    148  TIRKRFKVFLDSSLPVINYYDAKGKVRKVDAARPVEEVFES-VKAIFGPKN | EKAD-----
MtrUMK3    155  TIRKRFKVFLDSSLPVINYYDAKGKVRKVDAARPVEEVFES-VKAIFGPKN | EKAD-----
PvuUMK3    154  TIRKRFKVFLESSLPVINYYDAKGKVRKIDAARPVEEVFET-VKGIFAPKT | EKAE-----
TcaUMK3a   154  TIRKRFNVFLESSLPVIQYYKAKGKVREIDAAKPIEEVFEA-VKVVFTPKG | EKVTA----
RcoUMK3    156  TIRKRFKVFLESSIPVVEYYESKGKVRKIDAAKSIEEVFED-VKAIFAQKD | EKDSV----
CclUMK3    155  TIRKRFKVFLESSLPVVQYYEAKGKVRKIDAAKPVAEVFDA-VKAVFTPKD | EKVKHYSCT
PtrUMK3a   145  TIRKRFKVFLESSLPVVEYYDSKGKVQKVDAAKPIDEVFEV-VKAIFTPKD | EKVAV----
PtrUMK3b   145  TIRKRFNVFLESSLPVVEYYDSKGKVRKVDAAKSVEEVFEA-VKAIFTPKE | EKVAV----
```

▲152          ▲182     ▲194(P)

```
AcmUMK2      239  LYAFVQTPFSSIAVDNWRSLRTWQRLPISDSFPAVPQVSFD
OsaUMK2           -----------------------------------------
BdiUMK2           -----------------------------------------
SitUMK2           -----------------------------------------
SbiUMK2           -----------------------------------------
ZmaUMK2           -----------------------------------------
StuUMK2           -----------------------------------------
SlyUMK2           -----------------------------------------
PvuUMK2           -----------------------------------------
MtrUMK2a          -----------------------------------------
MtrUMK2b          -----------------------------------------
MguUMK2           -----------------------------------------
CsaUMK2           -----------------------------------------
PtrUMK2           -----------------------------------------
VviUMK2           -----------------------------------------
AcoUMK2           -----------------------------------------
EsaUMK2      246  SKSPLGSNLVEN-----------------------------
AthUMK2      246  PQSHLGSSLVENSS---------------------------
BstUMK2      246  PKSTLGSSLVENSS---------------------------
CruUMK2      248  IQKSTIGSSLVENSS--------------------------
RcoUMK2           -----------------------------------------
CclUMK2           -----------------------------------------
TcaUMK2           -----------------------------------------
MesUMK2           -----------------------------------------
PtrUMK4      225  -----------------------------------------
MesUMK4      357  ECNILQIGAQVSSII--------------------------
ZmaUMK4      206  IYLPFFFPIDCSLLIKP------------------------
BdiUMK4      211  SRAQTNPFKRWFLDLCCGCFDAQERRN--------------
OsaUMK4      212  ASGLSRAQMNPLKRWFFDFFCGCFGTKEEARN---------
SbiUMK4      211  MSSRVQSNPLKRFLDLLCGCFGTQEARS-------------
SitUMK4      211  SRAQSNPLKRFVDLFCGCFGTQEETN---------------
StuUMK4      210  KCKCLIL----------------------------------
SlyUMK4      210  KCKCLIL----------------------------------
AcmUMK4      214  NAEAGVEHRMCPALSKRLARCARKTKALFRKRAPV------
EsaUMK1      214  NKKISV-----------------------------------
AthUMK1           -----------------------------------------
BstUMK1           -----------------------------------------
CruUMK1           -----------------------------------------
OsaUMK3c          -----------------------------------------
CsaUMK3b          -----------------------------------------
AcmUMK3      220  EVFSR------------------------------------
BdiUMK3           -----------------------------------------
OsaUMK3a     220  TNL--------------------------------------
SbiUMK3a          -----------------------------------------
SitUMK3b          -----------------------------------------
OsaUMK3b          -----------------------------------------
ZmaUMK3b          -----------------------------------------
ZmaUMK3a          -----------------------------------------
SbiUMK3b          -----------------------------------------
SitUMK3a          -----------------------------------------
CsaUMK3a          -----------------------------------------
MguUMK3           -----------------------------------------
StuUMK3           -----------------------------------------
SlyUMK3           -----------------------------------------
VviUMK3           -----------------------------------------
TcaUMK3b          -----------------------------------------
AcoUMK3           -----------------------------------------
MesUMK3           -----------------------------------------
EsaUMK3           -----------------------------------------
CruUMK3           -----------------------------------------
BstUMK3           -----------------------------------------
AthUMK3           -----------------------------------------
MtrUMK3           -----------------------------------------
PvuUMK3           -----------------------------------------
TcaUMK3a          -----------------------------------------
RcoUMK3           -----------------------------------------
CclUMK3      214  IL---------------------------------------
PtrUMK3a          -----------------------------------------
PtrUMK3b          -----------------------------------------
```

**Figure A 1. Multiple alignment of UMK sequences from 23 vascular plants.**
The alignment was generated with MUSCLE (www.ebi.ac.uk) and shaded with pyBoxshade (github.com/mdbaron42/pyBoxshade). Red lines enclose the part of the alignment that was used for the construction of the phylogenetic tree. Amino acids with side chains involved in UMP or CMP binding via H bonds according to Schlichting and Reinstein, 1997 are shaded in red. Pink shading was used for amino acids lining the active site pocket involved in hydrophobic contacts with the monophosphate substrate. Amino acids with green shading are involved in trinucleotide (ATP) binding. Glutamate 76 of UMK3 from Arabidopsis missing in the UMK3 variant encoded by the mutant allele *umk3$_{\Delta E76}$* is shaded in yellow. The start of the N-terminal truncated UMK encoded at At3g60961 is indicated under the alignment (the sequence itself is not shown). Phosphorylation sites in AtUMK3 annotated in PhosPhat4.0 (phosphat.uni-hohenheim.de/) and by Mergner et al., 2020 are marked with black triangles. Phosphorylation sites in AtUMK1 are marked with grey triangles. Residue numbers at marked positions are corresponding to UMK3 from Arabidopsis.



**Figure A 2. Plants grown from a segregating GK723G02 seed batch.**
Red arrows indicate plants carrying a homozygous T-DNA insertion in the *UMK2* locus. The first true leaves of the indicated plants are yellow, which resembles the phenotype observed for a null mutation in the *UMK2*-neighboring locus At4g25270. This phenotype could not be complemented by expression of a *UMK2* transgene.
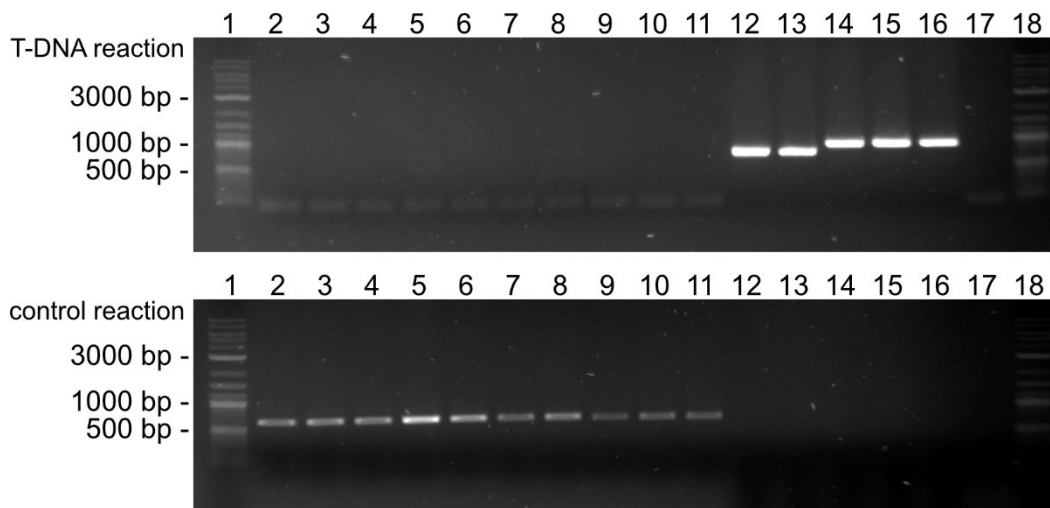
**Figure A 3. PCR products visualized on an agarose gel showing the absence of the transgene in *UMK* mutant lines.**

To show the absence of the T-DNAs used to induce the mutations in the *UMK* mutant lines, a T-DNA specific PCR with the primers P1164 and P1165 which amplify a T-DNA fragment of 779 or 932 bp depending on the vector was performed (upper panel). As DNA quality control, a 635 bp wild-type genomic DNA fragment was amplified from the same DNA preparations with P1686 and P1687 (lower panel). Lanes 1 and 18, DNA marker. Lanes 2-11, PCR products amplified from genomic DNA of the five single and three double mutants and the one triple mutant, as well as a wild type obtained from one of the segregating populations. Lanes 12-16, positive controls amplified using 1 ng of DNA from the constructs used for transformation (construct numbers H773, H774; H1133, H1135 and H983). Lane 17, negative control without DNA.
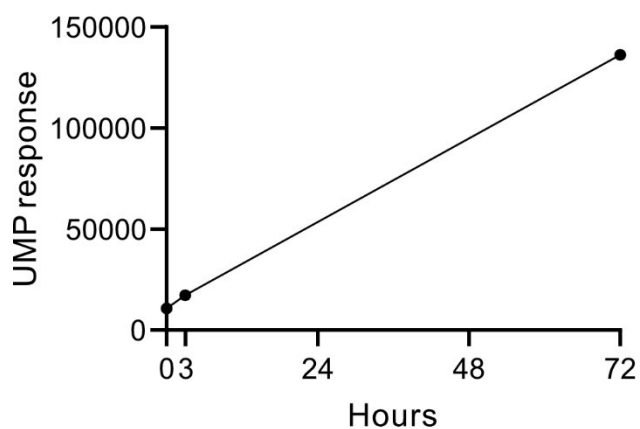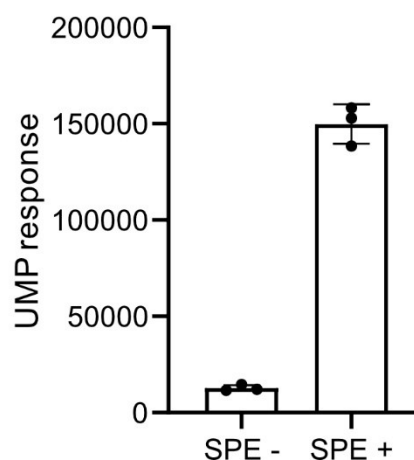


**Figure A 4. UDP-glucose-decay release UMP in buffer and during solid phase extraction.**
**A)** UDP-glucose was diluted to 1 mM in mobile phase A (5 mM ammonium acetate, pH 9.5) and UMP content quantified via LC-MS directly, after 3 hours and after 3 days. **B)** UMP content quantified from three solid phase extraction samples of Arabidopsis wild type leaf material. SPE -, no added UDP-glucose; SPE +, 100 nmol of UDP-glucose was added to the samples before starting the extraction.
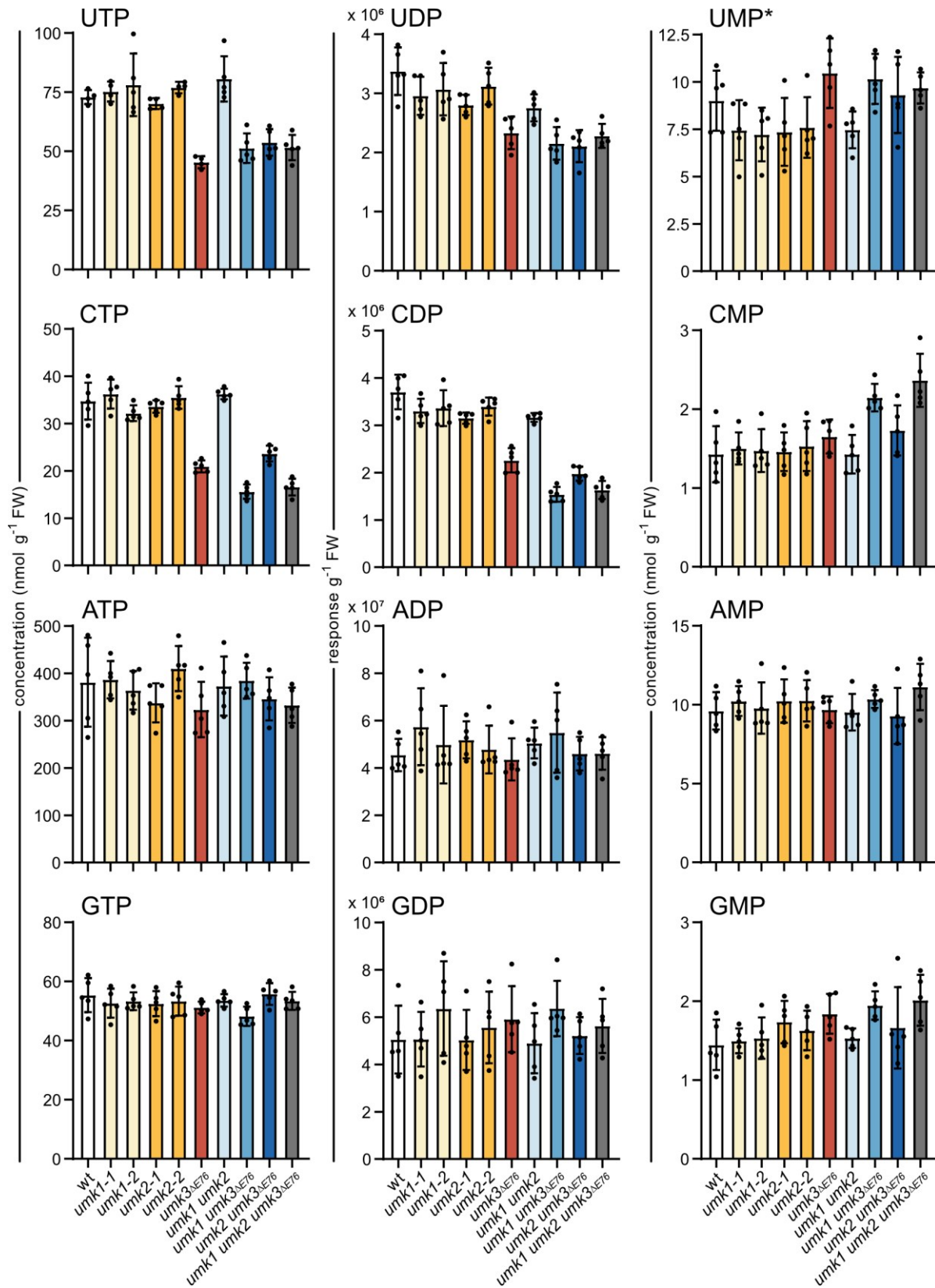
**Figure A 5. All quantified ribonucleotides from 18-day-old plants.**
Extended data from **Figure 20** and **Figure 21**. Methodology is explained in **Figure 20**.
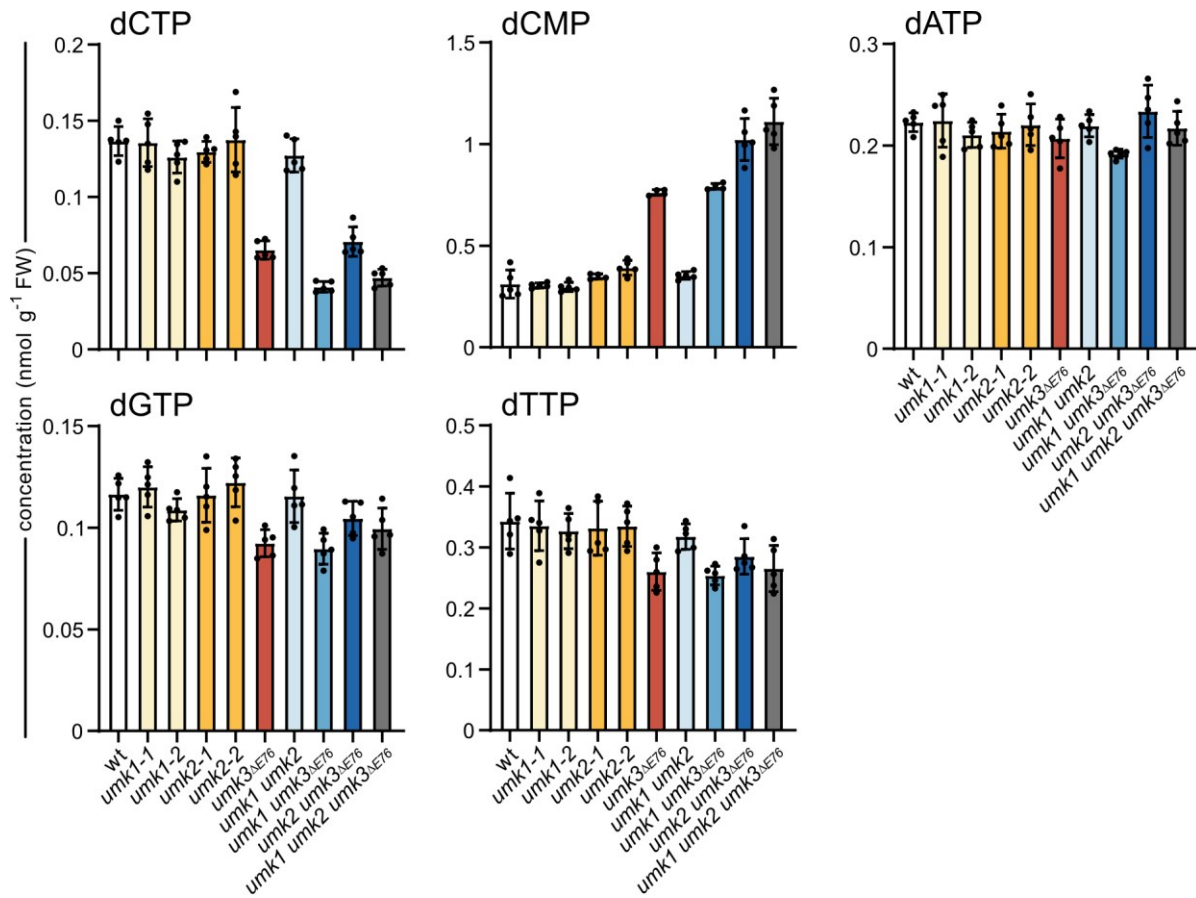
**Figure A 6. All quantified deoxynucleotides from 18-day-old plants.**
Extended data from **Figure 20** and **Figure 21**. Methodology is explained in **Figure 20**.

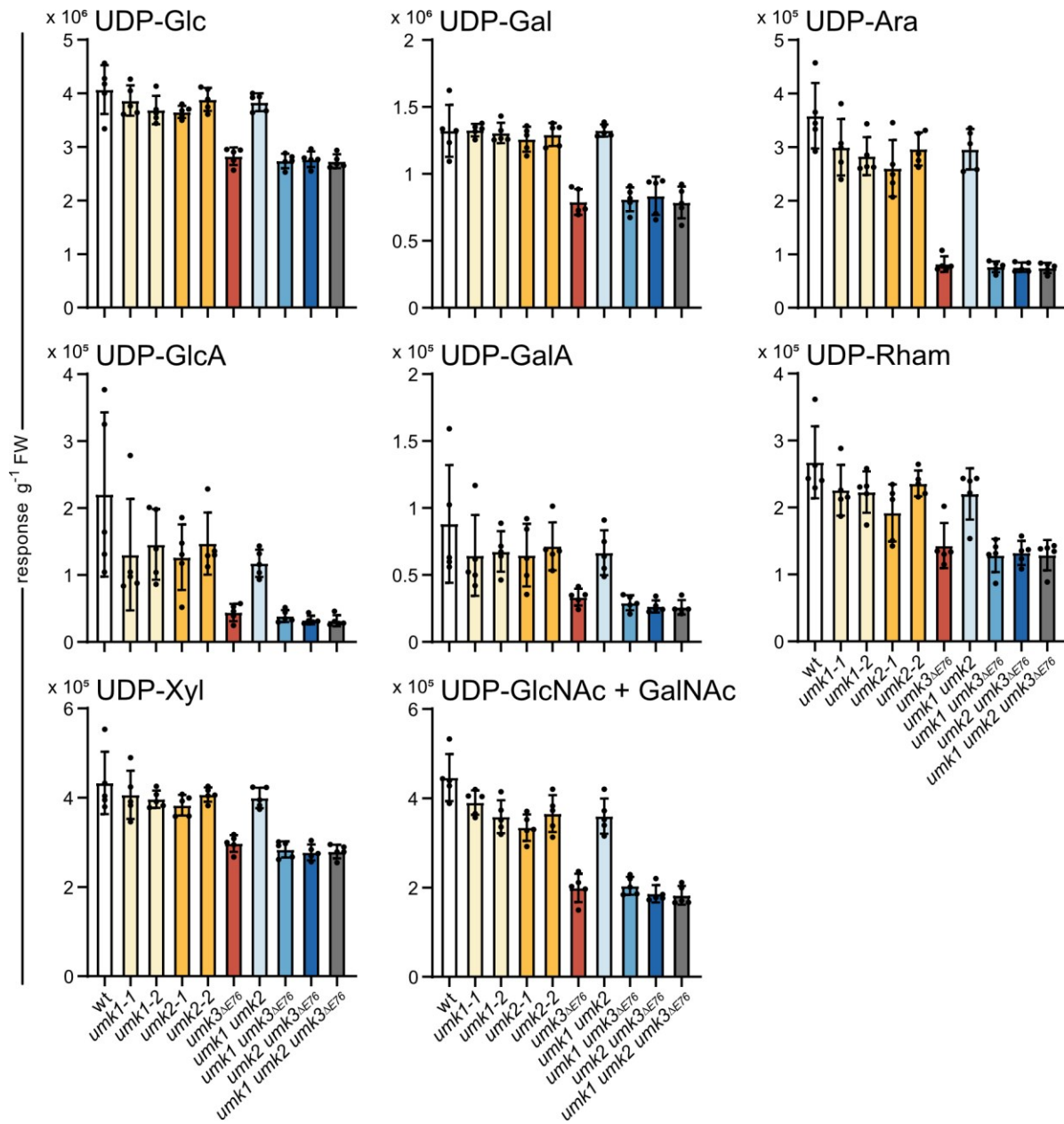**Figure A 7. All quantified UDP-sugars from 18-day-old plants.**
Extended data from **Figure 22**. Methodology is explained in **Figure 20**.

GAATTCAAAATGGTGTCTAAAGGAGAGGAGGATAACATGGCATCATTACCTGCTACCCACGAGTTGCACAT
TTTCGGAAGTATTAACGGAGTGGATTTCGATATGGTTGGACAAGGTACTGGAAATCCTAACGATGGTTATG
AAGAGCTTAACTTGAAGTCTACAAAAGGAGATCTTCAGTTTTCTCCTTGGATTTTGGTTCCACATATCGGT
TACGGATTCCATCAATACCTTCCTTATCCAGATGGAATGTCTCCATTTCAGGCTGCTATGGTTGATGGTTC
TGGATATCAAGTTCATAGAACTATGCAGTTCGAAGATGGAGCTTCTTTGACTGTTAACTACAGATACACAT
ACGAAG<mark>GTCAATATCTGTTCTAAATTTTGCATACTCCTTCAATTTTATGCGACATTTTTTTCCTTCTATTT
GTTCCAAAAAAAAAAGAAGAAGATGACGACATATTTATATATTTAGAAAAAATTTAACTTTTAACTTTAAT
ATGTTATTTTGCATGTGCAG</mark>GTTCTCATATTAAGGGAGAGGCTCAAGTTAAAGGTACTGGATTTCCTGCTG
ATGGTCCAGTTATGACTAATTCTCTTACAGCTGCTGATTGGTGTAGATCTAAGAAAACATACCCTAACGAT
AAGACTATTATCTCTACATTCAAATGGTCTTACACTACTGGTAACGGAAAGAGATATAGATCTACTGCTAG
AACTACATACACATTTGCTAAGCCTATGGCTGCTAATTACTTGAAAAACCAGCCAATGTATGTTTTCAGAA
AGACTGAACTTAAGCATTCTAAAACAGAGTTGAACTTTAAAGAATGGCAGAAGGCTTTCACCGATGTTATG
GGTATGGATGAGCTTTACAAATCAAAACTCTAACCCGGG

**Figure A 8. Nucleotide sequence of *mNeonGreen* gene.**
Intron is highlighted in yellow.

**Table A 1. Species names, protein abbreviations and corresponding locus identifiers in Phytozome V12.1.**

| species | abbreviation | locus identifier | remark |
|---|---|---|---|
| Aquilegia coerulea[1] | AcoUMK2 | Aqcoe7G075500.1 | |
| | AcoUMK3 | Aqcoe3G252000.1 | |
| Ananas comosus | AcmUMK2 | Aco014070.1 | |
| | AcmUMK3 | Aco025395.1 | |
| | AcmUMK4 | Aco017481.1 | |
| Arabidopsis thaliana | AthUMK1 | At3g60180.1 | |
| | AthUMK2 | At4g25280.1 | |
| | AthUMK3 | At5g26667.2 | |
| Brachypodium distachyon | BdiUMK2 | Bradi5g02200.1 | |
| | BdiUMK3 | XP_003562640.1[2] | |
| | BdiUMK4 | Bradi1g51830.2 | |
| Bochera stricta | BstUMK1 | Bostr.13158s0003.1 | |
| | BstUMK2 | Bostr.7867s0335.1 | |
| | BstUMK3 | Bostr.29827s0177.1 | |
| Capsella rubella | CruUMK1 | Carubv10019180m | N-terminus corrected by hand according to consensus |
| | CruUMK2 | Carubv10005525m | |
| | CruUMK3 | Carubv10001985m | |
| Citrus clementina[1] | CclUMK2 | Ciclev10005772m | |
| | CclUMK3 | Ciclev10024291m | |
| Cucumis sativus | CsaUMK2 | Cucsa.251530.1 | |
| | CsaUMK3a | Cucsa.257020.1 | |
| | CsaUMK3b | Cucsa.362400.1 | |
| Eutrema salsugineum | EsaUMK1 | Thhalv10006230m | |
| | EsaUMK2 | Thhalv10026032m | |
| | EsaUMK3 | Thhalv10004914m | |
| Manihot esculenta[1] | MesUMK2 | Manes.14G034500.1 | |
| | MesUMK3 | Manes.05G059600.1 | |
| | MesUMK4 | Manes.14G090800.1 | possible pseudogene[3] |
| Medicago truncatula | MtrUMK2a | Medtr4g035850.1 | |
| | MtrUMK2b | Medtr5g068940.1 | |
| | MtrUMK3 | Medtr8g009520.1 | |
| Mimulus guttatus[1] | MguUMK2 | Migut.F00456.2 | |
| | MguUMK3 | Migut.B01476.1 | |
| Oryza sativa | OsaUMK2 | LOC_Os04g01530.2 | |
| | OsaUMK3a | LOC_Os02g53790.1 | |
| | OsaUMK3b | LOC_Os07g43170.3 | |
| | OsaUMK3c | LOC_Os06g10200.2 | possible pseudogene[3] |
| | OsaUMK4 | LOC_Os06g02000.1 | |
| Phaseolus vulgaris | PvuUMK2 | Phvul.011G004800.1 | |
| | PvuUMK3 | Phvul.010G040800.1 | |

| | | | |
|---|---|---|---|
| Populus trichocarpa | PtrUMK2 | Potri.015G129000.1 | |
| | PtrUMK3a | Potri.014G043300.2 | |
| | PtrUMK3b | Potri.002G134600.1 | |
| | PtrUMK4 | Potri.014G104700.1 | possible pseudogene[3] |
| Ricinus communis[1] | RcoUMK2 | 30147.m014199 | |
| | RcoUMK3 | 29709.m001213 | |
| Theobroma cacao | TcaUMK2 | Thecc1EG014953t1 | last short exon is missing from annotation, but not relevant for tree |
| | TcaUMK3a | Thecc1EG042355t1 | |
| | TcaUMK3b | Thecc1EG043327t1 | |
| Setaria italica | SitUMK2 | Seita.7G006900.1 | |
| | SitUMK3a | Seita.1G339500.1 | |
| | SitUMK3b | Seita.6G192300.1 | |
| | SitUMK4 | Seita.4G007200.1 | |
| Solanum lycopersicum | SlyUMK2 | Solyc03g083610.2.1 | |
| | SlyUMK3 | Solyc01g088480.2.1 | |
| | SlyUMK4 | Solyc08g077300.2.1 | |
| Solanum tuberosum | StuUMK2 | PGSC0003DMT400049087 | |
| | StuUMK3 | PGSC0003DMT400004310 | |
| | StuUMK4 | PGSC0003DMT400062219 | |
| Sorghum bicolor | SbiUMK2 | Sobic.006G003300.1 | |
| | SbiUMK3a | Sobic.007G223100.1 | |
| | SbiUMK3b | Sobic.004G318100.1 | |
| | SbiUMK4 | Sobic.010G007500.1 | |
| Vitis vinifera[1] | VviUMK2 | GSVIVT01018596001 | |
| | VviUMK3 | GSVIVT01026782001 | |
| Zea mays | ZmaUMK2 | GRMZM2G079944_T01 | |
| | ZmaUMK3a | GRMZM5G801436_T01 | |
| | ZmaUMK3b | GRMZM2G149281_T01 | |
| | ZmaUMK4 | GRMZM2G141009_T01 | |

[1] Sequences with several strong deviations from highly conserved consensus were classified as possible pseudogenes and not included in the analysis. These are: Ciclev10013384m (*Citrus clementina*), 30074.m001348 (*Ricinus communis*), Manes.02G200000.1 (*Manihot esculenta*), Migut.N01708.1 (*Mimulus guttatus*), GSVIVT01027430001 (*Vitis vinifera*), Aqcoe6G240400.1 (*Aquilegia coerule*a), LOC_Os06g10200.3 (*Oryza sativa*), Aco019022.1 (*Ananas comosus*).

[2] Genebank accession number, because this gene is missing in the Phytozome annotation

[3] Proteins have several amino acid deviations from consensus at highly conserved positions

**Table A 2. List of primers.**

| Name | Sequence |
|---|---|
| P272 | TAGGTCTCCAAACGAAGACAAAAACAAAAAAAAAAGCACCGACTCG |
| P274 | TAGGTCTCCAAACGAAGACAAAAAC |
| P293 | CGGGTCTCAGGCAGAAGACTAATTGAACAAAGCACCAGTGG |
| P294 | CGGGTCTCAGGCAGAAGACTAATTG |
| P1099 | TCCCGGGATGGTGTCTAAAGGAGAG |
| P1100 | ATCTAGATTATTTGTAAAGCTCATCC |
| P1164 | AAAGCTGCAAATGTTACTGA |
| P1165 | GGCAACCTCGCATGAAAATAGTA |
| P1379 | TGAATTCAAAATGGAAACTCCTATCGATGCTC |
| P1380 | ACCCGGGAGTTTCAGATGCAAATAGAACTC |
| P1381 | TGAATTCAAAATGTGGAGACGCGTGG |
| P1382 | ACCCGGGAGATGAATTTTCTACCAAACTCG |
| P1383 | TATCGATAAAATGGGATCTGTTGATGCTG |
| P1384 | ACCCGGGGGCTTCAACCTAAATAAACGATC |
| P1577 | TTCGTTCTCTTTGGGAAATTAGA |
| P1578 | CTCGCTGTACCTCTTTGTATTCTTT |
| P1581 | GTAGCTGCGGTGAAGTAGGC |
| P1582 | CTGCCTGGATTCGGTATCAT |
| P1653 | TAGGTCTCCAGGAAGTGGAAAGTTTTAGAGCTAGAA |
| P1654 | ATGGTCTCATCCTGGACCACCTGCACCAGCCGGGAA |
| P1655 | TAGGTCTCCTTGTACCTTCTGGTTTTAGAGCTAGAA |
| P1656 | ATGGTCTCAACAATCTTCCCCTGCACCAGCCGGGAA |
| P1686 | TGGCGCGCCTCGACGAGTCAGTAATAAACG |
| P1687 | ACTCGAGCTGTTAATCAGAAAAACTCAGATTA |
| P1831 | ATTGACGAACATGAATGCCCTAGG |
| P1832 | AAACCCTAGGGCATTCATGTTCGT |
| P1833 | ATTGACCTTATTCGGAGCATCGAT |
| P1834 | AAACATCGATGCTCCGAATAAGGT |
| P1874 | AGGCGCGCCTCTTTTCTTCTGCTTTTAATAAATTTG |
| P1875 | ACTCAGAATTCACCTACAATAAG |
| P1942 | GTAAAACGACGGCCAGTTTCGCCTGGACTTGTCAAAC |
| P1943 | CCACTTCCAGGACCACC |
| P1946 | GTAAAACGACGGCCAGTCAGGTTCTGAAAATGGGTATGC |
| P1947 | CGAGGGAAACCATCAATGAGG |
| P2188 | TGGCGCGCCATCTCCTTAATTCGGTTGCTG |
| P2451 | GAGGCTATGCGCTCAAATGCAACA |
| P2452 | AAACTGTTGCATTTGAGCGCATAG |
| P2453 | GAGGAATGGAGCTTTCTCTTTCGG |
| P2454 | AAACCCGAAAGAGAAAGCTCCATT |
| P2455 | GAGGACACGGTTCTCCTCAGTTCG |
| P2456 | AAACCGAACTGAGGAGAACCGTGT |
| P2457 | GAGGTCTAACCTGGCTGAGCTTAA |
| P2458 | AAACTTAAGCTCAGCCAGGTTAGA |
| P2459 | GAGGGCGATCCCTTAAGCTCAGCC |
| P2460 | AAACGGCTGAGCTTAAGGGATCGC |
| P2461 | GAGGGCAGCTTCTGGGCTCAAAGT |

| | |
|---|---|
| P2462 | AAACACTTTGAGCCCAGAAGCTGC |
| P2465 | ACTCGAGGGTTATGGAAACGAAGAGAGAAG |
| P2471 | TAGGTCTCCTCAAATGCAACAGTTTTAGAGCTAGAA |
| P2472 | ATGGTCTCATTGAGCGCATAGTGCACCAGCCGGGAA |
| P2473 | TAGGTCTCCTTTCTCTTTCGGGTTTTAGAGCTAGAA |
| P2474 | ATGGTCTCAGAAAGCTCCATTTGCACCAGCCGGGAA |
| P2475 | TAGGTCTCCCTCCTCAGTTCGGTTTTAGAGCTAGAA |
| P2476 | ATGGTCTCAGGAGAACCGTGTTGCACCAGCCGGGAA |
| P2477 | TAGGTCTCCGGCTGAGCTTAAGTTTTAGAGCTAGAA |
| P2478 | ATGGTCTCAAGCCAGGTTAGATGCACCAGCCGGGAA |
| P2479 | TAGGTCTCCTTAAGCTCAGCCGTTTTAGAGCTAGAA |
| P2480 | ATGGTCTCATTAAGGGATCGCTGCACCAGCCGGGAA |
| P2481 | TAGGTCTCCTGGGCTCAAAGTGTTTTAGAGCTAGAA |
| P2482 | ATGGTCTCACCCAGAAGCTGCTGCACCAGCCGGGAA |
| P2565 | GTAAAACGACGGCCAGTGAAGATGTGGAGACGCGTG |
| P2566 | CAAACTATTGAGCTTTTCCTGATC |
| P2567 | GTAAAACGACGGCCAGTGGGGAATCTTTTGCAACAGAC |
| P2568 | GCCCATGACTGAAGAACAAGTG |
| P2569 | GTAAAACGACGGCCAGTGATGGGAAGATTGTTCCTTCAG |
| P2570 | GTACTACATCAGGGTCTGCTC |

Appendix

**Calculation A 1. Calculation of cytosolic UMP-concentration.**

Formula from Straube et al., 2021: $\dfrac{X\left[\frac{pmol}{g}\right]}{AMV\left[\frac{\mu L}{g}\right]} \times \dfrac{100\,\%}{Z\,\%} = Y\,[\mu M]$

X is the average amount of UMP from the metabolome analysis of the wild type: 9010 pmol/g FW

AMV is the average leaf mesophyll volume per unit total fresh weight (Straube et al., 2021): 600 µL g$^{-1}$

Z is the average relative cytosol volume of young and old leaves from Koffler et al., 2013: 5.43 %

Inserting these values into the formula: $\dfrac{9010\,\frac{pmol}{g}}{600\,\frac{\mu L}{g}} \times \dfrac{100\,\%}{5.43\,\%} = 276.6\,\mu M$

# 12. Acknowledgements

# 13. Curriculum vitae

| | | |
|---|---|---|
| **Name** | Jannis Rinne | |
| **Date of birth** | 18th September 1992 | |
| **Place of birth** | Hameln | |
| **Education** | 2018 – present: | Doctoral candidate, Institute of Plant Nutrition |
| | | Leibniz Universität Hannover |
| | 2015 – 2018 | Master of Science Pflanzenbiotechnologie |
| | | Leibniz Universität Hannover |
| | 2012 – 2015 | Bachelor of Science Pflanzenbiotechnologie |
| | | Leibniz Universität Hannover |
| | 2004 – 2011 | General university entrance qualification |
| | | Viktoria-Luise-Gymnasium Hameln |

**Publications**

**Jannis Rinne**, Claus-Peter Witte, Marco Herde; Loss of MAR1 function is a marker for co-selection of CRISPR-induced mutations in plants; Frontiers in Genome Editing; 2021.

Henryk Straube, Jannis Straube, **Jannis Rinne**, Lisa Fischer, Markus Niehaus, Claus-Peter Witte, Marco Herde; An inosine triphosphate pyrophosphatase safeguards plant nucleic acids from aberrant purine nucleotides; New Phytologist; 2023.

Vanessa Scherer, Leo Bellin, Serena Schwenkert, Martin Lehmann, **Jannis Rinne**, Claus-Peter Witte, Kathrin Jahnke, Andreas S Richter, Tobias Pruss, Anne Sophie Lau, Dario Leister, Torsten Möhlmann; UPP affects chloroplast development by interfering with chloroplast proteostasis; Plant Physiology (under review); 2023.

**Jannis Rinne**, Markus Niehaus, Nieves Medina-Escobar, Henryk Straube, Frank Schaarschmidt, Nils Rugen, Hans-Peter Braun, Marco Herde, Claus-Peter Witte; Three UMP kinases of *Arabidopsis thaliana* have different roles in pyrimidine nucleotide synthesis and (deoxy)CMP salvage; Plant Cell (under review); 2023

**Conference contributions**

**Jannis Rinne**, Marina Varbanova-Herde, Marco Herde, Claus-Peter Witte; CRISPRselect – Development of knockout markers for selection of CRISPR/Cas9-induced mutations; Poster at PLANT 2030 Status Seminar; March 2019; Potsdam.

**Jannis Rinne**, Marina Varbanova-Herde, Marco Herde, Claus-Peter Witte; CRISPRselect – Selection of CRISPR/Cas9-induced mutations via the endogenous gene Mar1; Talk at Keystone Symposia - Plant Genome Engineering: From Lab to Field; April 2020; Breckenridge, Colorado, USA (cancelled due to Corona).

**Jannis Rinne**, Claus-Peter Witte; Pyrimidine catabolism and possibly polyamine oxidation are sources for beta-alanine, which is transaminated in mitochondria and peroxisomes; Poster at Botanik-Tagung – International Conference of the German Society for Plant Sciences; August 2022; Bonn.

**Jannis Rinne**, Claus-Peter Witte; Molecular recycling in plant metabolism: beta-alanine is metabolic waste and precursor in one; Talk at Meeting of Molecular Biosciences at Campus Herrenhausen; March 2023.