

Review

Functional Gene Network of Prenyltransferases in *Arabidopsis thaliana*

Diana Kopcsayová^{1,2} and Eva Vranová^{1,*}

¹ Institute of Biology and Ecology, Pavol Jozef Šafárik University in Košice, 04180 Košice, Slovakia; diana.kopcsayova@gmail.com

² Institute of Pharmacology, University of Veterinary Medicine and Pharmacy in Košice, 04181 Košice, Slovakia

* Correspondence: eva.vranova@upjs.sk

Academic Editors: Ewa Swiezewska and Liliana Surmacz

Received: 7 November 2019; Accepted: 10 December 2019; Published: 12 December 2019



Abstract: Prenyltransferases (PTs) are enzymes that catalyze prenyl chain elongation. Some are highly similar to each other at the amino acid level. Therefore, it is difficult to assign their function based solely on their sequence homology to functional orthologs. Other experiments, such as in vitro enzymatic assay, mutant analysis, and mutant complementation are necessary to assign their precise function. Moreover, subcellular localization can also influence the functionality of the enzymes within the pathway network, because different isoprenoid end products are synthesized in the cytosol, mitochondria, or plastids from prenyl diphosphate (prenyl-PP) substrates. In addition to in vivo functional experiments, in silico approaches, such as co-expression analysis, can provide information about the topology of PTs within the isoprenoid pathway network. There has been huge progress in the last few years in the characterization of individual *Arabidopsis* PTs, resulting in better understanding of their function and their topology within the isoprenoid pathway. Here, we summarize these findings and present the updated topological model of PTs in the *Arabidopsis thaliana* isoprenoid pathway.

Keywords: prenyltransferases; isoprenoid diphosphate synthases; product specificity; co-expression analysis; mutants; *Arabidopsis*

1. Introduction

Isoprenoids form a large group of more than 55,000 natural metabolites despite the fact that they are all derived from just five carbon units of isopentenyl diphosphate (IPP) [1]. IPP and its isomer dimethylallyl diphosphate (DMAPP) are synthesized in a highly conserved manner in almost all living species. In plants, they are produced in the cytosol via the mevalonic acid (MVA) pathway and in plastids through the MVA-independent methylerythritol phosphate (MEP) pathway. However, the vast majority of organisms possess just one of these pathways—the MVA pathway is more widely spread, utilized by archaeobacteria, some Gram-positive bacteria, yeasts, fungi, and animals, and the MEP pathway provides the isoprenoid precursors in bacteria, cyanobacteria, and green algae [2].

Some groups of isoprenoids, such as cytokinins and hemiterpenoids, are derived directly from the DMAPP, but the great structural and functional diversity of isoprenoids is created by the activity of prenyltransferases (PTs). In a broader sense, PTs are any enzymes that catalyze the transfer of the prenyl group [3]. Possible acceptors are both aliphatic and aromatic compounds with low or even high molecular weight including proteins. In a narrower sense, PTs catalyze condensation between DMAPP or another prenyl-PP, which serves as an allylic donor, and IPP, which serves as an acceptor [4] (Figure 1). These are also known in the literature as isoprenyl-PP synthases (IDSs) [5]. In this review, we will use the term PT for convenience, as PT is more frequently used in the literature, but we will

refer to only the prenyl chain elongating enzymes that catalyze sequential condensation of allylic diphosphate and IPP.

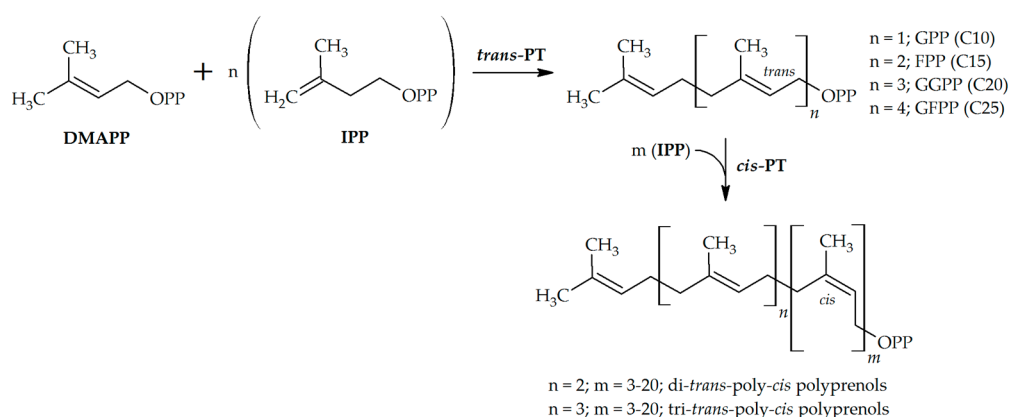


Figure 1. Schematic representation of the reactions catalyzed by *trans*-prenyltransferases and *cis*-prenyltransferases showing the stereochemistry of the double bond formed in the enzyme product. DMAPP, dimethylallyl diphosphate; IPP, isopentenyl diphosphate; PT, prenyltransferase; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; GFPP, geranylgeranyl diphosphate; and OPP, diphosphate ester.

The binding of both co-substrates to the enzyme in the presence of the divalent Mg salts creates carbocation on the allylic substrate, followed by the electrophilic attack on the double bond of the IPP. As a result, a new carbon bond joining the co-substrates is formed [6]. The product is then released from the active site or can act as a primer in the subsequent coupling with another IPP molecule. With every further elongation step, a linear prenyl-PP grows by C₅. The carbon chain length of the final isoprenoid polymer and its stereochemical properties are determined by the specificity of the PT, which is given to the PT by its structural features. The two main classes of PTs are recognized due to the differences in the spatial arrangement of units in the isoprenoid product, namely *trans*- or (*E*)-PTs and *cis*- or (*Z*)-PTs (Figure 1). Basically, shorter prenyl-PPs are formed via *trans*-condensation and polyprenyl-PPs via *cis*-condensation in plants [7]. Although *trans*- and *cis*-PTs catalyze similar reactions, they are evolutionarily distinct [7] (Figure 2).

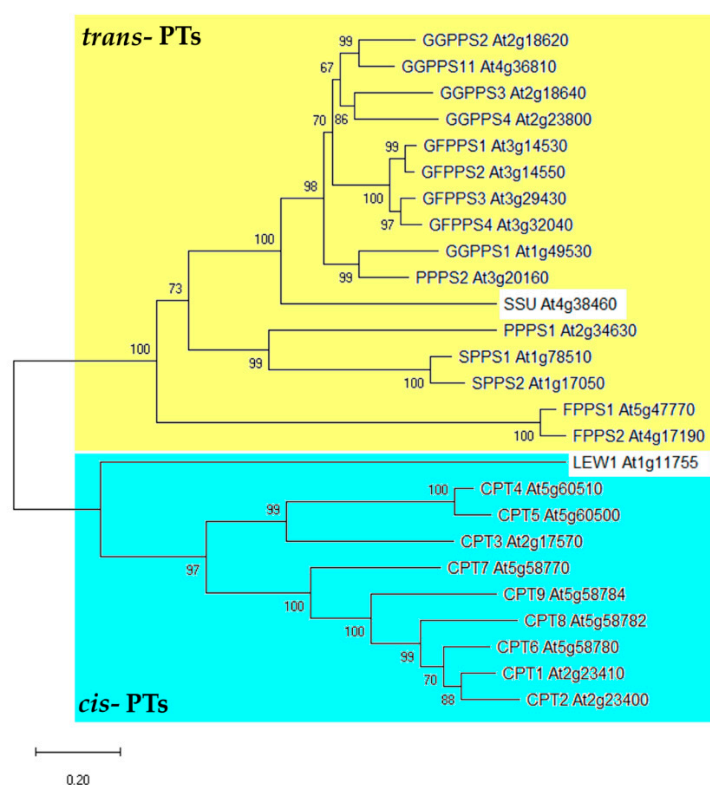


Figure 2. Phylogenetic analysis of the prenyltransferase family in *A. thaliana*. The phylogenetic tree illustrates the relatedness of PTs, which cluster into two major clades—*trans*- and *cis*-PTs. Enzyme abbreviations are listed in Table 1. Enzymes in white boxes are not PTs, but they interact with PTs—small subunit (SSU) with geranylgeranyl diphosphate synthase 11 (GGPPS11) [8,9] and LEW1 with *cis*-PT3 (CPT3) [10]. Because the tree was left unrooted, the branching order may not always reflect the evolutionary course that the enzymes have taken. The evolutionary analysis was conducted in MEGA X [11] using the neighbor-joining method with 500 bootstrap replicates. The scale bar represents 0.2 amino acid substitutions per site.

2. Prenyltransferases in *Arabidopsis thaliana*

IPP is a prenyl acceptor of the prenyl group originating from DMAPP, in the reaction catalyzed by PTs. IPP and DMAPP, both substrates of PTs, are interconvertible isomers, whose equilibrium is controlled by IPP isomerase (IPPI) [12]. The enzyme is encoded by *IPPI1* (*At3g02780*) and *IPPI2* (*At5g16440*) genes in *A. thaliana*. Both are transcribed into two different protein variants—short with peroxisomal targeting (IPPI1S, IPPI2S), supplying IPP for the cytosolic PTs, and long with either mitochondrial (IPPI1L) or plastidial targeting (IPPI2L), providing IPP for PTs in respective organelles [13–15] (Figure 3). In agreement with that, PT activity at the level of a plant cell is also compartmentalized [16].

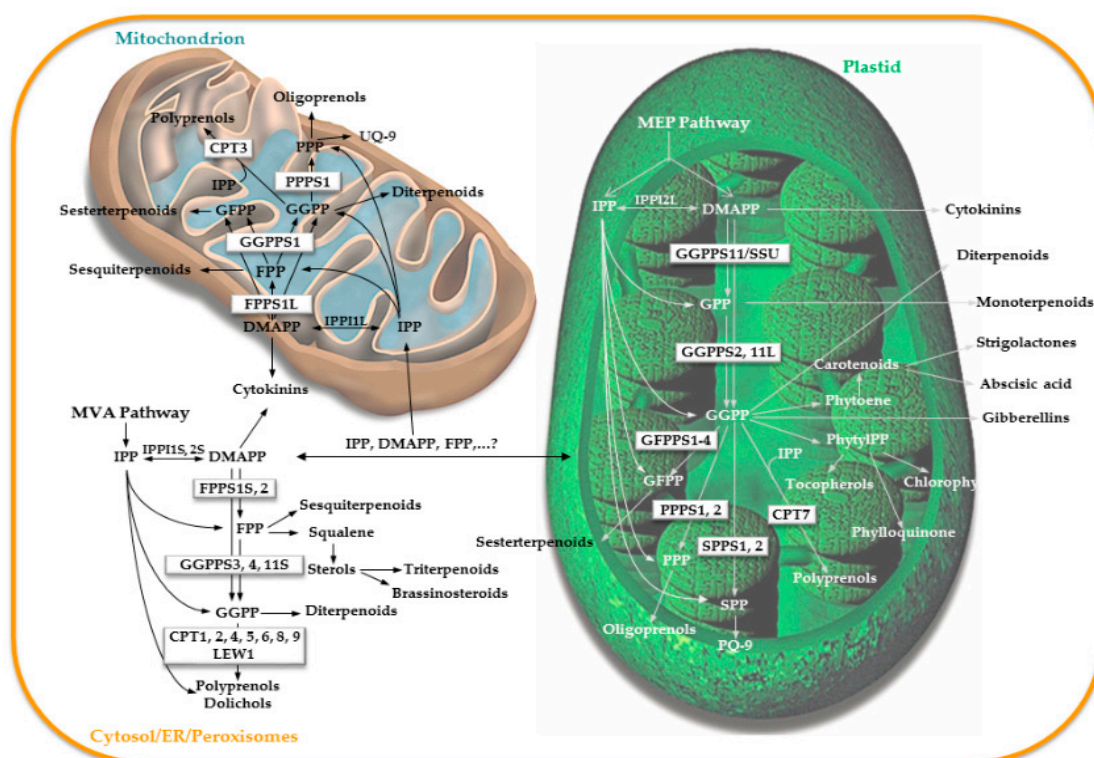


Figure 3. Topology of the prenyltransferase genes within the isoprenoid pathway of *A. thaliana*. Based on the isoprenoid pathway network constructed by Vranová et al. [16]. Abbreviations of the enzymes in boxes are given in Table 1. MVA, mevalonate; MEP, 2-C-methyl-D-erythritol 4-phosphate; IPPI, isopentenyl diphosphate (IPP, C5) isomerase; IPPI1S, 2S, short IPPI variants; IPPI1L, 2L, long IPPI variants; DMAPP (C5), dimethylallyl diphosphate; GPP (C10), geranyl diphosphate; FPP (C15), farnesyl diphosphate; GGPP (C20), geranylgeranyl diphosphate; GFPP (C25), geranylfarnesyl diphosphate; PPP (C ≥ 30), polyprenyl diphosphate; SPP (C45), solanesyl diphosphate; PQ-9, plastoquinone 9; and UQ-9, ubiquinone 9.

2.1. *Trans*-Prenyltransferases

C–C double bonds in prenyl-PP can be either in *trans* or *cis* conformation depending on the PT—*trans*-PTs form *trans* isomers, and *cis*-PTs form *cis*-isomers. Considering the length of the produced carbon backbone, *trans*-PTs can be classified as short-chain PTs (C_{10–25}), medium-chain PTs (C_{30–35}), and long-chain PTs (C_{40–50}) [7]. Members of the short-chain ones are in ascending order C₁₀ geranyl diphosphate synthase (GPPS), C₁₅ farnesyl diphosphate synthase (FPPS), C₂₀ geranylgeranyl diphosphate synthase (GGPPS), and C₂₅ geranylfarnesyl diphosphate synthase (GFPPS) (Table 1). Their respective products serve as substrates for several downstream enzymes and, as such, represent important branch points in isoprenoid biosynthesis.

C₁₀ geranyl diphosphate (GPP), the precursor of monoterpenoids, is the simplest product of the condensation reaction of DMAPP with IPP that is controlled by GPPS residing in plastids [8,17] (Figure 3, Table S1). Initially, the GPPS function was attributed to a homodimeric protein (At2g34630) [18]. Later, however, it was revealed that GPPS in *A. thaliana* exists only as a heterodimer composed of one large subunit (LSU) and one small subunit (SSU) [8,9]. While LSU as a homodimer also functions as GGPPS (GGPPS11, At4g36810), its activity is shifted towards enhanced GPP production, when it dimerizes with SSU (At4g38460). SSU itself is not catalytically active due to the lack of one of the two aspartate-rich motifs, which were shown to play a role in binding prenyl-PP substrates by PTs [19,20] (Figure 4).

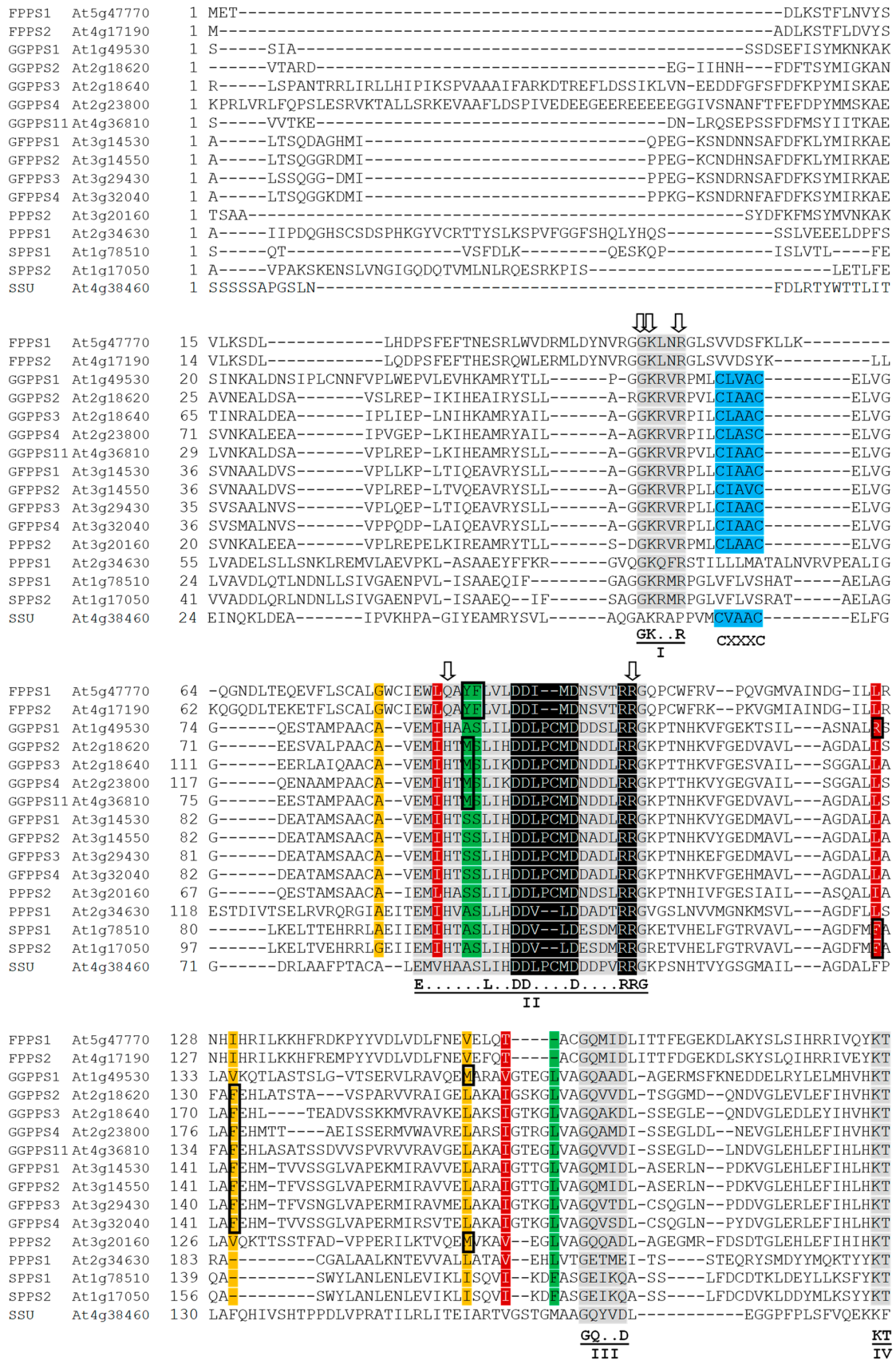


Figure 4. Cont.

FPPS1	At5g47770	193	AYYSFYLPVACALLMAGENLENHIDVKNVLDVDMGIYFQVQ	DDYLD	CFADPETLGK-IGTDIEDFKCSWL	
FPPS2	At4g17190	192	AYYSFYLPVACALLMAGENLENHTDVKTVLVDVDMGIYFQVQ	DDYLD	CFADPETLGK-IGTDIEDFKCSWL	
GGPPS1	At1g49530	201	AVLVEAAAVVGAI-MGGGSDEEIERLKSARCVGLMFQVM	DDVLD	DETKSSEELGKTAGKDLITGKLTYPK	
GGPPS2	At2g18620	194	AVLLEAATVLGAI-VGGGSDEEVEKLRRFARCI	GLLFQVV	DDILD	DETKSSEELGKTAGKDLIADKLTYPK
GGPPS3	At2g18640	233	GSLLEASAVIGAV-IGGGTEKEIEKVRNFARCI	GLLFQVV	DDILD	DETKSSEELGKTAGKDKVAGKLTYPK
GGPPS4	At2g23800	239	AVLLEAATVLGAI-IGGGSDDEIESVRKFARCI	GLLFQVV	DDILD	DETKSSEELGKTAGKDLQLAGKLTYPK
GGPPS11	At4g36810	201	AALLEASAVLGAI-VGGGSDEEIERLRKFARCI	GLLFQVV	DDILD	DETKSSEELGKTAGKDLIADKLTYPK
GFPPS1	At3g14530	207	AALLEAAAVLGI-MGGGTEQIEKLRKYARCI	GLLFQVV	DDILD	DETKSSEELGKTAGKDVMAKLTYP
GFPPS2	At3g14550	207	AALLEAAAVLGI-MGGGTEEEEIEKLRKYARCI	GLLFQVV	DDILD	DETKSSEELGKTAGKDVMAKLTYP
GFPPS3	At3g29430	206	AALLEAAAVLGI-MGGGTEEEEIEKLRKYARCI	GLLFQVV	DDILD	DETKSSEELGKTAGKDVMAKLTYP
GFPPS4	At3g32040	207	AALLEAAAVLGI-MGGGTEEEEIQKLRKYGRCI	GLLFQVV	DDILD	DETKSSEELGKTAGKDVMAKLTYP
PSPS1	At3g20160	191	AALLEAAAVMGI-MGGGSDEEIERLRSYARCI	GLMFQVV	DDVLD	DETKSSEELGKTAGKDLIAGKLTYP
PSPS2	At2g34630	237	ASLISNSCKAVAV-LTGQTAEVAVLAFEYGRNLGLAFQLI	DDILD	FTQSTASLKGKSLDIRHGVI	
SPPS1	At1g78510	193	ASLVAASTKGAAI-FSRVEPDVTEQMYEFGKNLGLSFQIV	DDILD	FTQSTEQLGKPAAGSDLAKGNLTAPV	
SPPS2	At1g17050	210	ASLVAASTK-GAAIFSKVESKVAEQMYQFGKNLGLSFQIV	DDILD	FTQSTEQLGKPAANDLAKGNITAPV	
SSU	At4g38460	189	GAMGECSAVCGGL-LGGATEDELQSLRRYGRAVGM	YQVDD	ITEDKKKSYDGGAEKGMMAEELK---	
					G . . FQ . . DD . . D GK D K	
					v	
FPPS1	At5g47770	262	VKALERCSEEQTKILYENYGKDPNSNAVKVDLYKELDLEGVFMEYESKS	---	YEKLTGAIEGHQSKAI--	
FPPS2	At4g17190	261	VKALERCSEEQTKILYENYGKAEPSNAVKVLYKELDLEGFMEYEEKS	---	YEKLTGLIEAHQSKAI--	
GGPPS1	At1g49530	270	VMGVDNA---REY-----	AKRLNREAQ--	EHLQGFDS----DKVVPLL-----	
GGPPS2	At2g18620	263	LMGLEKS---KDF-----	ADKLLSDAH--	EQLHGFDs----SRVKPLL-----	
GGPPS3	At2g18640	302	VIGVEKS---KEF-----	VEKLRDAR--	EHLQGFDS----DKVKPLI-----	
GGPPS4	At2g23800	308	LIGLEKS---KEF-----	VKRLTKDAR--	QHLQGFSS----EKVAPLV-----	
GGPPS11	At4g36810	270	IMGLEKS---REF-----	AEKLNREAR--	DQLLGFDS----DKVAPLL-----	
GFPPS1	At3g14530	276	LIGLEKS---REV-----	AEKLRREAE--	EQLLGFDP----SKAAPLV-----	
GFPPS2	At3g14550	276	LIGLERS---KEV-----	AEKLRREAE--	EQLLGFDP----SKAAPLV-----	
GFPPS3	At3g29430	275	LIGLERS---REV-----	AEKLRREAA--	EQLLGFDS----DKAAPLV-----	
GFPPS4	At3g32040	276	LIGLERS---REV-----	AEKLRREAA--	EQLLGFDS----NKVAPLV-----	
PSPS2	At3g20160	260	LMGVEKS---KEY-----	AERLNIEAR--	EHLQGFDP----IDKAPLV-----	
PSPS1	At2g34630	306	LFAMEEFPQLREV-----	VDQVEKDPNVDIALEYL	GKSKGIQRARELAMEHANLAAA	
SPPS1	At1g78510	262	IFALEREPRLEI-----	IESEFCEAGSLEEAE	IAVTKGGGKIKRAQELAREKADDAIKN	
SPPS2	At1g17050	279	IFALENEPRLEI-----	IESEFCEPGSLEEAE	IEIVRNRRGGIKRAQELAKEKAEALAKN	
SSU	At4g38460	255	-----	EKAKKE-----	LQVFDNKYGGDTLVPLV-----	
				↓		
FPPS1	At5g47770	328	-----	QAVLKSFLAKIYK--	RQK	
FPPS2	At4g17190	327	-----	QAVLKSFLAKIYK--	RQK	
GGPPS1	At1g49530	304	-----	SLADYIVK----	RQN	
GGPPS2	At2g18620	297	-----	ALANYIAK----	RQN	
GGPPS3	At2g18640	336	-----	ALTNFIAN----	RNH	
GGPPS4	At2g23800	342	-----	ALTTFIAN----	RNK	
GGPPS11	At4g36810	304	-----	ALANYIAY----	RQN	
GFPPS1	At3g14530	310	-----	ALASYIAC----	RHN	
GFPPS2	At3g14550	310	-----	ALASYIAC----	RHN	
GFPPS3	At3g29430	309	-----	ALASYIAC----	RHN	
GFPPS4	At3g32040	310	-----	ALASYIAC----	RHN	
PSPS2	At3g20160	294	-----	SLADYIVN----	RQN	
PSPS1	At2g34630	360	IGSLPETDNEVDKRSRRALID-LTH-RVITRNK			
SPPS1	At1g78510	316	LQCLPRSG-----	FRSALED-MVL-	YNLERID	
SPPS2	At1g17050	333	LNCLPRSG-----	FRSALEDMMVMF--	NLERID	
SSU	At4g38460	279	-----	TFVDYAAH--RHFLPL		

Figure 4. Multiple sequence alignment for *A. thaliana* trans-PTs, listed in Table 1. N-terminal sequences are omitted according to prediction by TargetP-2.0 [21] (<http://www.cbs.dtu.dk/services/TargetP>). Highly conserved regions are highlighted in gray. Only regions II and III are present in small subunit (SSU). Motifs essential for catalysis and binding of allylic substrate (first aspartate-rich motif (FARM), second aspartate-rich motif (SARM), and +5 and +6 arginines (RR)) are highlighted in black [19,20]. Basal residues active in IPP binding are marked with arrows [22]. The CXXXC motif highlighted in blue is critical for physical interaction between the large subunit (LSU) and the SSU [9]. The floor 1, 2, and 3 residues of the three-floor model [23] are indicated in green, red, and yellow, respectively. Blocking residues are boxed.

For the physical interaction between the subunits, the CXXXC (X is any hydrophobic amino acid residue) motif was found to be critical [9]. This motif is characteristic particularly for the SSU subfamily, but it is also present in GGPP and GFPP synthases and in polyprenyl-PP synthase 2 (PSPS2) (Figure 4). In *A. thaliana*, this type of interaction was reported specifically for GGPPS11 and SSU, earlier annotated as GGPPS12 [24]. Because both enzymes reside in plastids, GPP synthesis can be assigned solely to plastids. This is in agreement with monoterpenoid biosynthesis, which is located in plastids [16].

FPP is formed from DMAPP or GPP and IPP via action of FPPS [25] (Table S1). It is a major substrate for cytosolic and mitochondrial branches of isoprenoid biosynthesis (Figure 3). FPP-derived compounds, such as sterols, brassinosteroids, dolichols, and ubiquinones (UQ), are essential for plant growth and development. Others, such as triterpenoids and sesquiterpenoids, mediate defense reactions. Last, but not least, FPP plays a very important role in protein prenylation and hence in cell signaling [26]. In *A. thaliana*, two *FPPS* genes were identified and characterized—*FPPS1* (*At5g47770*) with two protein variants and *FPPS2* (*At4g17190*) [25,27–32]. The longer-variant FPPS1L is targeted to mitochondria, the shorter FPPS1S, together with the FPPS2, is localized to the cytosol. Research on this topic demonstrated that *FPPS* genes are redundant, thus one copy of the gene is enough to keep plants viable with only small changes in metabolite content [25,30–33]. On the other hand, total loss of FPPS activity is lethal due to severely impaired embryo development [31].

GGPP is a product of GGPPS, which accepts allylic DMAPP, GPP, and FPP in the condensation reaction with IPP [5,9,23] (Table S1). In plant cells, GGPP is used for the synthesis of gibberellins and other diterpenoids, as well as carotenoids; it also participates in chlorophyll, tocopherol, plastoquinone (PQ), UQ, and phylloquinone biosynthesis and enables protein geranylgeranylation (Figure 3). GGPPS is encoded by a gene family in *A. thaliana*. At first, 12 gene paralogs were predicted to form the *GGPPS* gene family (*GGPPS1-12*) [24], but with progressive research in the field and more sophisticated methods of analysis, several gene paralogs were excluded from this family.

Table 1. Enzymes and gene models of prenyltransferases in *A. thaliana*.

ENZYME			GENE						
Acronym	Name	EC Number	AGI	Gene Model	References	Alternative Names	References	Localization ^b	References
GPPS	Geranyl diphosphate synthase ^c	EC 2.5.1.1	<i>At4g36810/At4g38460</i>	<i>GGPPS11L^a/SSU</i>		<i>IDS11; GGPPS11; GGPS1/GGPPS12</i>	[5,24,34]/[24]	Pl/Pl	[17,34–36]/[8]
FPPS	Farnesyl diphosphate synthase	EC 2.5.1.10	<i>At5g47770</i>	<i>FPPS1S^a</i>	[24,37]			C	[25]
				<i>FPPS1L^a</i>	[24,37]		Mt	[37]	
			<i>At4g17190</i>	<i>FPPS2</i>	[24]		C	[25]	
GGPPS	Geranylgeranyl diphosphate synthase	EC 2.5.1.29	<i>At1g49530</i>	<i>GGPPS1</i>	[24]	<i>IDS1; GGPS6</i>	[5,34]	Mt	[5,17,34,35,38]
			<i>At2g18620</i>	<i>GGPPS2</i>	[24]	<i>IDS2</i>	[5]	Pl	[17]
			<i>At2g18640</i>	<i>GGPPS3</i>	[24]	<i>IDS3; GGPS4</i>	[5,34]	ER	[17,34]
			<i>At2g23800</i>	<i>GGPPS4</i>	[24]	<i>IDS4; GGPS2</i>	[5,34]	ER	[17,34,39]
			<i>At4g36810</i>	<i>GGPPS11S^a</i>	[36]			C	[36,40]
			<i>At4g36810</i>	<i>GGPPS11L^a</i>		<i>IDS11; GGPPS11; GGPS1</i>	[5,24,34]	Pl	[17,34,35]
GFPPS	Geranylarnesyl diphosphate synthase	EC 2.5.1.81	<i>At3g14530</i>	<i>GFPPS1</i>	[23]	<i>IDS6; GGPS6</i>	[5,24]	Pl	[5,17]
			<i>At3g14550</i>	<i>GFPPS2</i>	[23]	<i>IDS7; GGPPS7; GGPS3</i>	[5,24,34]	Pl	[5,17,34]
			<i>At3g29430</i>	<i>GFPPS3</i>	[23]	<i>IDS9; GGPPS9</i>	[5,24]	Pl	[5,17]
			<i>At3g32040</i>	<i>GFPPS4</i>	[23]	<i>IDS10; GGPPS10</i>	[5,24]	Pl	[5,17]
PPPS	Polyprenyl diphosphate synthase	EC 2.5.1.91	<i>At2g34630</i>	<i>PPPS1</i>	[41]	<i>GPPS; SPSS3</i>	[24,42]	Pl/Mt	[18,43]
			<i>At3g20160</i>	<i>PPPS2</i>	[23]	<i>IDS8; GGPPS8</i>	[5,24]	Pl	[17]
SPPS	Solanesyl diphosphate synthase	EC 2.5.1.85	<i>At1g78510</i>	<i>SPPS1</i>	[42]			Pl	[42,44]
			<i>At1g17050</i>	<i>SPPS2</i>	[42]			Pl	[44–46]
			<i>At1g11755</i>	<i>LEW1</i>	[47]			PM ^d	
			<i>At2g23410</i>	<i>CPT1</i>	[47]	<i>AtcPT3</i>	[48]	ER	[48,49]
CPT	<i>cis</i> -Prenyltransferase	EC 2.5.1.87	<i>At2g23400</i>	<i>CPT2</i>	[47]	<i>AtcPT2</i>	[48]	C ^d	
			<i>At2g17570</i>	<i>CPT3</i>	[47]	<i>AtcPT1</i>	[48]	Mt ^d	
			<i>At5g60510</i>	<i>CPT4</i>	[47]	<i>AtcPT9</i>	[48]	C ^d	
			<i>At5g60500</i>	<i>CPT5</i>	[47]	<i>AtcPT8</i>	[48]	C ^d	
			<i>At5g58780</i>	<i>CPT6</i>	[47]	<i>AtcPT5</i>	[48]	ER	[48,50]
			<i>At5g58770</i>	<i>CPT7</i>	[47]	<i>AtcPT4/AtHEPS</i>	[48,51]	Pl	[48,51]
			<i>At5g58782</i>	<i>CPT8</i>	[47]	<i>AtcPT6</i>	[48]	ER	[47]
<i>At5g58784</i>	<i>CPT9</i>	[47]	<i>AtcPT7</i>	[48]	ER ^d				

Abbreviations: EC, Enzyme Commission; AGI, The *Arabidopsis* Genome Initiative. ^a Alternative transcription start site. ^b Subcellular localization: C, cytosol; ER, endoplasmic reticulum; Mt, mitochondrion; Pl, plastid; PM, plasma membrane. ^c Heterodimeric GPP synthase. ^d SUBA prediction [52]; <http://suba.live/aboutSUBA4.html>.

One of them was *GGPPS5* (*At3g14510*). Because of a frame-shift mutation in its nucleotide sequence, *GGPPS5* encodes a nonfunctional protein and is therefore considered a pseudogene [17]. SSU has no GGPPS activity on its own, but upon dimerization with *GGPPS11*, changes its product specificity more towards GPP production [9]. In recent years, all the remaining members of the *GGPPS* family were reexamined for their function, and the analysis brought up surprising results (Table S1). Two independent studies provided evidence that *GGPPS6* (*At3g14530*), *GGPPS7* (*At3g14550*), *GGPPS9* (*At3g29430*), and *GGPPS10* (*At3g32040*) encode GFPPS enzymes rather than GGPPS [5,23]. Their names were therefore changed to *GFPPS1*, *GFPPS2*, *GFPPS3*, and *GFPPS4*, respectively [23]. Additionally, *GGPPS8* (*At3g20160*) also failed to synthesize GGPP [5,23] and instead produced a mix of prenyl-PPs ranging from C₂₀ to C₄₀, with the predominant C₃₀–C₃₅ species. *GGPPS8* was therefore renamed to *PPPS2* [23], and together with *PPPS1* (*At2g34630*), represents the medium chain-length PTs in *A. thaliana*. The remaining GGPPS enzymes showed solely or predominantly GGPPS activity and therefore remained members of the *A. thaliana* GGPPS family. However, the *GGPPS1* (*At1g49530*) classification is still controversial, because Nagel et al. [5] showed that this enzyme has both GFPPS and GGPPS activity, with GFPPS activity being the prominent one, whereas Wang et al. [23] showed that this enzyme has only GGPPS activity (Table S1). Thus, *GGPPS1* is considered to be a member of the *GGPPS* gene family, which has five members: mitochondrial *GGPPS1*, plastidial *GGPPS2* (*At2g18620*), *GGPPS11*, endoplasmic reticulum (ER)-bound *GGPPS3* (*At2g18640*), and *GGPPS4* (*At2g23800*). Their activities were confirmed in several studies by in vitro enzymatic assays, by genetic complementation of GGPPS-deficient *Escherichia coli*, or by analysis of *A. thaliana* mutants [5,9,17,23,34,36,53,54].

C₂₅ GFPP is a precursor of sesterterpenoids, which are compounds with multiple bioactive properties and thus promising pharmacological potential [55]. They were occasionally reported in plants, but their origin had not been addressed until recently. Only with the discovery of GFPPS activity in *A. thaliana* has the topic become more attractive. GFPP is in *A. thaliana* synthesized by *GFPPS1* (*At3g14530*), *GFPPS2* (*At3g14550*), *GFPPS3* (*At3g29430*), and *GFPPS4* (*At3g32040*) (Figure 3). In in vitro tests, these isozymes utilized preferentially GPP, FPP, and GGPP as prenyl donors and, with lower efficiency, DMAPP as well [5,23] (Table S1). Although GFPP is a predominant product of the GFPPSs, the enzymes also have considerable GPPS and GGPPS activity [5,23]. This can explain why complementation tests of carotogenic *E. coli* lacking functional GGPPS provided false positive results when using GFPPS [17,23]. *GFPPS* genes, at least in Brassicaceae, occur in the vicinity of the terpene synthase (*TPS*) genes forming *PT-TPS* clusters [23,56,57].

Long-chain *trans*-PTs in *A. thaliana* are represented by solanesyl diphosphate synthases (SPPSs). Their C₄₅ long product solanesyl diphosphate (SPP) can be found in prenylquinones as a prenyl side chain, which serves to anchor the molecule into the membrane [58]. Plant prenylquinones operate in electron transfer during photosynthesis in plastids (PQ) and during respiration in mitochondria (UQ). The length of their prenyl attachments vary among species—in *A. thaliana* plastoquinone 9 (PQ-9) and ubiquinone 9 (UQ-9) are predominant prenylquinone species [45], which means, they have mainly nonaprenyls (C₄₅) attached. The first *Arabidopsis* *SPPS1* (*At1g78510*) gene was described in 2003 by Hirooka et al. [59] The sequence was identified based on the similarity with *Arabidopsis* FPP and GGPP synthases and also bacterial long-chain PTs. According to in vitro enzymatic reaction, *SPPS1* could utilize FPP and GGPP but not DMAPP and GPP as substrates. Its predominant product was SPP (Table S1). Another long-chain PT (*SPPS2*; *At1g17050*), with 79% amino acid sequence identity to *SPPS1* (*At1g78510*), was described by Hirooka et al. in 2005 [46]. *SPPS2*, similar to *SPPS1*, used only FPP or GGPP but not DMAPP with IPP to synthesize SPP (Table S1). In vivo activity of both enzymes was checked in the background of *Schizosaccharomyces pombe* *d1p1* and *dps1* mutants lacking functional decaprenyl-PP synthase [45]. When either of the mutants expressed *SPPS1* or *SPPS2*, its ability to grow on minimal medium was restored. Both enzymes reside in plastids [42,44–46] (Figure 3), although *SPPS1* was initially described to localize to the ER [45,46]. In plastids, both enzymes interact with fibrillin FBN5, which is believed to be an indispensable structural component of PQ-9 biosynthesis through the direct stimulation of SPPS activity [44]. On the other hand, FBN5 did not interact

with a plastidial fraction of PPPS1 (PPPS1pl; At2g34630), which is a medium-chain polyprenyl-PP synthase also able to catalyze the solanesyl formation [41,43]. PPPS1 was firstly characterized as a homomeric GPPS, active in plastidial monoterpene biosynthesis [18]. The corresponding gene shared 24.2%–28.7% identity with *Arabidopsis* short-chain PTs known at that time. When Hirooka et al. studied *SPPS1* [59] and *SPPS2* [46], they noticed a relatively high level of sequence identity between their genes and the previously misidentified *GPPS* (39% and 40%, respectively). However, the information was communicated very marginally, and only later did it draw the needed attention. Hsieh et al. [41] found out that putative GPPS can make C₂₀–C₄₅ products out of GPP, FPP, or GGPP when IPP is in excess (Table S1). They reasoned that previous in vitro testing of GPPS could have been negatively affected by inappropriate reaction conditions in terms of the used substrates and their mutual ratio. Finally, the enzyme's participation in UQ biosynthesis was proved by successful complementation of the yeast *coq1* mutant [43]. PPPS1 has dual mitochondrial and plastidial targeting, probably thanks to a genuine bifunctional pre-sequence [60] (Figure 3). In *PPPS1*-silenced and overexpressing *Arabidopsis* lines only the amount of UQ-9 but not PQ-9 was found to vary [43]. Thus, PPPS1 was proposed to synthesize the side chain of UQ-9 in mitochondria but not that of PQ-9 in plastids. Summarizing the data, we can conclude that *SPPS1* and *SPPS2* contribute to PQ-9 biosynthesis, whereas PPPS1 is involved in UQ-9 biosynthesis. This functional difference is further supported by their separate evolutionary history—*SPPS1* and *SPPS2* are close relatives of cyanobacterial *trans*-long-chain PTs, whereas PPPS1 has strictly eukaryotic origin [42].

2.2. *Cis*-Prenyltransferases

In plants, *cis*-prenyltransferases (*cis*-PTs or CPTs) catalyze the synthesis of polyprenols and dolichols that are usually longer than C₅₀ in length [61]. Because CPTs use short all-*trans*-prenyl diphosphates—mostly FPP and GGPP as priming molecules—which conjugate with IPP in a *cis* orientation (Table S1), the resulting products exhibit mixed di-*trans*-poly-*cis* or tri-*trans*-poly-*cis* stereoisomerism, respectively [62] (Figure 1). While polyprenols are unsaturated at the terminal α -unit and are common in bacteria and plant photosynthetic tissues, dolichols are α -saturated and typically occur in yeasts, animals, and plant roots [49]. In fact, most organisms accumulate both forms of these polyisoprenoid alcohols, but the composition and ratio vary from species to species and also in a tissue-specific manner. The dynamic nature of polyisoprenoid metabolism can be observed during development or in response to various environmental stimuli suggesting diverse biological functions of these compounds [47]. Polyprenols, as well as dolichols, have the capacity to modulate the properties of membranes, they can serve as donors for protein prenylation, and dolichols, in particular, are inevitable saccharide carriers in protein glycosylation. Moreover, polyprenols that are quite abundant in chloroplasts can affect photosynthetic efficiency [51]. Current knowledge on the polyisoprenoid synthesis and function in plant cells is rather restricted and regulatory mechanisms controlling their biosynthesis are far from being understood.

Searching for plant CPTs, nine yeast CPT homologues (*CPT1*-9) and gene-encoding CPT-binding protein (CBP) *LEW1* were identified in the *A. thaliana* genome (Table 1). So far, only four CPTs have been characterized and proven to be functional CPT enzymes. The first *Arabidopsis* CPT (*CPT1*, At2g23410) was described by the independent studies of Oh et al. [63] and Cunillera et al. [64]. Interestingly, the protein produced a different spectrum of polyprenols in vitro and in planta (more than C₁₀₀) [49,63] and in vivo using *rer2* Δ yeasts that lack intrinsic CPT activity (less than C₁₀₀) [49,64]. Moreover, the range of detected polyisoprenoids differed slightly even between roots and leaves in *Arabidopsis* *CPT1* overexpressing lines [49]. Species and tissue-specific differences in chain length that were observed may be the result of different regulation of polyisoprenoid biosynthesis in different species and tissue background or may result from the availability of different substrates. Biosynthesis of short-chain polyisoprenoids in *A. thaliana* is mediated by heptaprenyl-PP synthase *CPT6*/*HEPS* (At5g58780), which catalyzes the formation of dolichol with seven isoprenoid units (Dol-7) [48,50] (Table S1). Both *CPT1* [49] and *CPT6* [48,50] were found to reside in the ER (Figure 3). *CPT7* (At5g58770) is so far the

only plastidial CPT that was characterized in *Arabidopsis* [51] (Figure 3). Following its subcellular localization, GGPP not FPP is likely the substrate for the synthesis of C₅₅ polyisoprenoids, and mutant analysis of the *cpt7-1* null mutant showed that CPT7 influences photosynthetic performance via modulation of thylakoid membrane dynamics.

There are two types of CPTs in plants—homomeric and heteromeric [65]. Homomeric ones act autonomously in the contrary to the heteromeric CPT enzymes, which need to dimerize with a CBP to gain its activity [66–68]. So far, the only CBP identified in *A. thaliana* is LEW1 (At1g11755), which was described to form a heterocomplex with CPT3 (At2g17570). It was shown that CPT3/LEW1 synthesizes long-chain polyisoprenoids, whereas individual CPT3 or LEW1 proteins were catalytically inactive [10]. Synthesis of polyisoprenoids by CPT/CBP heterocomplexes is typical for the eukaryotic lineage of CPTs, based on the phylogenetic analysis. Besides CPT3 (At2g17570), CPT4 (At5g60510) and CPT5 (At5g60500) also belong to this lineage in *A. thaliana* [48]. The remaining *Arabidopsis* CPTs cluster together in a prokaryotic clade, suggesting that they are self-sufficient homomeric enzymes. Recently, it was shown, that CBPs and homomeric CPTs have in common a highly conserved RXG (eventually NXG) motif in their C-terminal sequence, which is essential for catalysis [65] (Figure S1). On the other hand, the motif is absent in the heteromeric CPTs, and therefore, interaction with a CBP is required. Furthermore, CBPs are not able to bind substrates due to deletion, and thus, creating a CPT/CBP dimer is essential for the function of both proteins.

3. Primary Structure of PTs Can Be Indicative for Determination of Their Function

Protein functions are closely related to their structures. However, the structure was determined only for a small fraction of the myriad protein sequences deposited to databases. Therefore, functional characterization of novel sequences still heavily relies on homology-based approaches, in which experimentally determined functions are transferred to uncharacterized homologous sequences with various levels of confidence [69]. Such methods use the fact that homologous sequences sharing a high level of identity (at least 40%) and similarity are often isofunctional. However, *Arabidopsis* GGPP and GFPP synthases are functionally different, although they share more than 55% sequence identity [23]. Thus, the sequence alignment is not robust enough to assign the function to a protein, although it is still a convenient method for rough classification of unknown proteins.

Five highly conserved regions (I–V) can be recognized in the sequence of every *trans*-PT [70] (Figure 4). Regions II and V include aspartate-rich motifs DDX₍₂₋₄₎D, referred to as the first aspartate-rich motif (FARM) and the second aspartate-rich motif (SARM). These motifs and +5 and +6 arginines (RR) downstream from the FARM are essential for catalysis and binding of the allylic substrate [19,20]. The attachment of IPP is coordinated by the C terminus of the enzyme and dispersed conserved basal residues [22]. From the point of the tertiary structure, all conserved features are associated with the active site, which is located in a large deep cleft of the enzyme formed out solely of α -helices [23,41]. The capacity of the catalytic pocket, where prenyl-PP elongation takes place, is determined by residues, which protrude into the cavity. Residues with large side chains can obstruct the cavity and thus further prenyl elongation [71]. It is currently unknown whether the mechanism is also valid for CPTs.

To describe the residues, which are an obstacle in the enzyme elongation channel, several authors have adopted the term “floor”. So far, single-, double-, and three-floor models of *trans*-PT chain-length determination have been proposed [23,41,71–74]. The three-floor model suggested by Wang et al. [23] is the most precise one. The model was deduced from the observations made on the crystal structures of three representatives of *Arabidopsis* PTs—GGPPS11, GFPPS2, and PPPS2. The authors revealed that some residues within the enzyme pocket assemble to form a floor, and there are three such floors along the pocket and thus three potential points where the synthesis of a prenyl product can be arrested due to steric hindrance from the blocking residue(s). Following the model scheme, the FPPS activity can be predicted from the two bulky residues (Tyr and Phe) on floor 1 [71] (Figure 4).

The GGPPS activity could be predicted from Phe on floor 3 and medium-sized Met on floor 1, which only partially closes the cavity (Figure 4). If solely Phe on floor 3 is present, the reaction

should terminate with the release of GFPP (Figure 4). Dual GGPP/GFPP function could be indicated by Arg on floor 2 and Met on floor 3, which were detected in GGPPS1 (Figure 4). Finally, PPPS activity could be deduced from the lack of blocking residues, because there is only Met on floor 3 in PPPS2 (Figure 4). Predictions using the three-floor model are in agreement with experimental data on the product specificity of selected enzymes; thus, the model is functional at least with *Arabidopsis* short- and medium-chain *trans*-PTs. However, it would be interesting to see if it also works with long-chain *trans*-PTs. Mapping the model's sites on the respective residues in PPPS1 has shown that no bulky residue is present in the pocket, but there are two in the SPPS1 and SPPS2—two molecules of Phe on floor 1 and 2 (Figure 4). Following the concept of the three floors model, the lack of obstacles in PPPS1 suggests that it is indeed PPPS. On the other hand, two residues blocking the cavity in SPPS are rather unlikely and hence challenge the relevance of the model with respect to the prediction of long-chain *trans*-PTs. In that case, the crystal structure of SPPS would be needed to optimize the model, but so far it is not available. Yet, we can see that in case of PTs, not the total sequence similarity but either the presence of specific domains or amino acid residues have predictive power when analyzing the amino acid sequence.

In past decades, determination of the key residues forming the pocket had been preceded by resolving the three-dimensional (3D) structure of the enzyme and/or laborious site-directed mutagenesis experiments [22,23,41,73–76]. In recent years, however, we can witness considerable efforts to create *in silico* prediction models, which could reliably assign function to a protein based on its raw sequence [1,77,78]. Naturally, 3D conformations remain an invaluable source of data needed to design a powerful algorithm.

4. Prenyltransferase Mutants: A Source of Information about Their Function and Topology

Analysis of knockouts and knockdowns is the way to understand the functionality of pathway enzymes within an *in planta* context. Visible phenotypes, moreover, can have predictive function for the topology of yet uncharacterized enzymes and for our knowledge on carbon fluxes via the isoprenoid pathway network. Both specific inhibitors and gene mutants were extensively used to analyze early enzymatic steps in the isoprenoid pathway (MVA and MEP pathways). In the case of PTs, however, mainly mutants were used to characterize their *in planta* function (Table 2).

Mutants in the linear metabolic pathway are expected to have similar phenotypes. MVA pathway null mutants are embryo lethal or male sterile, affected in male gametophyte development or in pollen tube elongation, whereas MEP pathway mutants grow to the seedling stage, but remain albinos [2]. In fact, there is only one null pathway mutant in the MVA pathway that is embryo and not male gametophyte lethal. It is an *aact2* mutant impaired in acetoacetyl-CoA transferase, acting at the beginning of the MVA pathway [79]. It is expected that mutants in essential downstream pathways stemming from either MVA or MEP pathway will have identical or similar phenotypes. Indeed, mutants in the cytosolic/mitochondrial essential pathways, such as the sterol biosynthetic pathway, UQ biosynthetic pathway, protein Rab geranylgeranylation pathway, and dolichol biosynthetic pathway, have either impaired male gametophyte function or are embryo lethal [80–83]. Similarly, mutants in essential biosynthetic plastidial pathways, such as the carotenoid biosynthetic pathway, chlorophyll biosynthetic pathway, and PQ biosynthetic pathway, are albino or pale green [36,42,84,85]. In agreement with that, mutants in PTs that provide substrates for those specialized isoprenoid pathways should display similar phenotypes.

4.1. Prenyltransferase Mutants in the Cytosolic/Mitochondrial Branch of the Isoprenoid Pathway

The first step in the DMAPP and IPP condensation in the cytosol and mitochondria is FPP synthesis (Figure 3). There are two enzymes (FPPS1S, FPPS2) in the cytosol and one enzyme (FPPS1L) in the mitochondria that synthesize FPP. FPPS1L is encoded by the same locus as FPPS1S but only has mitochondrial targeting sequence. Neither the *fpps1* nor the *fpps2* null mutant (the *fpps1* mutant is mutant in both FPPS1S and FPPS1L enzymes) shows the visible phenotype, and the levels of sterols

and UQ-9, the major cytosolic and mitochondrial isoprenoids, are only slightly reduced in single mutants (Table 2). This demonstrates that both enzymes are highly redundant in *Arabidopsis*. Moreover, it shows that the function of FPPS1L in the mitochondria is not essential, excluding FPPS1L from being the only source of substrate for UQ-9 biosynthesis.

The *fpps1fpps2* double mutant is embryo lethal, arrested at the globular stage of development. The *fpps1fpps2* male gametophyte is only slightly affected in pollen tube elongation, which means that compared with MVA pathway mutants, which are mostly male gametophyte lethal, the *fpps1fpps2* phenotype is less severe [2,31]. FPP is in the cytosol mainly used for the synthesis of phytosterols and the null mutant in the cycloartenol synthase (*cas1*), which synthesizes cycloartenol for the phytosterol biosynthetic pathway, is also male gametophyte lethal, as most of the MVA pathway mutants [81]. The less severe phenotype of *fpps1fpps2* as compared with other mutants in sterol biosynthesis indicates that FPPS activity can be partly complemented, at least in some developmental stages. So far, we can only speculate about the source of the complementation. FPP can be imported from the wild-type cells of surrounding tissues or be synthesized by other PTs. For example, GGPPS11, which is also present in the cytosol, can synthesize a small amount of FPP (2.1%) [9] (Table S1).

Table 2. Mutations in prenyltransferase genes and their phenotypes in *A. thaliana*.

Gene	AGI	Mutagen	Allele	Mutant Line	Phenotype	References
FPPS1	At5g47770	T-DNA	<i>fppls1-1</i>	SAIL_310-D07	Wild-type phenotype with slightly reduced levels of sterols and UQ-9	[31]
		T-DNA	<i>fppls1-2</i>	SALK_073576		
FPPS2	At4g17190	T-DNA	<i>fppls2-1</i>	SAIL_328_G06	Wild-type phenotype with slightly reduced levels of sterols and UQ-9	[31]
		Ds	<i>fppls2-2</i>	GT7041		
FPPS1/FPPS2	At5g47770/At4g17190	T-DNA/Ds	<i>fppls1/fppls2</i>	SAIL_310_D07; SALK_073576/SAIL_328_G06; GT7041	Embryo lethal at the globular stage; slightly impaired pollen tube elongation	[31]
		amiRNA	<i>fppls1/fppls2</i>	amiFPPS1 (21%/26% mRNA), amiFPPS2 (16%/35% mRNA)		
SSU	At4g38460	T-DNA	<i>ssu-1</i> (<i>ggpps12-1</i>)	pst11416	Wild-type phenotype with reduced level of monoterpenoids, and wild-type level of carotenoids, chlorophylls, sesquiterpenoids	[8]
		T-DNA	<i>ssu</i> (<i>ggpps12</i>)	pst01798 (20% mRNA)		
GGPPS1	At1g49530	T-DNA	<i>ggpps1-1</i>	SAIL_559_G01	Wild-type phenotype with wild-type level of UQ-9, carotenoids, tocopherols, chlorophylls, PQ-9, phylloquinones, plastochromanol-8	[36]
GGPPS2	At2g18620	T-DNA	<i>ggpps2-1</i>	FLAG_134_B10	Wild-type phenotype with wild-type levels of carotenoids, chlorophylls, phylloquinones	[54]
GGPPS3	At2g18640	-	-	-	-	-
GGPPS4	At2g23800	-	-	-	-	-
GGPPS11	At4g36810	EMS mutagenesis	<i>ggpps11-1</i>	D163R point mutation	Variegated phenotype; germination delayed on the inhibitor of gibberellin biosynthesis; reduced level of chlorophylls and carotenoids	[54,86]
		T-DNA in chloroplast targeting sequence	<i>ggpps11-2</i>	SALK_015098	Albino seedling	[54,86]
		T-DNA	<i>ggpps11-3</i>	SALK_085914	Embryo lethal at the heart stage	[54,86]
		T-DNA	<i>ggpps11-4</i>	SAIL_712_D06		[54]
		T-DNA in 5'-UTR	<i>ggpps11-5</i>	SALK_140601	Pale green phenotype with reduced level of monoterpenoids, carotenoids, tocopherols, chlorophylls, PQ-9, phylloquinones, plastochromanol-8	[8,36,54]
GFPPS1	At3g14530	T-DNA	<i>gfpps1-1</i> (<i>ggpps6-1</i>)	SAIL_1148_A03	Wild-type phenotype with wild-type level of carotenoids, chlorophylls, phylloquinones	[54]
GFPPS2	At3g14550	T-DNA	<i>gfpps2-1</i> (<i>ggpps7-1</i>)	SALK_119280	Wild-type phenotype with wild-type level of carotenoids, chlorophylls, phylloquinones	[54]
GFPPS3	At3g29430	RNAi	<i>gfpps3</i> (<i>ggpps9</i>)	RNAi GGPPS9-1 (21% mRNA); RNAi GGPPS9-6 (16% mRNA)	Wild-type phenotype with wild-type level of carotenoids, chlorophylls, phylloquinones	[54]
GFPPS4	At3g32040	T-DNA	<i>gfpps4-1</i> (<i>ggpps10-1</i>)	SM_3_32015	Wild-type phenotypes with wild-type level of carotenoids, chlorophylls, phylloquinones	[54]

Table 2. Cont.

Gene	AGI	Mutagen	Allele	Mutant Line	Phenotype	References
PPPS1	At2g34630	T-DNA	<i>ppps1-1 (gps1-1)</i>	GABI_097_G02	Embryo lethal	[87]
		RNAi	<i>ppps1</i>	RNAi PPPS1-1-6 (10% mRNA)	Growth reduction; dwarfed plants with delayed flowering	[87]
		RNAi	<i>ppps1</i>	RNAi PPPS1-1-3 (20% mRNA)	Reduced level of UQ-9; wild-type level of PQ	[43]
		RNAi	<i>ppps1</i>	GPPS RNAi-2, GPPS RNAi-4 (10% mRNA)	Wild-type phenotype with wild-type level of monoterpenoids and sesquiterpenoids	[8]
PPPS2	At3g20160	T-DNA	<i>ppps2-1 (ggpps8-1)</i>	FLAG_470_E09	Higher level of carotenoids; reduced level of chlorophyll b	[54]
SPPS1	At1g78510	T-DNA	<i>spps1</i>	SALK_126948	Wild-type phenotype with reduced level of PQ and plastochromanol, and wild-type level of tocopherol and UQ	[42]
SPPS2	At1g17050	T-DNA	<i>spps2</i>	SALK_064292	Developmental delay; stunted phenotype and pale yellowish leaves at the high light; reduced level of PQ; no plastochromanol; higher level of tocopherol, and wild-type level of UQ	[42]
SPPS1/SPPS2	At1g78510/ At1g17050	T-DNA	<i>spps1/spps2</i>	SALK_126948/SALK_064292	Seedling lethal, albino phenotype; no PQ and plastochromanol; wild-type level of UQ and tocopherol	[42]
LEW1	At1g11755	EMS mutagenesis	<i>lew1-1</i>	G159A point mutation	Leaf-wilting phenotype, increased drought resistance, impaired plasma membrane integrity, impaired protein N-glycosylation, reduced the total plant content of main dolichols C ₇₅ , C ₈₀ by 85% and protein glycosylation defects	[88]
		T-DNA	<i>lew1-2</i>	SALK_032276	Lethal	[88]
CPT1	At2g23410	T-DNA	<i>cpt1-1</i>	SALK_038151	Extremely stunted growth and shorter root length; reduced level from Dol-18 to -23 (Dol-21 dominating)	[49]
		T-DNA	<i>cpt1-2</i>	SALK_032276		
		T-DNA	<i>cpt1-3</i>	SALK_100795		
		-	-	-		
CPT2	At2g23400	-	-	-	-	-
CPT3	At2g17570	-	-	-	-	-
CPT4	At5g60510	-	-	-	-	-
CPT5	At5g60500	-	-	-	-	-
CPT6	At5g58780	T-DNA	<i>cpt6-1</i>	SALK_071255	Reduced level of Dol-7 and short-chain dolichols (Dol-13 dominating)	[50]
			<i>cpt6-2</i>	SALK_064499		
CPT7	At5g58770	T-DNA	<i>cpt7-1</i>	SALK_022111	Decreased thylakoid membrane fluidity and photosynthetic performance; no polyprenols 9, 10, 11 (Pol-10 dominating); wild-type level of tocopherols, phylloquinone, carotenoids, PQ, and chlorophylls	[51]
		RNAi	<i>cpt7</i>	RNAi-23; RNAi-24 and RNAi-31 (15% mRNA)		
CPT8	At5g58782	-	-	-	-	-
CPT9	At5g58784	-	-	-	-	-

The second step in DMAPP and IPP condensation in the cytosol and mitochondria is GGPP synthesis (Figure 3). There are three GGPPS enzymes in the cytosol (GGPPS3, GGPPS4, and GGPPS11S) and one in the mitochondria (GGPPS1). *GGPPS11* transcript is, among all the *GGPPS* paralogs, the most abundant and the only one ubiquitously expressed in plants [17]. So far, *ggpps11* and *ggpps1* mutants have been isolated and characterized (Table 2). *ggpps1* has no visible phenotype and no decrease in metabolite levels, including UQ-9 [36]. On the contrary, the *ggpps11* mutant showed embryo lethal phenotype with the embryo arrested at the heart stage of development [36,54]. A similar phenotype is typical for the mutants deficient in the UQ biosynthesis [80]. The level of UQ-9 was tested in the *ggpps11-5* knockdown mutant, which has decreased *GGPPS11* transcript levels to about 50%. In this mutant, the level of UQ-9 was slightly decreased but not significantly [36], suggesting that the lack of UQ-9 is not the (main) reason for the embryo lethal phenotype of *ggpps11* and suggesting, together with the results of *ggpps1* analysis, that FPP rather than GGPP could be the source of prenyl-PP substrate for the UQ biosynthesis.

The level of UQ-9 was decreased significantly in *fpps1* and *fpps2* single mutants [31] and *FPPS*-silenced plants [32]. However, approximately 50% decrease in total *FPPS* activity, at the 16% level of transcripts for *FPPS1* and 35% for *FPPS2* in *FPPS*-silenced plants, caused only 30% decrease in UQ-9 levels [32]. This suggests that although *FPPS1* and *FPPS2* are contributing significantly to UQ biosynthesis, they are not the only substrates for SPP and subsequently UQ-9 biosynthesis. C_{45} solanesyl moiety, the predominant form of UQ prenyl chains in *Arabidopsis*, is synthesized by a mitochondrial *PPPS1* (*PPPS1mt*) that elongates an initial FPP or GGPP with IPP units [41,43]. The *ppps1* null mutant is, similarly to other UQ biosynthetic mutants, embryo lethal [80,87]. The reduction of *PPPS1* transcript levels by 80% in *PPPS1*-silenced plants caused almost proportional reduction in UQ-9 content (reduction by 75–80%) [43]. This indicates that while *PPPS1* is the only enzyme synthesizing C_{45} for UQ-9 biosynthesis, *FPPS* and *GGPPS* will have a rather promiscuous function in this process.

The question of whether the lack of *GGPPS* activity in the cytosol and mitochondria can to some degree be complemented by other sources of GGPP cannot be answered at the moment, because we do not have mutant(s) with null *GGPPS* activity in both the cytosol and mitochondria. What we can say so far is that the *ggpps11* mutant has a less severe phenotype than the Rab geranylgeranylation null mutant [82]. A complete lack of Rab prenylation, which uses GGPP as a substrate, causes male sterility.

So far, in *Arabidopsis*, mutants in three genes synthesizing polyprenols in the cytosol and/or mitochondria were characterized, *cpt1*, *cpt6*, and *lew1* (Table 2). In all three mutants, dolichol levels were reduced, but in each of them, different classes of dolichols were affected. While synthesis of long chain dolichols was reduced in the *cpt1* and *lew1* mutants [49], short chain dolichols were reduced in the *cpt6* mutant [50]. The phenotype was also very different in all the mutants. No visible phenotype was reported for *cpt6* [50], *cpt1* null mutation affected plant growth and root elongation [49], and *lew1* null mutation is lethal, although it is not known whether it is due to the embryo lethality or gametophyte lethality [88]. Mutants in later steps of dolichol biosynthesis, in polyprenol reductase (*PPRD2*), which catalyzes the conversion of polyprenol to dolichol, affects male gametophyte development and pollen germination [83]. Thus, the function of different CPTs in plant dolichol synthesis within the context of plant physiological processes still remains largely unknown.

4.2. Prenyltransferase Mutants in the Plastidial Branch of the Isoprenoid Pathway

The first product of DMAPP and IPP condensation in plastids is GPP, which in *Arabidopsis* is synthesized by the heteromeric enzyme *GGPPS11/SSU* (Figure 3). In agreement with that, both the knockdown *ggpps11* mutant and the null *ssu* mutant have reduced levels of monoterpenoids to approximately 50% [11] (Table 2). Heterodimer *GGPPS11/SSU* is so far considered to be the only *GPPS* in *Arabidopsis*. Nevertheless, because monoterpenoids are still synthesized in *Arabidopsis* in the *ssu* knockout mutant, it is likely that other *GPP*-synthesizing enzymes are active in *Arabidopsis*. In addition to *GGPPS11*, small *GPPS* activity was also observed in *GFPPSs* [5,23] (Table S1). Previously, *PPPS1* was classified as *GPPS*, because it was shown to synthesize *GPP* predominantly and was assumed to

function in monoterpene biosynthesis and gibberellin biosynthesis in plastids [18,87]. Nevertheless, *PPPS1*-RNAi plants do not have a decreased level of monoterpenoids [8], and GGPPS was shown to have PPPS (C₂₀–C₄₅) activity [41] (Table S1).

In agreement with the proposed housekeeping role of GGPPS11 [17,54], the *ggpps11* knockdown mutant has lower levels of chlorophylls, carotenoids, phylloquinones, PQs, and tocopherols [8,36,54]. In addition to GGPPS11, which is the main GGPP-synthesizing enzyme in plastids and essential for the synthesis of photosynthesis-related isoprenoid compounds [54], GGPPS2 is another GGPPS paralog in plastids (Figure 3). The *ggpps2* null mutant shows no visible developmental defects, and the level of carotenoids, chlorophylls, and prenylquinones is at the wild-type level (Table 2), demonstrating that GGPPS2 does not contribute significantly to the synthesis of photosynthesis-related plastidial isoprenoids that are essential for growth and development, and therefore, the function of this isozyme is yet to be determined [54].

The newly discovered group of PTs, GFPP synthases (GFPPS1-4) that synthesize GFPP, resides in plastids [17] (Figure 3). GFPPSs predominantly synthesize GFPP, which is a precursor of sesterterpenoids [5,23]. There are four GFPPSs in *Arabidopsis* plastids. GFPPS null mutants *gfpps1*, 2, 4 and *GFPPS3*-RNAi lines with 21% and 16% of mRNA levels, respectively, show no visible phenotypes [54] (Table 2).

Arabidopsis possesses two paralogous SPP synthases, SPPS1 and SPPS2, which assemble the side chain of PQ-9 in plastids (Figure 3). The *spps1spps2* double knockout is devoid of PQ-9 and plastochromanol and cannot grow photoautotrophically [42] (Table 2). Because the preferred substrate for SPPS is GGPP and not DMAPP or GPP [59] (Table S1) and the *ggpps11* mutant has decreased PQ-9 and plastochromanol levels [54], we can predict that the main enzyme generating substrate for SPPS in plastids is GGPPS11.

PPPS responsible for the synthesis of prenyl chain of UQ-9 in the mitochondria is also localized in plastids (PPPS1pl) because of double targeting [18,43] (Figure 3). While PPPS1mt synthesizes the side chain of UQ, PPPS1pl is not involved in PQ biosynthesis [43], and the function of PPPS1pl thus remains to be discovered. In tomato, downregulation of *PPPS1* transcript led to a decreased level of gibberellins [87], but it was not tested whether PPPS1pl has the same function in *Arabidopsis*.

There is another PPPS in plastids, PPPS2, which synthesizes carbon chains of C_{20–40} [23] (Figure 3, Table S1). The null *ppps2* mutant has no visible phenotype, and from all photosynthetically related metabolites that were analyzed, the level of chlorophyll b was slightly decreased and levels of β -carotene and neoxanthin slightly increased. The levels of other metabolites, such as chlorophyll a and phylloquinone were unchanged [54] (Table 2).

There is one CPT predicted to localize to plastids (Figure 3). The null mutant in plastidial CPT7 has a decreased level of plastidial polyprenols (C₄₅–C₅₅ with a maximum of C₅₀) by 100% (Table 2). Affected plants have impaired photosystem II operating efficiency, and their thylakoids exhibit a decreased rate of electron transport but are not lethal [51].

5. Co-Expression Analysis and Prediction of Isoprenoid Fluxes

Co-expression analysis determines clusters of genes that are co-regulated together at expression levels. Metabolic pathway genes involved in the same physiological process are often co-regulated, which was also shown for genes in the isoprenoid pathway [89–91]. Therefore, gene association analysis based on the gene co-expression can be very instrumental in uncovering novel metabolic pathway genes or in determining their topology in the pathway gene network.

There are several examples where gene co-expression analysis resulted in better understanding of the prenyl-PP synthase function. For example, co-expression analysis with plastidial GGPPS and GFPPS isozymes showed that only *GGPPS11*, which is the ubiquitously expressed plastidial GGPPS paralog and the only one highly expressed in green tissues [17], is significantly co-expressed with all genes encoding MEP pathway enzymes and almost all genes from pathways synthesizing photosynthesis-related isoprenoids. The result of this co-expression analysis was also supported by the

metabolite profiling of *ggpps11* mutant [54] (Table 2). Co-expression analysis also correctly predicted subcellular localization of SPPS1 in plastids [42], although previous experimental evidence showed that this protein is associated with the ER [45,46]. This prompted further experiments that proved that *SPPS1* is a paralog of *SPPS2* and both enzymes are responsible for the plastidial synthesis of SPP and are essential for PQ and plastochromanol biosynthesis [42]. Co-expression analysis was also instrumental in assigning function to the PPPS1 protein. The PPPS1 protein was originally annotated as GPPS. It was shown that GPPS produces GPP in vitro, and GFP fusion of the protein (GPPS-GFP) localizes to plastids [18]. Nevertheless, the embryo lethal phenotype of the knockout mutant was rather nontypical for a plastidial isoprenoid pathway enzyme. In search of *Arabidopsis* isoprenoid pathway genes that are co-expressed with the *Arabidopsis* Ubi/Coq orthologs (UQ biosynthesis genes), GPPS was surprisingly found to be associated with the gene co-expression cluster. The enzymatic function of GPPS was reevaluated, and it was shown that GPPS actually synthesizes PPPs of different length (C₂₅–C₄₅), with C₃₅ being the predominant prenyl-PP, and the enzyme was renamed to be PPP synthase 1 (PPPS1) [41,43]. Localization experiments using the expression of the GFP fusion protein driven by the native promoter have shown that PPPS1 localizes both to plastids and to the mitochondria [43].

We applied the same co-expression approach, which we used earlier for the plastidial *GGPPS* and *GFPPS* gene paralogs [54], to short *trans*-PTs to see whether we can extract additional information as to the function of PTs within the isoprenoid pathway gene network (Table 3, Table S2). Microarray expression data from different *A. thaliana* organs and tissues were used for analysis as described in Table S2. In the cytosol, as expected, *FPPS1* and *FPPS2* genes are the most co-expressed with the genes from the MVA and sterol biosynthetic pathways and both to the same extent supporting functional redundancy of both enzymes. As to the other functionally uncharacterized PTs and their involvement in the biosynthetic pathways in the cytosol and mitochondria, we can see that both *FPPS* and *GGPPS* genes and gene paralogs have connections to the genes of the UQ biosynthetic pathway and to the protein prenylation, and there is likely no strict allocation of a single *FPPS* or *GGPPS* gene paralog to those pathways. This seems not to be true for the diterpenoid biosynthesis, where *GGPPS3* has the most gene connections to *TPSs*. In plastids, the synthesis of isoprenoid compounds associated with photosynthesis is associated with the single *GGPPS11* gene paralog, as observed previously [54], except in the phylloquinone biosynthetic pathway, which is here, with the current gene expression set, more connected to *GGPPS2* and other *PTs*. Genes synthesizing terpenoids and gibberellins are also instead associated with several isoprenoid synthases and their paralogs. Surprisingly, monoterpenoids show more connections to *PPPS1*, which was previously annotated as *GPPS*, than with *SSU*.

Table 3. *trans*-Prenyltransferase genes and isoprenoid pathway gene co-expressions in *A. thaliana*.

CYTOSOL AND MITOCHONDRIA										
Pathway		MVA	Sterol	Triterpenoids	Sesquiterpenoids/Diterpenoids	Protein Prenylation	Ubiquinone			
Total No. of Genes per Pathway		9	24	13	13	7	4			
FPPS1	<i>At5g47770</i>	6	13	2	1	4	0			
FPPS2	<i>At4g17190</i>	7	14	0	1	3	1			
GGPPS1	<i>At1g49530</i>	0	0	2	0	0	3			
GGPPS3	<i>At2g18640</i>	0	4	3	6	1	2			
GGPPS11	<i>At4g36810</i>	0	11	2	0	1	0			
GGPPS4	<i>At2g23800</i>	2	2	3	1	0	2			
PPPS1	<i>At2g34630</i>	2	3	0	2	1	4			
PLASTIDS										
Pathway		MEP	Plastoquinone	Chlorophyll	Carotenoid	Phylloquinon	Gibberellins	Monoterpenoids	Diterpenoids/Sesquiterpenoids	Sesterterpenoids
Total No. of Genes per Pathway		7	6	37	32	7	23	6	13	3
SSU	<i>At4g38460</i>	5	5	25	22	0	5	2	1	0
GGPPS11	<i>At4g36810</i>	7	3	29	18	1	7	1	0	0
GGPPS2	<i>At2g18620</i>	0	0	3	1	5	2	2	7	1
GFPPS1	<i>At3g14530</i>	0	0	1	0	3	2	0	5	2
GGPPS3	<i>At3g29430</i>	0	0	4	0	4	2	2	8	1
GGPPS4	<i>At3g32040</i>	0	0	3	0	5	7	2	7	1
PPPS2	<i>At3g20160</i>	0	0	4	0	4	4	2	8	1
PPPS1	<i>At2g34630</i>	0	0	0	2	4	2	4	2	1
SPPS1	<i>At1g78510</i>	6	6	29	20	0	4	2	1	0
SPPS2	<i>At1g17050</i>	7	4	29	22	0	7	2	1	0

0%



100%

Table 3 shows number of pathway genes that are significantly co-expressed with the *trans*-PTs. The intensity of the highlighted background is directly proportional to the relative number of co-expressed genes/total number of pathway genes. Data for co-expression analysis were obtained from BAR/Expression Browser (http://bar.utoronto.ca/affydb/cgi-bin/affy_db_exprss_browser_in.cgi) and log₂ transformed (see Table S2 for log₂ transformed expression data). In general, AtGeneExpress_plus-Extended Tissue Series data/Average of replicate treatments were downloaded. Genes that are missing were not present on the microarrays, and therefore, they are not included in the analysis. Pearson correlation coefficients and the corresponding false discovery rate (FDR) *p*-values were calculated. The threshold for significance is *p*-value ≤ 0.05. A detailed description of the analysis can be found in [54]. Genes in categories mono-, di-, sesqui-, sester- and triterpenoids contain genes encoding only terpene cyclases. Genes in the category diterpenoids/sesquiterpenoids and sesquiterpenoids/diterpenoids partly overlap. In both categories were kept the genes that have no experimentally proven one or other activity, and plastidial genes were considered diterpenoids and kept only in the category diterpenoids/sesquiterpenoids.

Metabolic gene clusters within dynamic chromosomal regions could provide another mechanism of transcriptional co-regulation. It has become clear in recent years that many plant secondary metabolic pathways or some of their genes are clustered [92]. These clusters have arisen by recruitment of genes from elsewhere in the genome through duplication and neofunctionalization. *GFPPS–TPS* gene clusters, where *GFPPS* genes and *TPS* genes are in close proximity were also found in *Arabidopsis* [23,56,57].

6. Conclusions

In the last decade, we have witnessed enormous progress in the characterization of *Arabidopsis* isoprenoid pathway PTs. Enzymes that were initially defined only by their approximate enzymatic function are now more precisely characterized at the enzymatic level, and their biological role, as well as topology, within the isoprenoid pathway network is clearly defined. PTs with essential function in the cytosol, mitochondria, and plastids are known as metabolic processes controlled by those PTs. What still remains as a challenge is to define at the gene level metabolic routes (fluxes) between PTs and TPSs and between PTs and pathways where prenyl-PP as a substrate is ambiguous; to understand minor, dynamic, developmentally controlled fluxes; and, last but not least, to explore more physiological functions of *cis*-PTs.

Supplementary Materials: The following are available online. Figure S1: Multiple sequence alignment for *A. thaliana cis*-PTs listed in Table 1; Table S1: Substrate and product specificity and kinetics of prenyltransferases in *A. thaliana*; Table S2: Gene expression data.

Author Contributions: Both authors contributed equally to this work.

Funding: This work was supported by grant VEGA 1/0926/17 from the Scientific grant agency of the Slovak Ministry of Education, Science, Research and Sport and of the Slovak Academy of Sciences.

Acknowledgments: We appreciate the help of Diana Coman Schmid (ETH Zurich) with co-expression analysis.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wallrapp, F.H.; Pan, J.-J.; Ramamoorthy, G.; Almonacid, D.E.; Hillerich, B.S.; Seidel, R.; Patskovsky, Y.; Babbitt, P.C.; Almo, S.C.; Jacobson, M.P.; et al. Prediction of function for the polyprenyl transferase subgroup in the isoprenoid synthase superfamily. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E1196–E1202. [[CrossRef](#)] [[PubMed](#)]
2. Vranová, E.; Coman, D.; Gruissem, W. Network Analysis of the MVA and MEP Pathways for Isoprenoid Synthesis. *Annu. Rev. Plant. Biol.* **2013**, *64*, 665–700. [[CrossRef](#)] [[PubMed](#)]
3. Liang, P.-H.; Ko, T.-P.; Wang, A.H.-J. Structure, mechanism and function of prenyltransferases. *Eur. J. Biochem.* **2002**, *269*, 3339–3354. [[CrossRef](#)] [[PubMed](#)]
4. Winkelblech, J.; Shu-Ming Li, A.-F. Prenyltransferases as key enzymes in primary and secondary metabolism. *Appl. Microbiol. Biotech.* **2015**, *99*, 7379. [[CrossRef](#)]
5. Nagel, R.; Bernholz, C.; Vranová, E.; Košuth, J.; Bergau, N.; Ludwig, S.; Wessjohann, L.; Gershenzon, J.; Tissier, A.; Schmidt, A. *Arabidopsis thaliana* isoprenyl diphosphate synthases produce the C25 intermediate geranylgeranyl diphosphate. *Plant. J.* **2015**, *84*, 847–859. [[CrossRef](#)]
6. Koyama, T.; Ogura, K. Isopentenyl Diphosphate Isomerase and Prenyltransferases. In *Isoprenoids Including Carotenoids and Steroids*; Cane, D.E., Ed.; Elsevier: Amsterdam, The Netherlands, 1999; Volume 2, pp. 69–95.
7. Vandermoten, S.; Haubruge, É.; Cusson, M. New insights into short-chain prenyltransferases: Structural features, evolutionary history and potential for selective inhibition. *Cell Mol. Life Sci.* **2009**, *66*, 3685–3695. [[CrossRef](#)]
8. Chen, Q.; Fan, D.; Wang, G. Heteromeric Geranyl(geranyl) Diphosphate Synthase Is Involved in Monoterpene Biosynthesis in *Arabidopsis* Flowers. *Mol. Plant.* **2015**, *8*, 1434–1437. [[CrossRef](#)]
9. Wang, G.; Dixon, R.A. Heterodimeric geranyl(geranyl)diphosphate synthase from hop *Humulus lupulus* and the evolution of monoterpene biosynthesis. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9914–9919. [[CrossRef](#)]

10. Kwon, M.; Kwon, E.J.G.; Ro, D.K. Chapter Six - cis-Prenyltransferase and Polymer Analysis from a Natural Rubber Perspective. In *Methods in Enzymology*; O'Connor, S.E., Ed.; Elsevier: Amsterdam, The Netherlands, 2016; Volume 576, pp. 121–145. [[CrossRef](#)]
11. Kumar, S.; Stecher, G.; Li, M.; Niyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
12. Campbell, M.; Hahn, F.M.; Poulter, C.D.; Leustek, T. Analysis of the isopentenyl diphosphate isomerase gene family from *Arabidopsis thaliana*. *Plant. Mol. Biol.* **1998**, *36*, 323–328. [[CrossRef](#)]
13. Okada, K.; Kasahara, H.; Yamaguchi, S.; Kawaide, H.; Kamiya, Y.; Nojiri, H.; Yamane, H. Genetic evidence for the role of isopentenyl diphosphate isomerases in the mevalonate pathway and plant development in *Arabidopsis*. *Plant. Cell Physiol.* **2008**, *49*, 604–616. [[CrossRef](#)] [[PubMed](#)]
14. Phillips, M.A.; D'Auria, J.C.; Gershenzon, J.; Pichersky, E. The *Arabidopsis thaliana* Type I Isopentenyl Diphosphate Isomerases Are Targeted to Multiple Subcellular Compartments and Have Overlapping Functions in Isoprenoid Biosynthesis. *Plant. Cell* **2008**, *20*, 677–696. [[CrossRef](#)] [[PubMed](#)]
15. Sapir-Mir, M.; Mett, A.; Belausov, E.; Tal-Meshulam, S.; Frydman, A.; Gidoni, D.; Eyal, Y. Peroxisomal Localization of *Arabidopsis* Isopentenyl Diphosphate Isomerases Suggests That Part of the Plant Isoprenoid Mevalonic Acid Pathway Is Compartmentalized to Peroxisomes. *Plant. Physiol.* **2008**, *148*, 1219–1228. [[CrossRef](#)] [[PubMed](#)]
16. Vranová, E.; Hirsch-Hoffmann, M.; Gruissem, W. AtIPD: A Curated Database of *Arabidopsis* Isoprenoid Pathway Models and Genes for Isoprenoid Network Analysis. *Plant. Physiol.* **2011**, *156*, 1655–1660. [[CrossRef](#)]
17. Beck, G.; Coman, D.; Herren, E.; Ruiz-Sola, M.Á.; Rodríguez-Concepción, M.; Gruissem, W.; Vranová, E. Characterization of the GGPP synthase gene family in *Arabidopsis thaliana*. *Plant. Mol. Biol.* **2013**, *82*, 393–416. [[CrossRef](#)] [[PubMed](#)]
18. Bouvier, F.; Suire, C.; d'Harlingue, A.; Backhaus, R.A.; Camara, B. Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells. *Plant. J.* **2000**, *24*, 241–252. [[CrossRef](#)]
19. Joly, A.; Edwards, P.A. Effect of site-directed mutagenesis of conserved aspartate and arginine residues upon farnesyl diphosphate synthase activity. *J. Biol. Chem.* **1993**, *268*, 26983–26989.
20. Song, L.; Poulter, C.D. Yeast farnesyl-diphosphate synthase: Site-directed mutagenesis of residues in highly conserved prenyltransferase domains I and II. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3044–3048. [[CrossRef](#)]
21. Armenteros, J.J.A.; Salvatore, M.; Winther, O.; Emanuelsson, O.; von Heijne, G.; Elofsson, A.; Nielsen, H. Detecting Sequence Signals in Targeting Peptides Using Deep Learning. *Life Sci. Alliance* **2019**, *2*, e201900429. [[CrossRef](#)]
22. Hosfield, D.J.; Zhang, Y.; Dougan, D.R.; Broun, A.; Tari, L.W.; Swanson, R.V.; Finn, J. Structural Basis for Bisphosphonate-mediated Inhibition of Isoprenoid Biosynthesis. *J. Biol. Chem.* **2004**, *279*, 8526–8529. [[CrossRef](#)]
23. Wang, C.; Chen, Q.; Fan, D.; Li, J.; Wang, G.; Zhang, P. Structural Analyses of Short-Chain Prenyltransferases Identify an Evolutionarily Conserved GFPPS Clade in Brassicaceae Plants. *Mol. Plant.* **2016**, *9*, 195–204. [[CrossRef](#)] [[PubMed](#)]
24. Lange, B.M.; Ghassemian, M. Genome organization in *Arabidopsis thaliana*: A survey for genes involved in isoprenoid and chlorophyll metabolism. *Plant. Mol. Biol.* **2003**, *51*, 925–948. [[CrossRef](#)] [[PubMed](#)]
25. Keim, V.; Manzano, D.; Fernández, F.J.; Closa, M.; Andrade, P.; Caudepón, D.; Bortolotti, C.; Vega, M.C.; Arró, M.; Ferrer, A. Characterization of *Arabidopsis* FPS Isozymes and FPS Gene Expression Analysis Provide Insight into the Biosynthesis of Isoprenoid Precursors in Seeds. *PLoS ONE* **2012**, *7*, 49109. [[CrossRef](#)] [[PubMed](#)]
26. Running, M.P. The role of lipid post-translational modification in plant developmental processes. *Front. Plant. Sci.* **2014**, *5*, 50. [[CrossRef](#)]
27. Delourme, D.; Lacroute, F.; Karst, F. Cloning of an *Arabidopsis thaliana* cDNA coding for farnesyl diphosphate synthase by functional complementation in yeast. *Plant. Mol. Biol.* **1994**, *26*, 1867–1873. [[CrossRef](#)]
28. Cunillera, N.; Arró, M.; Delourme, D.; Karst, F.; Boronat, A.; Ferrer, A. *Arabidopsis thaliana* Contains Two Differentially Expressed Farnesyl-Diphosphate Synthase Genes. *J. Biol. Chem.* **1996**, *271*, 7774–7780. [[CrossRef](#)]

29. Masferrer, A.; Arro, M.; Manzano, D.; Schaller, H.; Fernandez-Busquets, X.; Moncalean, P.; Fernandez, B.; Cunillera, N.; Boronat, A.; Ferrer, A. Overexpression of *Arabidopsis thaliana* farnesyl diphosphate synthase (FPS1S) in transgenic *Arabidopsis* induces a cell death/senescence-like response and reduced cytokinin levels. *Plant. J.* **2002**, *30*, 123–132. [[CrossRef](#)]
30. Manzano, D.; Fernández-Busquets, X.; Schaller, H.; González, V.; Boronat, A.; Arró, M.; Ferrer, A. The metabolic imbalance underlying lesion formation in *Arabidopsis thaliana* overexpressing farnesyl diphosphate synthase (isoform 1S) leads to oxidative stress and is triggered by the developmental decline of endogenous HMGR activity. *Planta* **2004**, *219*, 982–992. [[CrossRef](#)]
31. Closa, M.; Vranová, E.; Bortolotti, C.; Bigler, L.; Arró, M.; Ferrer, A.; Grussem, W. The *Arabidopsis thaliana* FPP synthase isozymes have overlapping and specific functions in isoprenoid biosynthesis, and complete loss of FPP synthase activity causes early developmental arrest. *Plant. J.* **2010**, *63*, 512–525. [[CrossRef](#)]
32. Manzano, D.; Andrade, P.; Caudepón, D.; Altabella, T.; Arró, M.; Ferrer, A. Suppressing Farnesyl Diphosphate Synthase Alters Chloroplast Development and Triggers Sterol-Dependent Induction of Jasmonate- and Fe-Related Responses. *Plant. Physiol.* **2016**, *172*, 93–117. [[CrossRef](#)]
33. Bhatia, V.; Maisnam, J.; Jain, A.; Sharma, K.K.; Bhattacharya, R. Aphid-repellent pheromone *E*- β -farnesene is generated in transgenic *Arabidopsis thaliana* over-expressing farnesyl diphosphate synthase2. *Ann. Bot.* **2015**, *115*, 581–591. [[CrossRef](#)] [[PubMed](#)]
34. Okada, K.; Saito, T.; Nakagawa, T.; Kawamukai, M.; Kamiya, Y. Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in *Arabidopsis*. *Plant. Physiol.* **2000**, *122*, 1045–1056. [[CrossRef](#)] [[PubMed](#)]
35. Sakakibara, H.; Kasahara, H.; Ueda, N.; Kojima, M.; Takei, K.; Hishiyama, S.; Asami, T.; Okada, K.; Kamiya, Y.; Yamaya, T.; et al. *Agrobacterium tumefaciens* increases cytokinin production in plastids by modifying the biosynthetic pathway in the host plant. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9972–9977. [[CrossRef](#)] [[PubMed](#)]
36. Ruiz-Sola, M.Á.; Barja, M.V.; Manzano, D.; Llorente, B.; Schipper, B.; Beekwilder, J.; Rodriguez-Concepcion, M. A Single *Arabidopsis* Gene Encodes Two Differentially Targeted Geranylgeranyl Diphosphate Synthase Isoforms. *Plant. Phys.* **2016**, *172*, 1393–1402. [[CrossRef](#)] [[PubMed](#)]
37. Cunillera, N.; Boronat, A.; Ferrer, A. The *Arabidopsis thaliana* FPS1 Gene Generates a Novel mRNA That Encodes a Mitochondrial Farnesyl-diphosphate Synthase Isoform. *J. Biol. Chem.* **1997**, *272*, 15381–15388. [[CrossRef](#)]
38. Zhu, X.; Suzuki, K.; Saito, T.; Okada, K.; Tanaka, K.; Nakagawa, T.; Matsuda, H.; Kawamukai, M. Geranylgeranyl pyrophosphate synthase encoded by the newly isolated gene GGPS6 from *Arabidopsis thaliana* is localized in mitochondria. *Plant. Mol. Biol.* **1997**, *35*, 331–341. [[CrossRef](#)]
39. Lee, J.; Lee, H.; Kim, J.; Lee, S.; Kim, D.H.; Kim, S.; Hwang, I. Both the Hydrophobicity and a Positively Charged Region Flanking the C-Terminal Region of the Transmembrane Domain of Signal-Anchored Proteins Play Critical Roles in Determining Their Targeting Specificity to the Endoplasmic Reticulum or Endosymbiotic Organelles in *Arabidopsis* Cells. *Plant. Cell* **2011**, *23*, 1588. [[CrossRef](#)]
40. Vranová, E.; Kopcsayová, D.; Košuth, J.; Colinas, M. Mutant-Based Model of Two Independent Pathways for Carotenoid-Mediated Chloroplast Biogenesis in *Arabidopsis* Embryos. *Front. Plant. Sci.* **2019**, *10*, 1034. [[CrossRef](#)]
41. Hsieh, F.-L.; Chang, T.-H.; Ko, T.-P.; Wang, A.H.J. Structure and Mechanism of an *Arabidopsis* Medium/Long-Chain-Length Prenyl Pyrophosphate Synthase. *Plant. Phys.* **2011**, *155*, 1079–1090. [[CrossRef](#)]
42. Block, A.; Fristedt, R.; Rogers, S.; Kumar, J.; Barnes, B.; Barnes, J.; Elowsky, C.G.; Wamboldt, Y.; Mackenzie, S.A.; Redding, K.; et al. Functional Modeling Identifies Paralogous Solanesyl-diphosphate Synthases That Assemble the Side Chain of Plastoquinone-9 in Plastids. *J. Biol. Chem.* **2013**, *288*, 27594–27606. [[CrossRef](#)]
43. Ducluzeau, A.-L.; Wamboldt, Y.; Elowsky, C.G.; Mackenzie, S.A.; Schuurink, R.C.; Basset, G.J.C. Gene network reconstruction identifies the authentic *trans*-prenyl diphosphate synthase that makes the solanesyl moiety of ubiquinone-9 in *Arabidopsis*. *Plant. J.* **2012**, *69*, 366–375. [[CrossRef](#)] [[PubMed](#)]
44. Kim, E.-H.; Lee, Y.; Kim, H.U. Fibrillin 5 Is Essential for Plastoquinone-9 Biosynthesis by Binding to Solanesyl Diphosphate Synthases in *Arabidopsis*. *Plant. Cell* **2015**, *27*, 2956–2971. [[CrossRef](#)] [[PubMed](#)]
45. Jun, L.; Saiki, R.; Tatsumi, K.; Nakagawa, T.; Kawamukai, M. Identification and Subcellular Localization of Two Solanesyl Diphosphate Synthases from *Arabidopsis thaliana*. *Plant. Cell Physiol.* **2004**, *45*, 1882–1888. [[CrossRef](#)] [[PubMed](#)]

46. Hirooka, K.; Izumi, Y.; An, C.-I.; Nakazawa, Y.; Fukusaki, E.; Kobayashi, A. Functional analysis of two solanesyl diphosphate synthases from *Arabidopsis thaliana*. *Biosci. Biotechnol. Biochem.* **2005**, *69*, 592–601. [[CrossRef](#)] [[PubMed](#)]
47. Surmacz, L.; Swiezewska, E. Polyisoprenoids—Secondary metabolites or physiologically important superlipids? *Biochem. Biophys. Res. Commun.* **2011**, *407*, 627–632. [[CrossRef](#)] [[PubMed](#)]
48. Kera, K.; Takahashi, S.; Sutoh, T.; Koyama, T.; Nakayama, T. Identification and characterization of a *cis,trans*-mixed heptaprenyl diphosphate synthase from *Arabidopsis thaliana*. *FEBS J.* **2012**, *279*, 3813–3827. [[CrossRef](#)]
49. Surowiecki, P.; Onysk, A.; Manko, K.; Swiezewska, E.; Surmacz, L. Long-Chain Polyisoprenoids Are Synthesized by AtCPT1 in *Arabidopsis thaliana*. *Molecules* **2019**, *24*, 2789. [[CrossRef](#)]
50. Surmacz, L.; Plochocka, D.; Kania, M.; Danikiewicz, W.; Swiezewska, E. *cis*-Prenyltransferase AtCPT6 produces a family of very short-chain polyisoprenoids in planta. *Biochim. Biophys. Acta* **2014**, *1841*, 240–250. [[CrossRef](#)]
51. Akhtar, T.A.; Surowiecki, P.; Siekierska, H.; Kania, M.; Van Gelder, K.; Rea, K.A.; Virta, L.K.A.; Vatta, M.; Gawarecka, K.; Wojcik, J.; et al. Polyisoprenoids Are Synthesized by a Plastidial *cis*-Prenyltransferase and Influence Photosynthetic Performance. *Plant. Cell* **2017**, *29*, 1709–1725. [[CrossRef](#)]
52. Hooper, C.M.; Castleden, I.R.; Tanz, S.K.; Aryamanesh, N.; Millar, A.H. SUBA4: The interactive data analysis centre for *Arabidopsis* subcellular protein locations. *Nucleic Acids Res.* **2016**, *45*, 1064–1074. [[CrossRef](#)]
53. Zhu, X.; Suzuki, K.; Okada, K.; Tanaka, K.; Nakagawa, T.; Kawamukai, M.; Matsuda, K. Cloning and functional expression of a novel geranylgeranyl pyrophosphate synthase gene from *Arabidopsis thaliana* in *Escherichia coli*. *Plant. Cell Physiol.* **1997**, *38*, 357–361. [[CrossRef](#)] [[PubMed](#)]
54. Ruiz-Sola, M.Á.; Coman, D.; Beck, G.; Barja, M.V.; Colinas, M.; Graf, A.; Welsch, R.; Rütimann, P.; Bühlmann, P.; Bigler, L.; et al. *Arabidopsis* Geranylgeranyl Diphosphate Synthase 11 is a hub isozyme required for the production of most photosynthesis-related isoprenoids. *New Phytol.* **2016**, *209*, 252–264. [[CrossRef](#)] [[PubMed](#)]
55. Evidente, A.; Kornienko, A.; Lefranc, F.; Cimmino, A.; Dasari, R.; Evidente, M.; Mathieu, V.; Kiss, R. Sesterterpenoids with Anticancer Activity. *Curr. Med. Chem.* **2015**, *22*, 3502–3522. [[CrossRef](#)] [[PubMed](#)]
56. Huang, A.C.; Kautsar, S.A.; Hong, Y.J.; Medema, M.H.; Bond, A.D.; Tantillo, D.J.; Osbourn, A. Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6005–6014. [[CrossRef](#)]
57. Shao, J.; Chen, Q.-W.; Lv, H.-J.; He, J.; Liu, Z.-F.; Lu, Y.-N.; Liu, H.-L.; Wang, G.-D.; Wang, Y. (+)-Thalianatriene and (–)-Retigeranin B Catalyzed by Sesterterpene Synthases from *Arabidopsis thaliana*. *Org. Lett.* **2017**, *19*, 1816–1819. [[CrossRef](#)]
58. Nowicka, B.; Kruk, J. Occurrence, biosynthesis and function of isoprenoid quinones. *Biochim. Biophys. Acta* **2010**, *1797*, 1587–1605. [[CrossRef](#)]
59. Hirooka, K.; Bamba, T.; Fukusaki, E.; Kobayashi, A. Cloning and kinetic characterization of *Arabidopsis thaliana* solanesyl diphosphate synthase. *Biochem. J.* **2003**, *370*, 679–686. [[CrossRef](#)]
60. Mackenzie, S.A. Plant organellar protein targeting: A traffic plan still under construction. *Trends Cell. Biol.* **2005**, *15*, 548–554. [[CrossRef](#)]
61. Liang, P.-H. Reaction Kinetics, Catalytic Mechanisms, Conformational Changes, and Inhibitor Design for Prenyltransferases. *Biochemistry* **2009**, *48*, 6562–6570. [[CrossRef](#)]
62. Swiezewska, E.; Danikiewicz, W. Polyisoprenoids: Structure, biosynthesis and function. *Prog. Lipid. Res.* **2005**, *44*, 235–258. [[CrossRef](#)]
63. Oh, S.K.; Han, K.H.; Ryu, S.B.; Kang, H. Molecular Cloning, Expression, and Functional Analysis of a *cis*-Prenyltransferase from *Arabidopsis thaliana*. Implications in Rubber Biosynthesis. *J. Biol. Chem.* **2000**, *275*, 18482–18488. [[CrossRef](#)] [[PubMed](#)]
64. Cunillera, N.; Arro, M.; Fores, O.; Manzano, D.; Ferrer, A. Characterization of dehydrodolichyl diphosphate synthase of *Arabidopsis thaliana*, a key enzyme in dolichol biosynthesis. *FEBS Lett.* **2000**, *477*, 170–174. [[CrossRef](#)]
65. Grabińska, K.A.; Edani, B.H.; Park, E.J.; Kraehling, J.R.; Sessa, W.C. A conserved C-terminal RXG motif in the NgBR subunit of *cis*-prenyltransferase is critical for prenyltransferase activity. *J. Biol. Chem.* **2017**, *292*, 17351–17361. [[CrossRef](#)] [[PubMed](#)]

66. Qu, Y.; Chakrabarty, R.; Tran, H.T.; Kwon, E.-J.G.; Kwon, M.; Nguyen, T.-D.; Ro, D.-K. A lettuce (*Lactuca sativa*) homolog of human Nogo-B receptor interacts with *cis*-prenyltransferase and is necessary for natural rubber biosynthesis. *J. Biol. Chem.* **2015**, *290*, 1898–1914. [[CrossRef](#)]
67. Epping, J.; van Deenen, N.; Niephaus, E.; Stolze, A.; Fricke, J.; Huber, C.; Eisenreich, W.; Twyman, R.M.; Prüfer, D.; Schulze Gronover, C. A rubber transferase activator is necessary for natural rubber biosynthesis in dandelion. *Nat. Plants* **2015**, *1*, 15048. [[CrossRef](#)]
68. Brasher, M.I.; Surmacz, L.; Leong, B.; Pitcher, J.; Swiezewska, E.; Pichersky, E.; Akhtar, T.A. A two-component enzyme complex is required for dolichol biosynthesis in tomato. *Plant. J.* **2015**, *82*, 903–914. [[CrossRef](#)]
69. Addou, S.; Rentzsch, R.; Lee, D.; Orengo, C.A. Domain-Based and Family-Specific Sequence Identity Thresholds Increase the Levels of Reliable Protein Function Transfer. *J. Mol. Biol.* **2009**, *387*, 416–430. [[CrossRef](#)]
70. Chen, A.; Kroon, P.A.; Poulter, C.D. Isoprenyl diphosphate synthases: Protein sequence comparisons, a phylogenetic tree, and predictions of secondary structure. *Protein Sci.* **1994**, *3*, 600–607. [[CrossRef](#)]
71. Tarshis, L.C.; Proteau, P.J.; Kellogg, B.A.; Sacchettini, J.C.; Poulter, C.D. Regulation of product chain length by isoprenyl diphosphate synthases. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15018–15023. [[CrossRef](#)]
72. Ohnuma, S.-i.; Hirooka, K.; Tsuruoka, N.; Yano, M.; Ohto, C.; Nakane, H.; Nishino, T. A Pathway Where Polyprenyl Diphosphate Elongates in Prenyltransferase: insight into a common mechanism of chain length determination of prenyltransferases. *J. Biol. Chem.* **1998**, *273*, 26705–26713. [[CrossRef](#)]
73. Guo, R.-T.; Kuo, C.-J.; Chou, C.-C.; Ko, T.-P.; Shr, H.-L.; Liang, P.-H.; Wang, A.H.J. Crystal Structure of Octaprenyl Pyrophosphate Synthase from Hyperthermophilic *Thermotoga maritima* and Mechanism of Product Chain Length Determination. *J. Biol. Chem.* **2004**, *279*, 4903–4912. [[CrossRef](#)] [[PubMed](#)]
74. Chang, T.-H.; Guo, R.-T.; Ko, T.-P.; Wang, A.H.J.; Liang, P.-H. Crystal Structure of Type-III Geranylgeranyl Pyrophosphate Synthase from *Saccharomyces cerevisiae* and the Mechanism of Product Chain Length Determination. *J. Biol. Chem.* **2006**, *281*, 14991–15000. [[CrossRef](#)] [[PubMed](#)]
75. Gabelli, S.B.; McLellan, J.S.; Montalvetti, A.; Oldfield, E.; Docampo, R.; Amzel, L.M. Structure and mechanism of the farnesyl diphosphate synthase from *Trypanosoma cruzi*: Implications for drug design. *Proteins* **2006**, *62*, 80–88. [[CrossRef](#)] [[PubMed](#)]
76. Kavanagh, K.L.; Dunford, J.E.; Bunkoczi, G.; Russell, R.G.G.; Oppermann, U. The Crystal Structure of Human Geranylgeranyl Pyrophosphate Synthase Reveals a Novel Hexameric Arrangement and Inhibitory Product Binding. *J. Biol. Chem.* **2006**, *281*, 22004–22012. [[CrossRef](#)] [[PubMed](#)]
77. Kulmanov, M.; Khan, M.A.; Hoehndorf, R.; Wren, J. DeepGO: Predicting protein functions from sequence and interactions using a deep ontology-aware classifier. *Bioinformatics* **2018**, *34*, 660–668. [[CrossRef](#)] [[PubMed](#)]
78. Bepler, T.; Berger, B. Learning Protein Sequence Embeddings Using Information from Structure. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
79. Jin, H.; Song, Z.; Nikolau, B.J. Reverse genetic characterization of two paralogous acetoacetyl CoA thiolase genes in *Arabidopsis* reveals their importance in plant growth and development. *Plant. J.* **2012**, *70*, 1015–1032. [[CrossRef](#)]
80. Okada, K.; Ohara, K.; Yazaki, K.; Nozaki, K.; Uchida, N.; Kawamukai, M.; Nojiri, H.; Yamane, H. The *AtPPT1* gene encoding 4-hydroxybenzoate polyprenyl diphosphate transferase in ubiquinone biosynthesis is required for embryo development in *Arabidopsis thaliana*. *Plant. Mol. Biol.* **2004**, *55*, 567–577. [[CrossRef](#)]
81. Babiychuk, E.; Bouvier-Navé, P.; Compagnon, V.; Suzuki, M.; Muranaka, T.; Van Montagu, M.; Kushnir, S.; Schaller, H. Allelic mutant series reveal distinct functions for *Arabidopsis* cycloartenol synthase 1 in cell viability and plastid biogenesis. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3163–3168. [[CrossRef](#)]
82. Gutkowska, M.; Wnuk, M.; Nowakowska, J.; Lichocka, M.; Stronkowski, M.M.; Swiezewska, E. Rab geranylgeranyl transferase β subunit is essential for male fertility and tip growth in *Arabidopsis*. *J. Exp. Bot.* **2015**, *66*, 213–224. [[CrossRef](#)]
83. Jozwiak, A.; Gutkowska, M.; Gawarecka, K.; Surmacz, L.; Buczkowska, A.; Lichocka, M.; Nowakowska, J.; Swiezewska, E. POLYPRENOL REDUCTASE2 Deficiency Is Lethal in *Arabidopsis* Due to Male Sterility. *Plant. Cell* **2015**, *27*, 3336–3353. [[CrossRef](#)]
84. Motohashi, R.; Ito, T.; Kobayashi, M.; Taji, T.; Nagata, N.; Asami, T.; Yoshida, S.; Yamaguchi-Shinozaki, K.; Shinozaki, K. Functional analysis of the 37 kDa inner envelope membrane polypeptide in chloroplast biogenesis using a Ds-tagged *Arabidopsis* pale-green mutant. *Plant. J.* **2003**, *34*, 719–731. [[CrossRef](#)] [[PubMed](#)]

85. Yamasato, A.; Tanaka, R.; Tanaka, A. Loss of the N-terminal domain of chlorophyllide a oxygenase induces photodamage during greening of *Arabidopsis* seedlings. *BMC Plant Biol.* **2008**, *8*, 64. [[CrossRef](#)] [[PubMed](#)]
86. Ruppel, N.J.; Kropp, K.N.; Davis, P.A.; Martin, A.E.; Luesse, D.R. Mutations in Geranylgeranyl Diphosphate Synthase 1 affect chloroplast development in *Arabidopsis thaliana* (Brassicaceae) *Am. J. Bot.* **2013**, *100*, 2074–2084. [[CrossRef](#)] [[PubMed](#)]
87. van Schie, C.C.N.; Ament, K.; Schmidt, A.; Lange, T.; Haring, M.A.; Schuurink, R.C. Geranyl diphosphate synthase is required for biosynthesis of gibberellins. *Plant. J.* **2007**, *52*, 752–762. [[CrossRef](#)]
88. Zhang, H.; Ohyama, K.; Boudet, J.; Chen, Z.; Yang, J.; Zhang, M.; Muranaka, T.; Maurel, C.; Zhu, J.-K.; Gong, Z. Dolichol Biosynthesis and Its Effects on the Unfolded Protein Response and Abiotic Stress Resistance in *Arabidopsis*. *Plant. Cell* **2008**, *20*, 1879–1898. [[CrossRef](#)]
89. Wille, A.; Zimmermann, P.; Vranova, E.; Furholz, A.; Laule, O.; Bleuler, S.; Hennig, L.; Prelic, A.; von Rohr, P.; Thiele, L.; et al. Sparse graphical Gaussian modeling of the isoprenoid gene network in *Arabidopsis thaliana*. *Genome Biol.* **2004**, *5*, 92. [[CrossRef](#)]
90. Ghassemian, M.; Lutes, J.; Tepperman, J.M.; Chang, H.-S.; Zhu, T.; Wang, X.; Quail, P.H.; Markus Lange, B. Integrative analysis of transcript and metabolite profiling data sets to evaluate the regulation of biochemical pathways during photomorphogenesis. *Arch. Biochem. Biophys.* **2006**, *448*, 45–59. [[CrossRef](#)]
91. Meier, S.; Tzfadia, O.; Vallabhaneni, R.; Gehring, C.; Wurtzel, E. A transcriptional analysis of carotenoid, chlorophyll and plastidial isoprenoid biosynthesis genes during development and osmotic stress responses in *Arabidopsis thaliana*. *BMC Syst. Biol.* **2011**, *5*, 77. [[CrossRef](#)]
92. Nützmann, H.-W.; Scazzocchio, C.; Osbourn, A. Metabolic Gene Clusters in Eukaryotes. *Annu. Rev. Genet.* **2018**, *52*, 159–183. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).