The *cis*-prenyltransferase protein family in *Taraxacum koksaghyz*

Boje Müller[1,†] (iD), Eva Niephaus[1,†], Wolfgang Eisenreich[2], Jan Niklas Bröker[1], Richard M. Twyman[3], Dirk Prüfer[1,4] (iD) and Christian Schulze Gronover[1,*] (iD)

[1]*Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Johann-Krane-Weg 42, 48149 Münster, Germany,*
[2]*Structural Membrane Biochemistry, Bavarian NMR Center (BNMRZ), Technical University of Munich (TUM), Lichtenbergstr. 4, 85747 Garching, Germany,*
[3]*TRM Ltd, PO Box 493, Scarborough YO11 9FJ, UK, and*
[4]*Institute of Plant Biology and Biotechnology, University of Münster, Schlossplatz 8, 48143 Münster, Germany*

SUMMARY

The *cis*-prenyltransferase (*cis*PT) enzyme family is involved in diverse biological processes that require the synthesis of linear isoprenoid compounds. *Taraxacum koksaghyz* is a rubber-producing species and potential crop that has eight *cis*PT homologs (TkCPT1–8) but their distribution and functions are unclear. We compared the structural organization and sequence homology of the proteins, and defined two groups: TkCPT and TkCPT-like (TkCPTL) proteins that form heteromeric *cis*PT enzymes (TkCPT1–4), and TkCPT proteins that function as homomeric *cis*PTs (TkCPT5–8). We found that *TkCPT1* and *TkCPT2* are predominantly expressed in latex whereas *TkCPT3* and *TkCPT6-8* are predominantly expressed in leaves. *TkCPT4* was constitutively expressed in all *T. koksaghyz* tissues and *TkCPT5* mRNA was detected in flowers. The TkCPT1–4 subunits localized to the endoplasmic reticulum whereas TkCPT5–7 were located in chloroplasts. TkCPT1-4 interacted with TkCPTL1-2, forming heteromeric complexes that complemented yeast lacking *cis*PT. Homomeric TkCPT6 could also complement yeast lacking *cis*PT but we observed no *cis*PT activity for TkCPT5, TkCPT7, or TkCPT8 in yeast functional complementation assays. TkCPT1/TkCPTL1 and TkCPT2/TkCPTL1 expressed in yeast produced extra-long-chain polyisoprenes, whereas TkCPT3/TkCPTL1 and TkCPT4/TkCPTL1 produced long-chain dolichols and polyisoprenes, TkCPT5 and TkCPT6 produced medium-chain polyisoprenes, and TkCPT7 and TkCPT8 catalyzed the formation of nerol. The complexity of *cis*PT proteins in *T. koksaghyz* suggests that they synthesize different metabolites in a tissue-specific manner, and thus play distinct roles in isoprenoid metabolism. This is the first comprehensive analysis of the localization, interactions, and products of the entire *T. koksaghyz cis*PT family *in vivo*, also revealing a novel pentaprenol found specifically in flowers.

Keywords: linear isoprenoids, protein complex, Russian dandelion, dolichol, natural rubber.

INTRODUCTION

Isoprenoids are a functionally and structurally diverse group of metabolites derived from the five-carbon building block isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP). In plants, these C_5 units are synthesized via the cytosolic mevalonate and plastidial methylerythritol phosphate pathways, respectively (Vranová et al., 2013). For the biosynthesis of linear isoprenoids, a defined number of IPP molecules must undergo sequential condensation reactions with an allylic starter molecule, such as DMAPP or farnesyl diphosphate (FPP), in a reaction catalyzed by prenyltransferases (Grabińska et al., 2016). These are defined as *trans*-prenyltransferases (*trans*PTs) or *cis*-prenyltransferases (*cis*PTs) according to the stereochemical configuration of their products. Although *trans*PTs and *cis*PTs catalyze similar condensation reactions, their structures, and enzymatic mechanisms are distinct (Grabińska et al., 2016; Kopcsayová & Vranová, 2019).

The *cis*PTs have five conserved regions that contain amino acid residues necessary for catalytic activity and substrate binding (Kharel & Koyama, 2003). The products

synthesized by *cis*PTs vary significantly in chain length and range from C_{10} compounds such as neryl diphosphate (NPP) to natural rubber (NR), which mainly consists of high-molecular-weight poly(*cis*-1,4-isoprene) with more than 10 000 units. The *cis*PTs are thus assigned to categories based on product chain length: short-chain ($<C_{25}$), medium-chain (C_{25}–C_{60}), long-chain ($>C_{60}$), and extra-long-chain ($>C_{10\,000}$) *cis*PTs (Grabińska et al., 2016).

The Russian dandelion (*Taraxacum koksaghyz*) produces large amounts of NR, which is synthesized by the extra-long-chain *cis*PTs TkCPT1 and TkCPT2 as part of the rubber transferase complex. However, other plants produce diverse and versatile *cis*PTs spanning all four of the chain length categories (Akhtar et al., 2013; Kopcsayová & Vranová, 2019). We therefore set out to identify the complete set of *cis*PT genes in the *T. koksaghyz* genome and determine the corresponding mRNA and protein expression profiles. We also investigated the subcellular compartmentalization of the enzymes and their ability to complement yeast strains with a deficiency in dolichol biosynthesis as a means to investigate their physiological functions in primary and secondary metabolism.

Most eukaryotic *cis*PTs produce polyisoprenyl diphosphates ($>C_{50}$). These are converted into dolichols or polyisoprenes, which differ in the hydrogenation status of the terminal (α) residue. Dolichols possess a hydrogenated terminal α-isoprene residue. Phosphorylated dolichols are indispensable carrier lipids for glycans during protein glycosylation in yeast, mammals, and plants (Surmacz & Swiezewska, 2011). Conversely, polyisoprenes possess an unsaturated α-isoprene residue and their physiological functions are poorly understood (Hofmann, 2017). For example, polyisoprenes produced by chloroplast-specific *cis*PTs in Arabidopsis (*Arabidopsis thaliana*) and tomato (*Solanum lycopersicum*) influence photosynthetic performance (Akhtar et al., 2017; van Gelder et al., 2018). Few short-chain *cis*PTs have been characterized thus far, but examples include glandular trichome-specific *cis*PTs in solanaceous species that synthesize the C_{10} molecule NPP and the C_{15} molecule *cis,cis*-FPP, which are allyl precursors for the synthesis of monoterpenes and sesquiterpenes (Sallaud et al., 2009; Schilmiller et al., 2009).

Many *cis*PTs are homomeric enzymes but others function as heteromeric complexes. Prokaryotic and most eukaryotic short-chain and medium-chain *cis*PTs are represented by a single protein that forms an autonomous homodimer (Chen et al., 2020; Grabińska et al., 2016). For example, the protist *Giardia lamblia* possesses a homomeric undecaprenyl diphosphate synthase (UPPS) closely related to prokaryotic *cis*PTs, and this produces the C_{55} compound undecaprenyl diphosphate (UPP), which is further reduced to dolichol (Grabińska et al., 2010). In contrast, eukaryotic long-chain *cis*PTs are generally complexes containing two essential subunits, a *cis*PT (CPT) and a

CPT-like (CPTL) protein. The only known exception is the homomeric Arabidopsis *cis*PT AtCPT3 (At2g23410) (Surowiecki et al., 2019). CPTL subunits are closely related to functional *cis*PT sequences but they lack conserved amino acids in the *cis*PT homology domains that are necessary for *cis*PT catalytic activity. Instead, CPTL proteins contain a C-terminal R*X*G motif, which they share with homomeric *cis*PTs, and this is also necessary for *cis*PT activity (Grabińska et al., 2017). Two *T. koksaghyz* CPTL proteins have been characterized thus far (TkCPTL1 and TkCPTL2). TkCPTL1 is known to interact with TkCPT1 and TkCPT2 to form a subunit of the rubber transferase complex (Niephaus et al., 2019).

NR *cis*PTs have also been found in other NR producers, such as the rubber tree (*Hevea brasiliensis*; HbHRT1–2/HbHRBP), guayule (*Parthenium argentatum*; PaCPT1–3/PaCBP), lettuce (*Lactuca sativa*; LsCPT3/LsCPTL2), and another dandelion species, *Taraxacum brevicorniculatum* (TbCPT1–3/TbRTA) (Epping et al., 2015; Lakusta et al., 2019; Qu et al., 2015; Yamashita et al., 2016). In *T. brevicorniculatum*, a close relative of *T. koksaghyz*, TbCPT2, TbCPT3, and the rubber transferase activator (TbRTA) are identical to TkCPT1, TkCPT2, and TkCPTL1 in *T. koksaghyz*. However, NR levels are significantly lower in *T. brevicorniculatum* plants due to the low-level expression of *TbCPT1–3* (Post et al., 2012). Numerous representative heteromeric CPT/CPTL complexes required for long-chain dolichol biosynthesis have also been described in other species (Brasher et al., 2015; Harrison et al., 2011; Kwon et al., 2016). Examples include HsDHDDS/HsNgBR in humans (*Homo sapiens*), Rer2/Nus1 and Srt1/Nus1 in yeast (*Saccharomyces cerevisiae*), AtCPT1 (At2g17570)/AtLEW1 in Arabidopsis, and SlCPT3/SlCBP in tomato. However, *cis*PTs responsible for dolichol biosynthesis have yet to be found in *T. koksaghyz*, and downregulating the TkCPTL subunits did not affect the dolichol content (Niephaus et al., 2019). Here we characterized all eight *cis*PTs in *T. koksaghyz* (TkCPT1–8) to better understand their expression, subcellular localization, interactions with TkCPTL proteins, and physiological functions. As the first comprehensive analysis of the entire *T. koksaghyz* *cis*-PT family *in vivo*, our work provides unprecedented insight into the regulation of isoprenoid metabolism in this NR-producing species.

RESULTS

Identification of *cis*PT genes in *T. koksaghyz*

We screened RNA-Seq data and the published *T. koksaghyz* genome sequence, revealing six new *cis*PT genes that we named *TkCPT3–8* (Table 1) in addition to the known genes *TkCPT1* and *TkCPT2* (identical to *TbCPT2* and *TbCPT3*) encoding subunits of the rubber transferase complex (Epping et al., 2015; Lin et al., 2017; Post et al., 2012;

Table 1 DNA and protein sequence information corresponding to the eight TkCPTs

Name	cDNA (bp)	Tissue[a]	Protein (aa)	MW (kDa)	Transit peptide (likelihood)[b]	References
TkCPT1	927	Latex	309	35	No (0.998)	Schmidt et al. (2010)
TkCPT2	906	Latex	302	34.3	No (0.998)	Schmidt et al. (2010)
TkCPT3	822	Leaf	273	31.8	No (0.989)	This study
TkCPT4	840	Latex	279	32.0	No (0.794)	This study
TkCPT5	873	Flower	290	33.1	ChlTP (0.717)	This study
TkCPT6	873	Leaf	290	32.6	No/ChlTP (0.485/0.452)	This study
TkCPT7	933	Latex	310	35.0	ChlTP (0.692)	This study
TkCPT8	903	Peduncle	300	34.3	ChlTP (0.871)	This study

[a]Tissue used for amplification of the coding sequence.
[b]Predicted signal peptide (ChlTP, chloroplast transit peptide).

Schmidt et al., 2010). Gene-specific primers were designed and the full-length coding sequences of all eight genes were amplified using total RNA isolated from different *T. koksaghyz* tissues.

The deduced TkCPT protein sequences (including TkCPT1 and TkCPT2) were aligned with the human CPT subunit (HsDHDDS), the *Micrococcus luteus cis*PT homodimer (MlUPPS), and the human CPTL subunit (HsNgBR), revealing that all eight dandelion TkCPTs contained the five highly conserved regions that characterize *cis*PTs (Figure 1a). TkCPT5–7 shared the C-terminal R*X*G motif with HsNgBR and MlUPPS. This motif appears in homomeric *cis*PTs as well as CPTL subunits and is necessary for catalytic activity (Grabińska et al., 2017). We omitted other plant sequences from the alignment to better highlight the taxonomic differences and emphasize the sequence motifs. We included the *M. luteus* and human sequences because these are among the best-characterized prokaryotic and eukaryotic CPTs and CPTLs. Additional proteins were included to represent other phylogenetic groups, allowing us to examine the relationship between TkCPTs and *cis*PTs from species representing diverse taxa (Figure 1b).

The CPTL proteins, which are similar enough to *cis*PTs for inclusion in the phylogenetic tree, formed a separate group from the *cis*PTs represented by two clusters: the CPT subunits of heteromeric *cis*PTs and the homomeric *cis*PTs. TkCPT3 and TkCPT4 clustered with the CPT subunits of heteromeric *cis*PTs. Most of these subunits depend on a CPTL subunit for *cis*PT activity. We found that TkCPT3 is closely related to CPT components (HbHRT1–2, LsCPT3, PaCPT3, and TkCPT1–2) that are thought to be part of the rubber transferase complex (Asawatreratanakul et al., 2003; Lakusta et al., 2019; Niephaus et al., 2019; Post et al., 2012; Qu et al., 2015). However, rubber transferase activity (the synthesis of extra-long-chain polymers) has not been observed for any of these enzymes *in vitro*. TkCPT4 grouped with other plant *cis*PTs, among which lettuce LsCPT1, tomato SlCPT3, and Arabidopsis AtCPT1 (At2g17570) produce long-chain polyisoprenoids (Akhtar

et al., 2013; Kwon et al., 2016; Qu et al., 2015). TkCPT5–8 clustered with homomeric plant *cis*PTs, which do not require other subunits for *cis*PT activity and mainly produce short-chain and medium-chain isoprenyl diphosphates. The conserved glycine at position 130 was replaced by methionine in TkCPT8, suggesting that TkCPT1–7 prefer the substrate *trans,trans*-FPP (C_{15}) whereas TkCPT8 prefers the C_5 precursor DMAPP (Kutsukawa et al., 2022). Moreover, the conserved tryptophan at position 337 in TkCPT1–6 may indicate these enzymes produce long-chain polyisoprenes, whereas TkCPT7 and TkCPT8 (with alanine and phenylalanine at position 337, respectively) are more likely to synthesize short-chain products (Kutsukawa et al., 2022).

Spatial expression profiles of *TkCPT1–8*

Given that *cis*PT genes in other plant species exhibit diverse spatial expression profiles, we measured the level of *TkCPT1–8* mRNA in different tissues of *T. koksaghyz*. Total RNA isolated from the latex, roots, leaves, peduncles, and flowers of 12-week-old wild-type *T. koksaghyz* plants was amplified by quantitative real-time PCR (qRT-PCR). *TkCPT1* and *TkCPT2* mRNAs were most abundant in the latex (Figure 2a,b) whereas *TkCPT3* and *TkCPT6–8* mRNAs were most abundant in the leaves (Figure 2c,f–h). *TkCPT3* was expressed at a higher level than *TkCPT6–8*. *TkCPT4* mRNA levels were low in all tissues (Figure 2d) and *TkCPT5* mRNA was only detected at high levels in flowers (Figure 2e).

Subcellular localization of TkCPTs

The biosynthesis of isoprenoids is compartmentalized in the cell, so the subcellular localization of *cis*PTs can indicate which pathways are involved, and therefore the most likely physiological functions. The prediction tool TargetP-2.0 identified a putative N-terminal chloroplast transit peptide (ChlTP) on TkCPT5, TkCPT7, and TkCPT8, whereas no transit peptides were predicted (likelihood >0.6) for TkCPT1–4 (Table 1). For TkCPT6, the prediction was

(a)

(b)

ambiguous (the presence or absence of a ChlTP both gave a likelihood of approximately 0.5). However, given that TkCPT6 clustered with plastid-localized *cis*PTs from other organisms (SlCPT4, SlCPT5, and AtCPT4/At5g58770) and those with a predicted ChlTP (LsCPT2), it is most likely that TkCPT6 is resident in the chloroplast. TkCPT1–3 featured a C-terminal di-lysine motif (K(X)KK), which can cause the retention of proteins in the endoplasmic reticulum (ER) (Benghezal et al., 2000). The same motif was present in TkCPT4, but its position differed from the canonical −3, −4 or −3, −5 relative to the C-terminus. None of the TkCPTs appeared to feature a transmembrane domain.

To determine the subcellular localization of TkCPTs *in planta*, each protein was transiently expressed in *Nicotiana benthamiana* leaves as a C-terminal fusion with the fluorophore Cerulean. When co-expressing Cerulean-tagged TkCPT1–4 and the N-terminus of the Arabidopsis ER marker protein CYP51G tagged with mRFP, we observed an overlapping fluorescence signal suggesting TkCPTs 1–4 are associated with the ER (Figure 3). As a control, we expressed Cerulean fusions with untagged mRFP, which is localized in the cytosol. The spatial analysis of fluorescence revealed that the untagged mRFP extended to a wider region than the CPT-Cerulean fusion proteins, indicating that the CPTs are more likely to be attached to the ER than distributed in the cytosol (shown representatively for CPT1 in Figure 3). However, we cannot exclude the possibility that some CPT is also present in the cytosol. The co-expression of Cerulean-tagged TkCPT5–7 with the tobacco (*Nicotiana tabacum*) ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) ChlTP linked to the fluorophore Venus as a plastid stromal marker also revealed co-localization (Figure 4). Chlorophyll autofluorescence was captured to confirm the location of chloroplasts. We did not detect a fluorescence signal for TkCPT8-Cerulean.

TkCPT–TkCPTL protein interactions

For TkCPT1 and TkCPT2 (or the identical proteins TbCPT2 and TbCPT3), we previously reported bimolecular fluorescence complementation (BiFC) experiments showing that these proteins interact with TkCPTL1 (or its ortholog TbRTA) and TkCPTL2 to form a *cis*PT complex (Epping et al., 2015; Niephaus et al., 2019). Because TkCPT3 and TkCPT4 cluster with other CPT subunits to form

heteromeric complexes, we expected these proteins to interact with the TkCPTL subunits. We therefore carried out BiFC experiments in *N. benthamiana* leaves using the split-mRFP system (Figure 5). We did not expect the plastid-localized TkCPT5–8 to interact with TkCPTLs, but we nevertheless tested the corresponding BiFC combinations as confirmation (Figure S1). The co-expression of TkCPT3 or TkCPT4 fused to NmRFP with TkCPTL1 or TkCPTL2 fused to CmRFP yielded strong red fluorescence, indicating that TkCPT3 interacted with TkCPTL1/TkCPTL2, and that TkCPT4 interacted with TkCPTL1/TkCPTL2. As expected, TkCPT5–8 did not interact with TkCPTL1/TkCPTL2. All other possible combinations of N-terminal and C-terminal fusions of the different split-mRFP tags were tested with the TkCPT and TkCPTL proteins and showed the same interaction patterns. As a control, NmRFP and CmRFP were tested with the corresponding counterparts fused to CPT/CPTL proteins but no red fluorescence was observed (Figure 5).

Functional analysis of *T. koksaghyz cis*PTs

We previously showed, using a yeast complementation assay, that TkCPT1 or TkCPT2 (TbCPT2 or TbCPT3) alone could not suppress the growth defects of yeast strains deficient in dolichol biosynthesis, but required TkCPTL1 (TbRTA) or TkCPTL2 for *cis*PT activity (Epping et al., 2015; Niephaus et al., 2019). However, the combination of TkCPT2 and TkCPTL2 was insufficient for functional complementation. Therefore, we tested the other TkCPT–TkCPTL complexes (and TkCPTs alone) to determine which are functional *cis*PT enzymes. We used two *S. cerevisiae* deletion strains deficient for dolichol biosynthesis: the *rer2Δsrt1Δ* strain, which only contains the yeast CPTL subunit (Nus1) but no CPT subunit (Rer2 and Srt1), and the *nus1Δ rer2Δsrt1Δ* strain, which lacks the yeast *cis*PT heteromers Nus1/Rer2 and Nus1/Srt1 (Grabińska et al., 2010; Park et al., 2014). To ensure growth, both strains carried a plasmid containing the *G. lamblia* gene encoding a homomeric *cis*PT and the *URA3* gene encoding orotidine 5′-phosphate decarboxylase, which converts 5-fluoroorotic acid (5-FOA) into the toxic compound 5-fluorouracil. Accordingly, only yeast cells that lose the plasmid and are transformed with a replacement functional *cis*PT can grow on a medium containing 5-FOA but not on a medium without uracil. TkCPT3 and TkCPT4 allowed growth in the
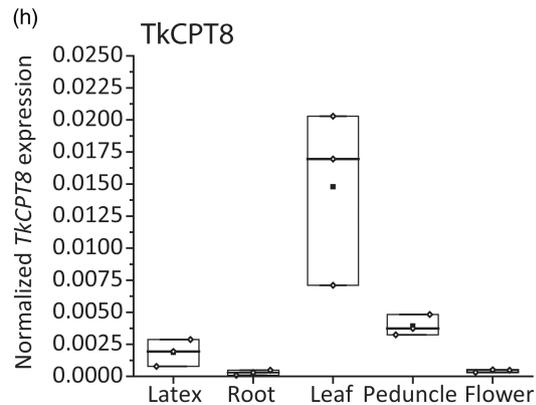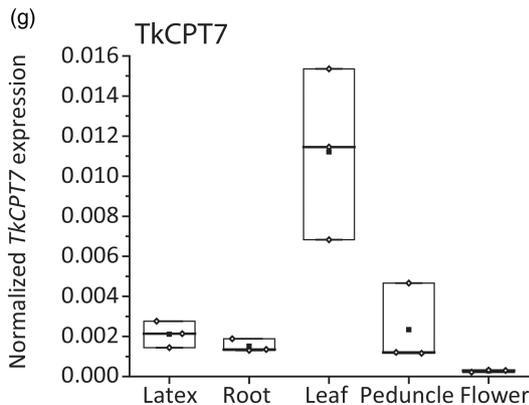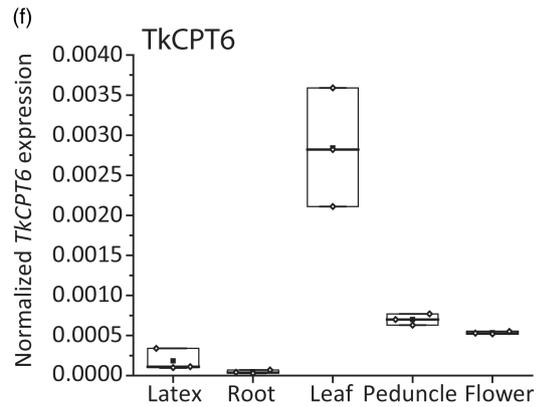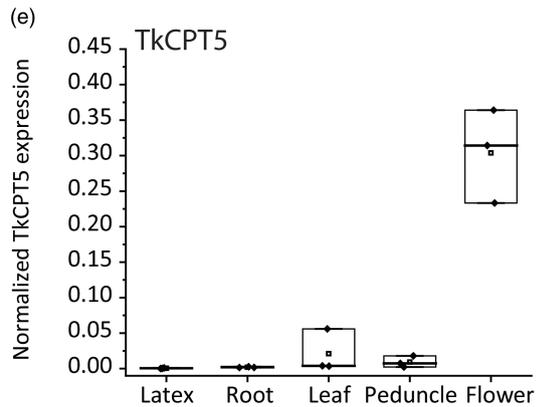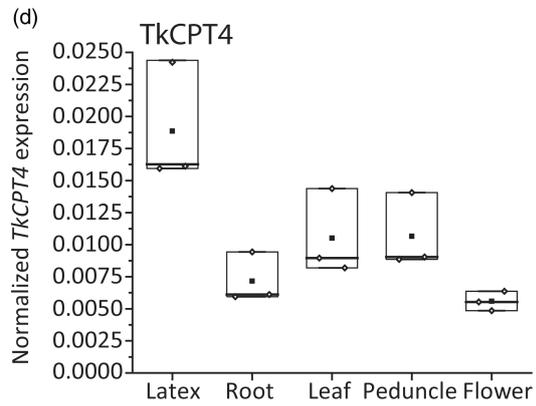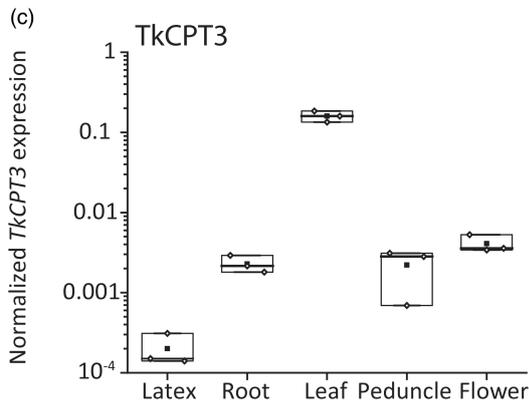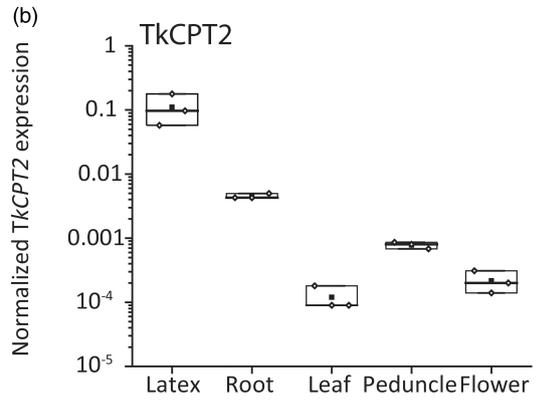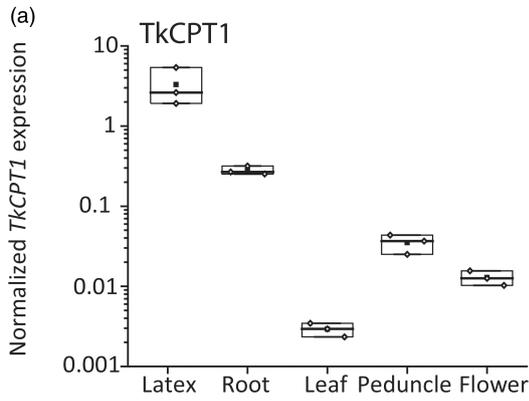
Figure 2. Spatial expression profiles of *TkCPTs*.
(a–h) The mRNA levels of (a) *TkCPT1*, (b) *TkCPT2*, (c) *TkCPT3*, (d) *TkCPT4*, (e) *TkCPT5*, (f) *TkCPT6*, (g) *TkCPT7*, and (h) TkCPT8 were measured by qRT-PCR. Data points represent pools containing cDNA from four 12-week-old wild-type *Taraxacum koksaghyz* plants. The box plots enclose values from the first to the third quartile. Filled squares represent the mean, and the horizontal line shows the median.
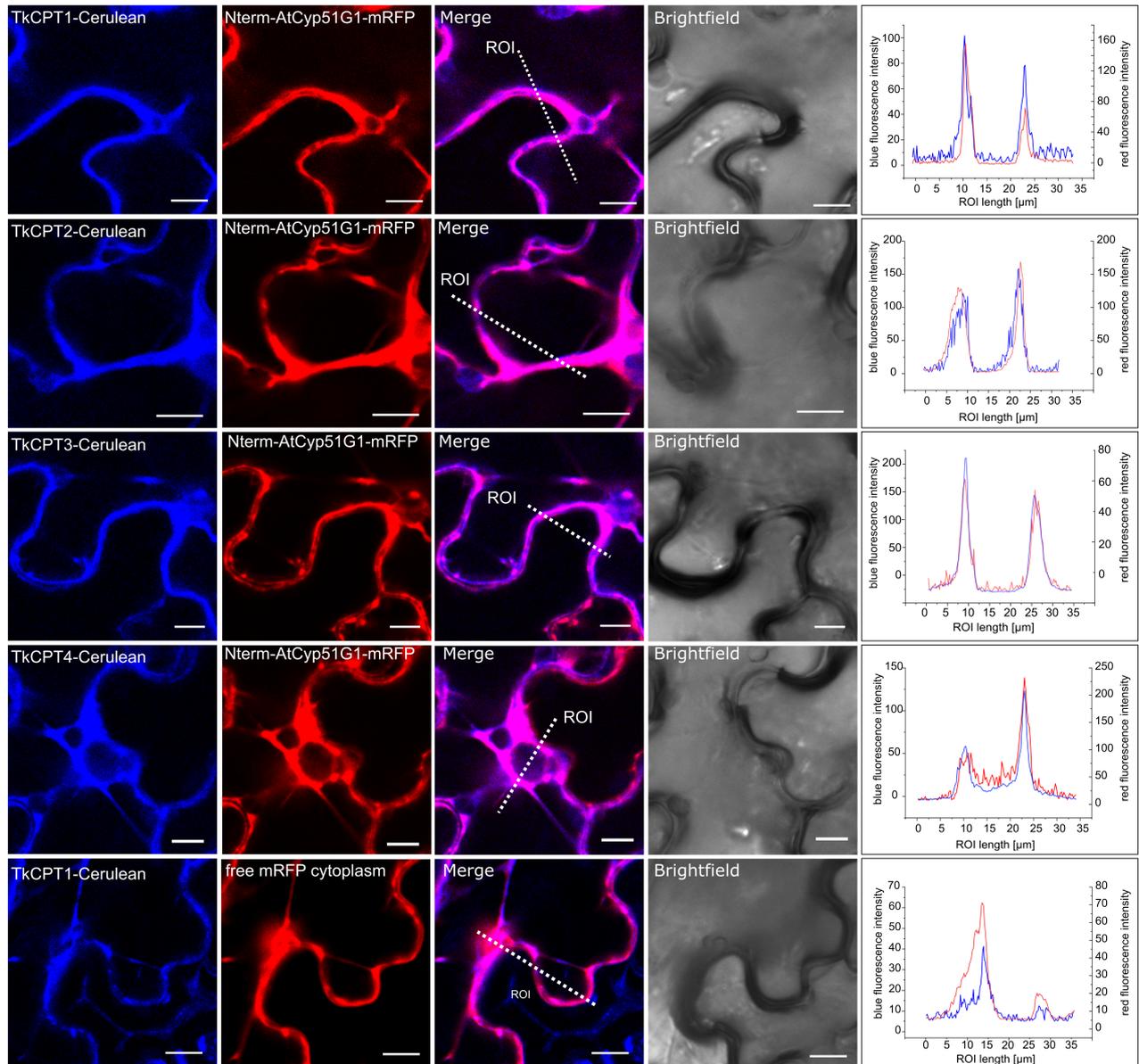
Figure 3. Subcellular localization of the catalytic subunits of heteromeric TkCPTs.
TkCPT-Cerulean fusion proteins were co-expressed with the mRFP-tagged 40-amino-acid N-terminus of the Arabidopsis ER marker protein CYP51G1 in *Nicotiana benthamiana* leaf epidermal cells. Fluorescence was detected by CLSM. Scale bar = 10 μm. ROI, region of interest.

presence of 5-FOA when co-expressed with TkCPTL1 or TkCPTL2 in the triple-deletion strain *nus1Δrer2Δsrt1Δ*, suggesting that both can form an active *cis*PT complex when combined with either of the TkCPTL proteins (Figure 6a,b). The loss of the *URA3* plasmid was confirmed by the inability of the yeast strains to grow on a medium without uracil.

However, each of the TkCPTs alone was insufficient to complement the *nus1Δrer2Δsrt1Δ* strain. Moreover, TkCPT3 and TkCPT4 were nonfunctional when combined with Nus1 because the double-deletion strain *rer2Δsrt1Δ* transformed with TkCPT3 or TkCPT4 alone did not grow on a medium containing 5-FOA. Western blot analysis
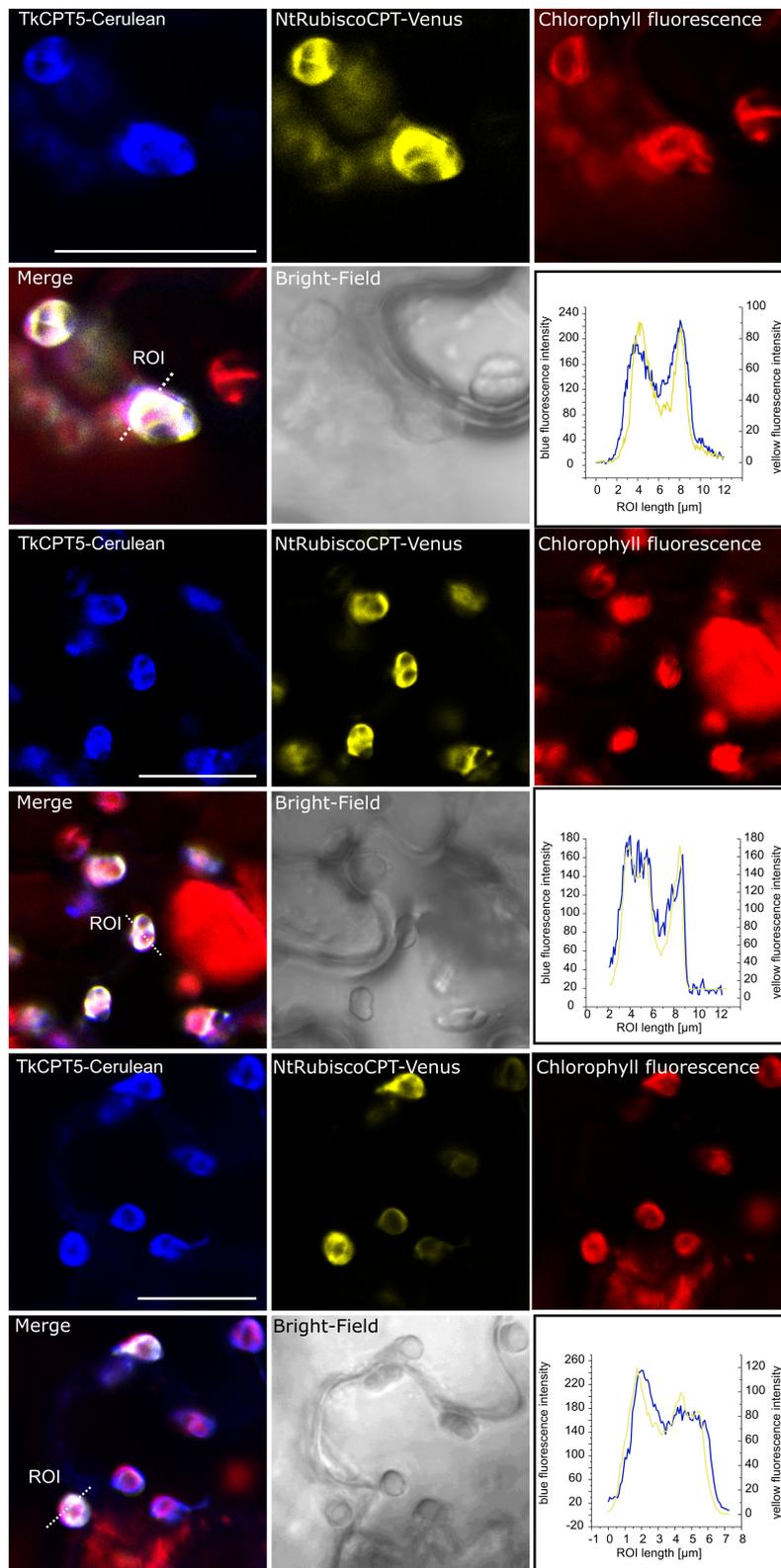
Figure 4. Subcellular localization of homomeric TkCPTs.
TkCPT-Cerulean fusion proteins were co-expressed with the 78-amino-acid chloroplast transit peptide of tobacco RuBisCO (NtRuBisCO-ChlTP) fused to Venus in *Nicotiana benthamiana* leaf epidermal cells. Fluorescence was detected by CLSM. Scale bar = 15 μm. ROI, region of interest.

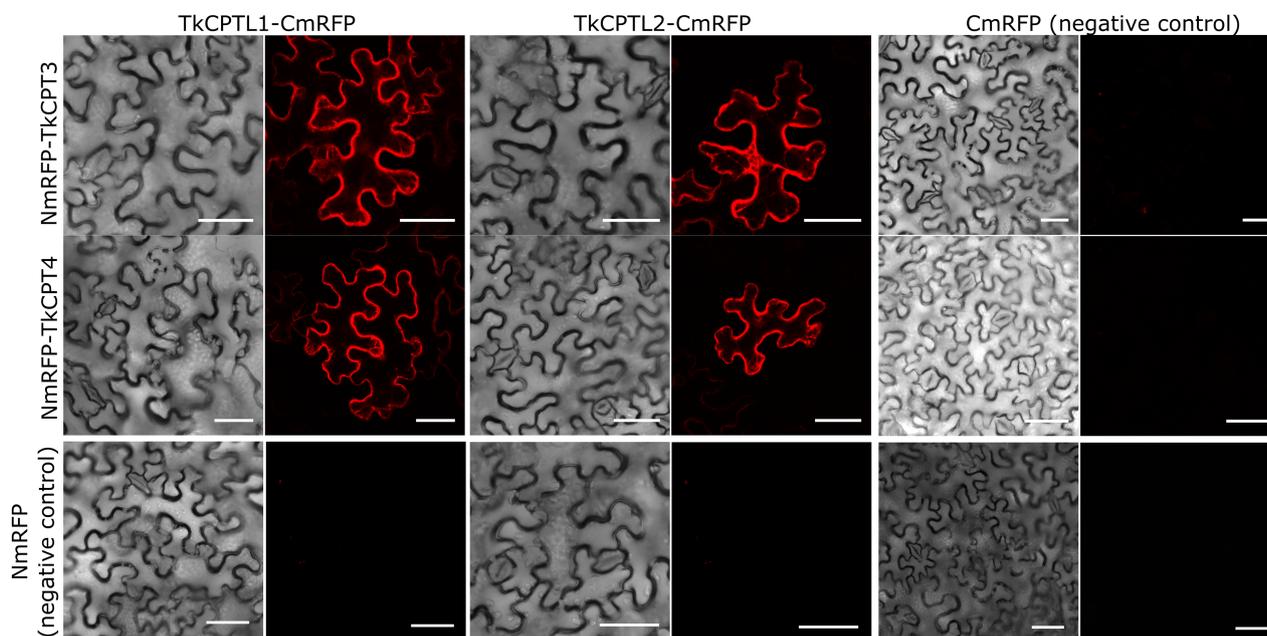**Figure 5.** BiFC analysis of heteromeric TkCPT subunits.
Nicotiana benthamiana leaf epidermal cells were infiltrated with pairwise combinations of NmRFP-TkCPT fusions and TkCPTL1-CmRFP or TkCPTL2-CmRFP. The NmRFP-TkCPTs were expressed with free CmRFP and the TkCPTL-CmRFPs with free NmRFP as negative controls. Fluorescence was detected by CLSM. Scale bar = 50 µm.

confirmed the expression of HA-TkCPT3, HA-TkCPT4, 3xcMyc-TkCPTL1, and 3xcMyc-TkCPTL2 in all transformed *S. cerevisiae* strains (Figure 6c,d).

No growth was observed for either yeast strain on medium containing 5-FOA when TkCPT5, TkCPT7, or TkCPT8 were expressed alone or co-expressed with TkCPTL1 or TkCPTL2, indicating the absence of functional complementation (Figure 7a,e,g). Western blot analysis confirmed the presence of HA-tagged TkCPT and 3xcMyc-tagged TkCPTL proteins, ruling out insufficient protein expression as an explanation (Figure 7b,f,h). In contrast, TkCPT6 appeared to confer complete *cis*PT activity (all the transformed yeast strains grew on 5-FOA medium but not on medium lacking uracil), suggesting TkCPT6 is a homomeric *cis*PT (Figure 7c). The expression of TkCPT6 was also confirmed by Western blot (Figure 7d).

Proteins were extracted from yeast strains expressing TkCPT3/TkCPTL2, TkCPT4/TkCPTL2, and TKCPT6 that were grown in the presence of 5-FOA, as well as original strains *rer2Δsrt1Δ* and *nus1Δrer2Δsrt1Δ* that were grown in SD-Ura. The proteins were fractionated by SDS–PAGE and probed with an antibody against the enzyme carboxypeptidase Y (CPY) encoded by the *PRC1* gene. This protein is hypoglycosylated in the *rer2Δ* mutant, and full glycosylation is not restored when the growth of the strain is complemented by the Arabidopsis protein AtCPT6 (Surmacz et al., 2014). Our Western blots showed that full glycosylation was restored to the original strains *rer2Δsrt1Δ* and

nus1Δrer2Δsrt1Δ when growth was complemented by the *Giardia lamblia cis*PT. Full glycosylation was also restored by TkCPT3/TkCPTL2 or TkCPT4/TkCPTL2 but not by TkCPT6, the latter resulting in lower-molecular-weight bands representing incomplete glycosylation (Figure 8).

For the assignment of TkCPTs to specific products, the corresponding genes were expressed in an engineered yeast strain and the extracted products were analyzed by thermal field flow fractionation (ThFFF), which uses an evaporative light scattering detector (ELSD) to determine the total mass of the eluted compound and a multi-angle light scattering (MALS) detector to determine its molecular weight. The formation of long-chain isoprenoids was confirmed in cells expressing TkCPT1/TkCPTL1 and TkCPT2/TkCPTL1 (Figure S2). The cells expressing TkCPT1/TkCPTL1 showed elevated MALS and ELSD signals 24–40 min after injection of the corresponding *n*-hexane extract (Figure S2a). These signals were not observed in the control yeast cells (Figure S2c). In cells expressing TkCPT2/TkCPTL1, only the MALS signal (Figure S2b) appeared stronger in the corresponding timeframe relative to the control (Figure S2d). Computational analysis revealed higher yields and higher product molecular weights ($>10^6$ g mol^{-1}) for TkCPT1/TkCPTL1 compared to TkCPT2/TkCPTL1 (Table 2). However, we also monitored a MALS signal in the control that might be attributed to the endogenous yeast dolichols, and therefore the monitored signals could not clearly prove a higher abundance of
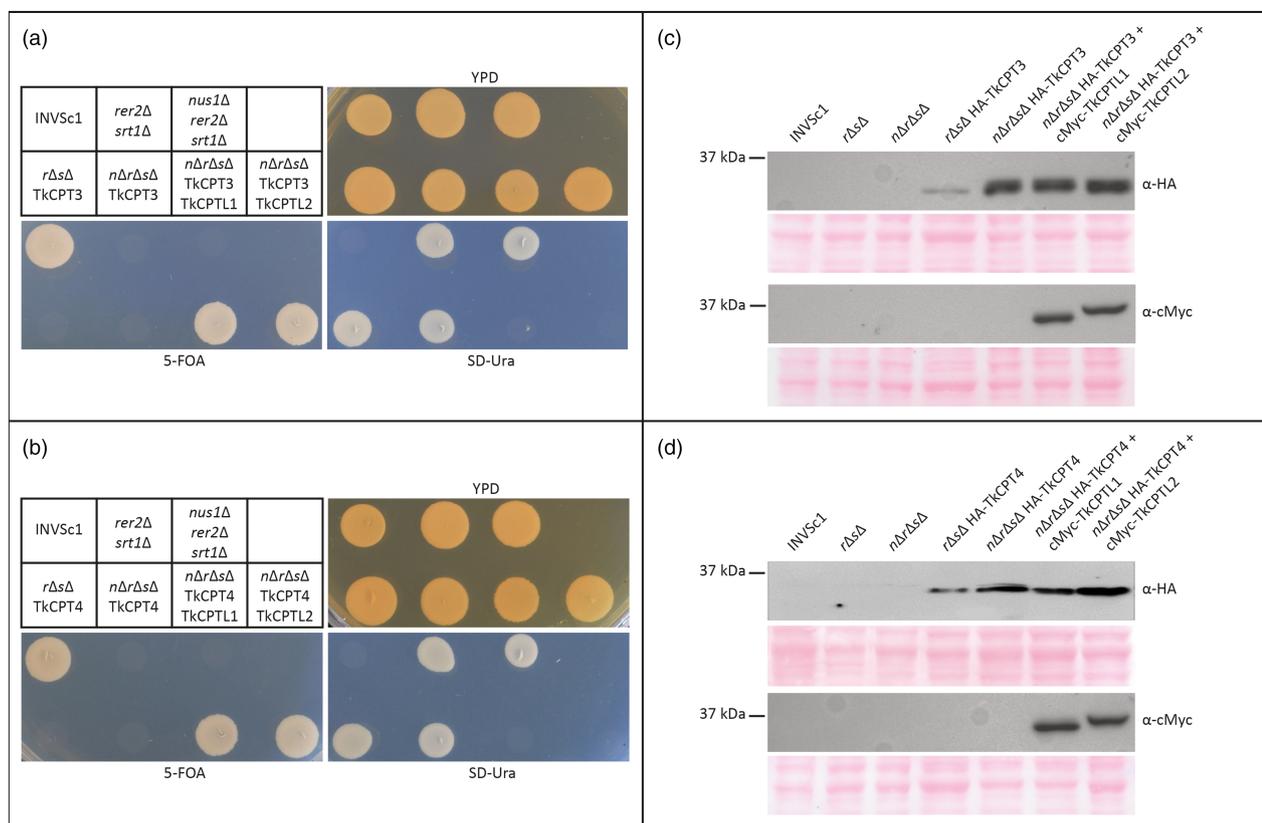
Figure 6. Functional complementation assay for putative heteromeric TkCPT subunits in *Saccharomyces cerevisiae*.

(a, b) The deletion strains *rer2Δsrt1Δ* and *nus1Δrer2Δsrt1Δ* were transformed either with a TkCPT subunit (TkCPT3/TkCPT4) alone or with a TkCPT and a TkCPTL subunit (TkCPTL1/TkCPTL2). Yeast cells were dropped on YPD, a complete synthetic medium containing 0.1% 5-fluoroorotic acid (5-FOA) or a synthetic defined medium lacking uracil (SD-Ura). The wild-type strain INVSc1 was used as a control.
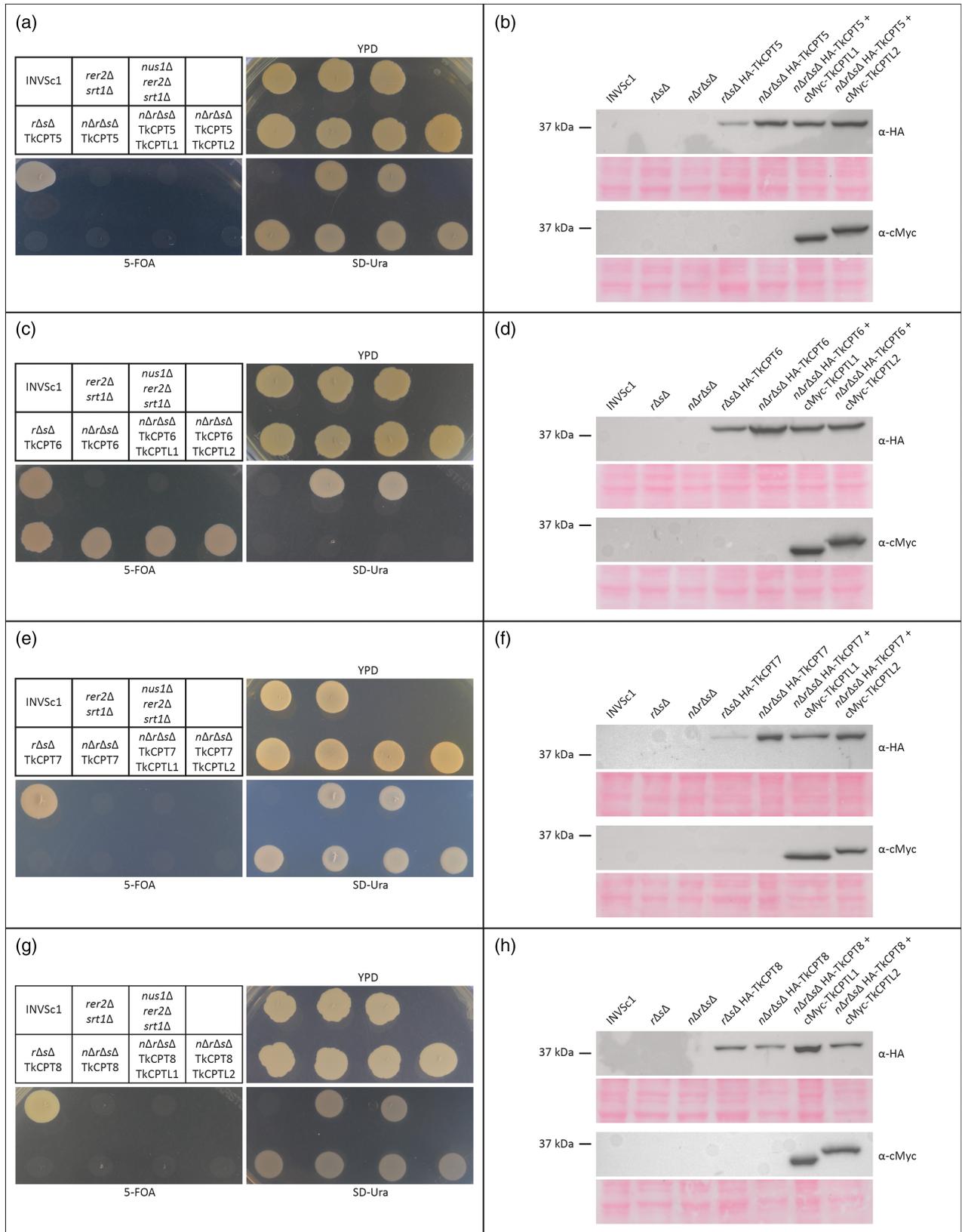
(c, d) Western blots of protein extracts from yeast strains. N-terminal HA-tagged TkCPTs were detected using an HA-specific antibody and N-terminal triple cMyc-tagged TkCPTLs were detected using a cMyc-specific antibody. A mouse IgG-specific secondary antibody was coupled to horseradish peroxidase (or alkaline phosphatase for the detection of HA-TkCPT4). The anticipated protein sizes were 35 kDa for HA-TkCPT3 and HA-TkCPT4, 33.8 kDa for 3xcMyc-TkCPTL1, and 34.1 kDa for 3xcMyc-TkCPTL2. Protein transfer to the nitrocellulose membranes was confirmed by Ponceau S staining.

isoprene-derived polymers with TkCPT2/TkCPTL1. To examine whether a low abundance of TkCPT2/TkCPTL1 resulted in the biosynthesis of polyisoprenes, codon-harmonized versions of these genes (TkCPT2*/TkCPTL1*, Table S2) were synthesized and expressed in yeast. Expression of the codon-harmonized genes led to a stronger signal compared to the original sequence in Western blot analysis using a polyclonal anti-CPT antibody (Figure S2f; Epping et al., 2015).

The ^{1}H-NMR spectra of crude *n*-hexane extracts showed strong signals for glycerolipids and other apolar compounds (Figure S3). Despite the overlapping peaks, minor signals matching the chemical shift values of a poly (*cis*-1,4-isoprene) reference were observed, in particular at 5.17 and 1.72 ppm (Figure S3), supporting the presence of poly(*cis*-1,4-isoprene) or closely related compounds in the crude extracts. Based on the integral values (relative to the solvent signal), the poly(*cis*-1,4-isoprene) content as a proportion of cell dry weight (CDW) in yeast expressing

the TkCPTL1/TkCPT1 complex (75 µg g^{-1} CDW) exceeded that of the control strain (33 µg g^{-1} CDW) and was almost double that of cells grown at 30°C (rox1::PGAL1-tHMGR PGAL10-ERG13 Perg9Δ::PCTR3 PGPD-TkCPT1 PTEF1-TkCPTL1 nus1Δ; 41 µg g^{-1} CDW), confirming the enhanced activity of the TkCPTL1/TkCPT1 complex during low-temperature yeast cultivation. However, the precise length of the polyisoprenes could not be determined due to the low signal-to-noise ratio of the specific polyisoprene signals.

For the expression of TkCPT3/TkCPTL1 and TkCPT4/TkCPTL1, we used LC–MS to monitor the accumulation of polyisoprenes and dolichols with chain lengths of C_{55}–C_{85} (Figure 9a). The LC–MS method was optimized using a commercial dolichol mixture (13–21) (Figure S4). TkCPT3/TkCPTL1 synthesized a dominant C_{70} product consisting of 14 isoprene units, whereas TkCPT4/TkCPTL1 produced C_{80} polyisoprenes with 16 isoprene units. For comparison, we analyzed polyisoprenes and dolichols in

Figure 7. Functional complementation of putative homomeric TkCPTs in *Saccharomyces cerevisiae*.

(a–c, e, g) The deletion strains *rer2Δsrt1Δ* and *nus1Δrer2Δsrt1Δ* were transformed either with a TkCPT (TkCPT5, TkCPT6, TkCPT7, or TkCPT8) alone or with a combination of a TkCPT and a TkCPTL subunit (TkCPTL1 or TkCPTL2). Yeast cells were dropped on YPD, a complete synthetic medium containing 0.1% 5-fluoroorotic acid (5-FOA) or a synthetic defined medium lacking uracil (SD-Ura). The wild-type strain INVSc1 was used as a control.

(b, d, f, h) Western blots of protein extracts from yeast strains. N-terminal HA-tagged TkCPTs were detected using an HA-specific antibody and N-terminal triple cMyc-tagged TkCPTLs were detected using a cMyc-specific antibody. A mouse IgG-specific secondary antibody was coupled to horseradish peroxidase. The anticipated protein sizes were 37.1 kDa for HA-TkCPT5, 36.6 kDa for HA-TkCPT6, 39 kDa for HA-TkCPT7, 38.3 kDa for HA-TkCPT8, 33.8 kDa for 3xcMyc-TkCPTL1, and 34.1 kDa for 3xcMyc-TkCPTL2. Protein transfer to the nitrocellulose membranes was confirmed by Ponceau S staining.
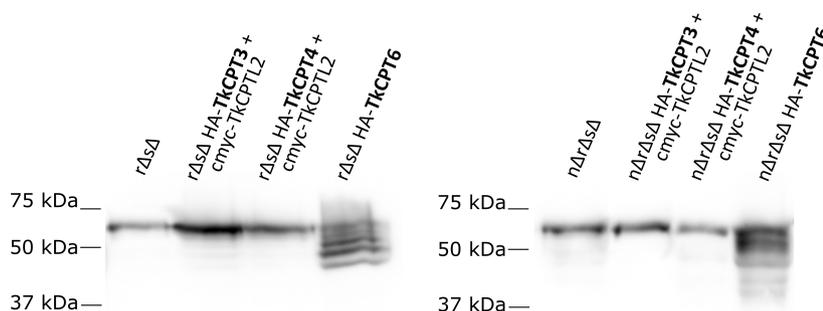
Figure 8. Western blot analysis of carboxypeptidase Y (CPY) encoded by the *PRC1* gene.

The CPY protein has an apparent molecular weight of approximately 61 kDa in the original strains *rer2Δsrt1Δ* and *nus1Δrer2Δsrt1Δ* when complemented by the *G. lambdia cis*PT or TkCPT3/TkCPTL2 or TkCPT4/TkCPTL2, indicating complete glycosylation. However, whereas TkCPT6 complements the growth deficiency of the yeast strains, allowing them to grow on 5-FOA, the detection of additional smaller isoforms of PRC1 reveals incomplete glycosylation of the protein.

Table 2 Molecular weights of products determined by thermal field flow fractionation of extracts from yeast cells expressing TkCPT1/TkCPTL1 or TkCPT2/TkCPTL1. M_n is the number average molecular weight, M_w is the weight average molecular weight and M_z is the Z average molecular weight, which has a greater weighting with regards to higher weights. For both combinations, we analyzed $n = 3$ independent clones

| Expression construct in yeast	M_n	M_w	M_z	Yield (mg g^{-1} yeast)
TkCPT1/TkCPTL1	3.3×10^6	4.0×10^6	5.6×10^6	0.43
TkCPT2/TkCPTL1	$>10^4$	$>10^4$	$>10^4$	0.05
TkCPT2*/TkCPTL1*	1.9×10^5	2.7×10^5	6.6×10^5	0.18

*Codon-harmonized and improved cultivation conditions.

different *T. koksaghyz* tissues (Figure 9b). Latex, root, and leaf tissues produced dominant C_{75} polyisoprenes/dolichols (15 isoprene units) whereas peduncle and flower tissues produced similar amounts of C_{75} and C_{80} polyisoprenes/dolichols (15 and 16 isoprene units, respectively).

For the expression of *TkCPT5–8* in yeast, we cloned truncated expression cassettes lacking transit peptides. *TkCPT5* produced an additional GC–MS signal that we also observed in *T. koksaghyz* flower tissue (Figure 10a,b). The mass spectrum of this compound was similar to that of geranylgeraniol and hexaprenol, kindly provided by Kruk and Szymańska (2023; Figure 10c–f). However, it eluted at 295°C (RT = 15.32 min), which is between the elution temperatures of geranylgeraniol (C_{20}, 258°C, RT = 13.07 min)

and hexaprenol (C_{30}, 330°C, RT = 17.82 min; Figure S5a). The largest recorded ion of the geranylgeraniol fragmentation was $m/z = 290$ and that of the unknown compound was $m/z = 358$, suggesting pentaprenol (C_{25}, $m/z = 358.4$, representing the monoisotopic mass). For confirmation, we also analyzed the extracts by LC–MS and extracted the corresponding ions from the MS data. The ions that could be assigned to pentaprenol were $m/z = 376.4$ ($M + NH_4^+$), 341.4 ($M - OH^-$), and 381.3 ($M + Na^+$). These were detected in the extracts of yeast expressing *TkCPT5* and in *T. koksaghyz* flowers (Figure 10g). Geranylgeraniol and pentaprenol standards were compared to the TkCPT5 product by LC–MS (Figure S5b). The geranylgeraniol eluted first at 10.8 min with a predominant m/z of 291.3 (corresponding to the protonated monoisotopic ion), whereas pentaprenol and hexaprenol were detected as ammonium ions ($m/z = 376.4$, 444.4) at 15 and 16 min, respectively.

The expression of TkCPT6 in yeast followed by LC–MS produced several additional signals in the UV chromatogram that could not be attributed to specific *T. koksaghyz* tissues (Figure 11a,b). When we screened for the specific NH_4^+ ions in the transformed yeast we detected hexaprenol ($m/z = 444.4$), heptaprenol ($m/z = 512.4$), octaprenol ($m/z = 580.5$), nonaprenol ($m/z = 648.6$), and decaprenol ($m/z = 716.6$) (Figure 11c). Each polyisoprene ion and corresponding dihydro-polyisoprene eluted at a similar time, but the latter showed an additional mass of 2 due to the reduction of a double bond to form dihydro-polyisoprene ($m/z = 446.4$, 514.4, 582.5, 650.6, and 718.6). In the control
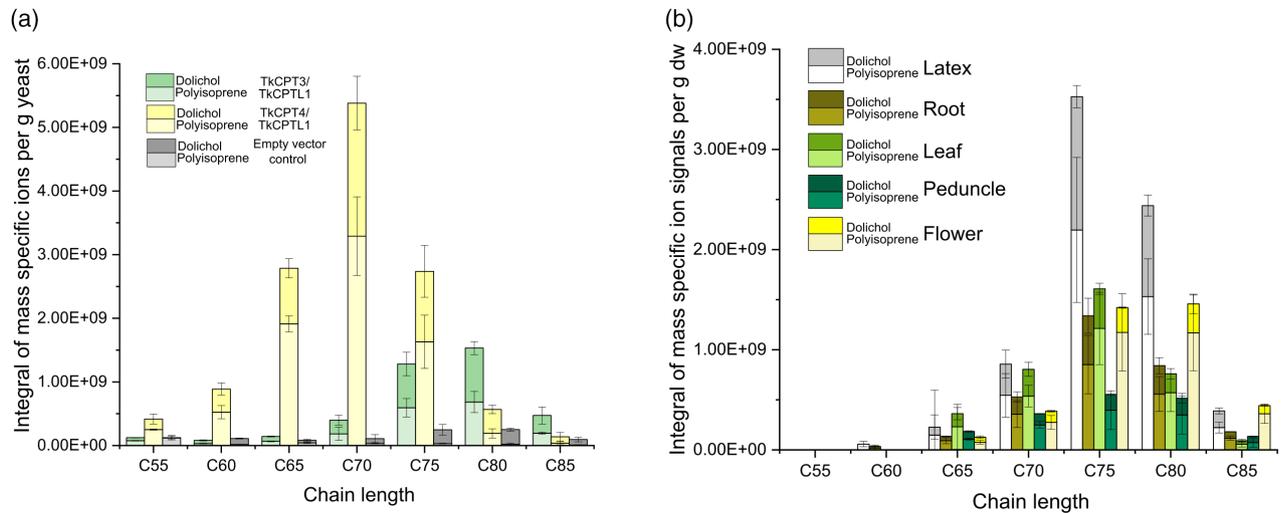
Figure 9. Polyisoprenes/dolichols in yeast cells expressing TkCPT3/TkCPTL1 and TkCPT4/TkCPTL1 and in different *Taraxacum koksaghyz* tissues.
(a) Dolichols and polyisoprenes were extracted from yeast expressing TkCPT3/TkCPTL1 and TkCPT4/TkCPTL1 or the empty vector control and were separated on a C18 column for LC–MS. The specific *m/z* signals at particular elution times were integrated, revealing a mixture of polyisoprenes/dolichols with C_{70} products most abundant for TkCPT3/TkCPTL1 and C_{80} products most abundant for TkCPT4/TkCPTL1.
(b) In *T. koksaghyz*, C_{75} products were predominant in the latex, root, leaf, and peduncle tissues, whereas C_{80} products were the most abundant in the flowers (dw, dry weight). For both combinations, we analyzed $n = 3$ independent expression clones. For the analysis of plant material, we pooled material from $n = 3$ plants for extraction.

yeast, the signal of $m/z = 444.4$ probably corresponds to (*S*)-2,3-oxidosqualene, which is typically present in yeast as a sterol precursor (Figure 11d). In *T. koksaghyz*, we detected only nonaprenol ($m/z = 648.6$) in leaves with the correct retention time (Figure 11e). However, an ion with a mass identical to hexaprenol ($m/z = 444.4$) was detected in latex and root tissue but eluted at 16.6 min, later than the hexaprenol compound in yeast. This may reflect the presence of triterpenoids such as taraxasterol in the roots and latex due to oxidosqualene cyclase activity.

The expression of TkCPT7 or TkCPT8 in yeast produced nerol that could be detected by GC–MS (Figure 12a, b). No saponification was necessary, so we assume that endogenous yeast phosphatase activity (Ignea et al., 2019) produced nerol from NPP. However, nerol was not detected in *T. koksaghyz* flower, stem, leaf, or root tissue, suggesting that nerol (or more likely the corresponding NPP) is not the end product, and is probably converted into monoterpenes as a precursor for further IPP polymerization or for the production of other metabolites *in planta*.

Table 3 summarizes the properties of the eight TkCPT enzymes. TkCPTs 1–4 form heteromers with TkCPTL proteins and are associated with the ER. The expression of TkCPT1–4, each in combination with TkCPTL1, rescues a yeast mutant deficient for dolichol synthesis. TkCPT1 and TkCPT2 produce very long-chain polyisoprenes in yeast, whereas TkCPT3 and TkCPT4 produce long-chain dolichol/-polyisoprenes. TkCPT1 and TkCPT2 are strongly expressed in the latex and their primary function is likely to be the

synthesis of NR. In contrast, TkCPT3 (mainly expressed in latex) and TkCPT4 (mainly expressed in leaves) are more likely to form dolichol/polyisoprenes for *N*-glycosylation. TkCPTs 5–8 do not interact with TkCPTL proteins and TkCPT5–7 fusions with a C-terminal fluorophore were translocated into the chloroplast stroma. Among the homomeric enzymes, only TkCPT6, which catalyzed the formation of the largest polyisoprene chain ($C_{30–50}$), was able to rescue the yeast *nus1Δrer2Δsrt1Δ* mutant. TkCPT5 is predominantly expressed in *T. koksaghyz* flowers, which contain pentaprenol, and the same compound was produced in yeast expressing TkCPT5. In contrast, TkCPT7–8 are mainly expressed in *T. koksaghyz* leaves and produce nerol when expressed in yeast, but this compound was not detected in *T. koksaghyz* fresh leaf tissue suggesting that NPP is converted to other compounds such as monoterpenoids *in planta*.

DISCUSSION

The activity of *cis*PTs in *T. koksaghyz* has been studied mainly in connection with NR biosynthesis. However, these homomeric and heteromeric enzymes can synthesize a variety of other linear isoprene compounds that are likely to be involved in diverse biological processes. This is supported by the presence of eight CPTs (TkCPT1–8) and two CPTL proteins (TkCPTL1–2) in *T. koksaghyz*. Similar complexity has been observed in Arabidopsis and tomato, where the functions of some *cis*PTs have already been determined. We therefore set out to characterize the full set of CPTs in *T. koksaghyz*.
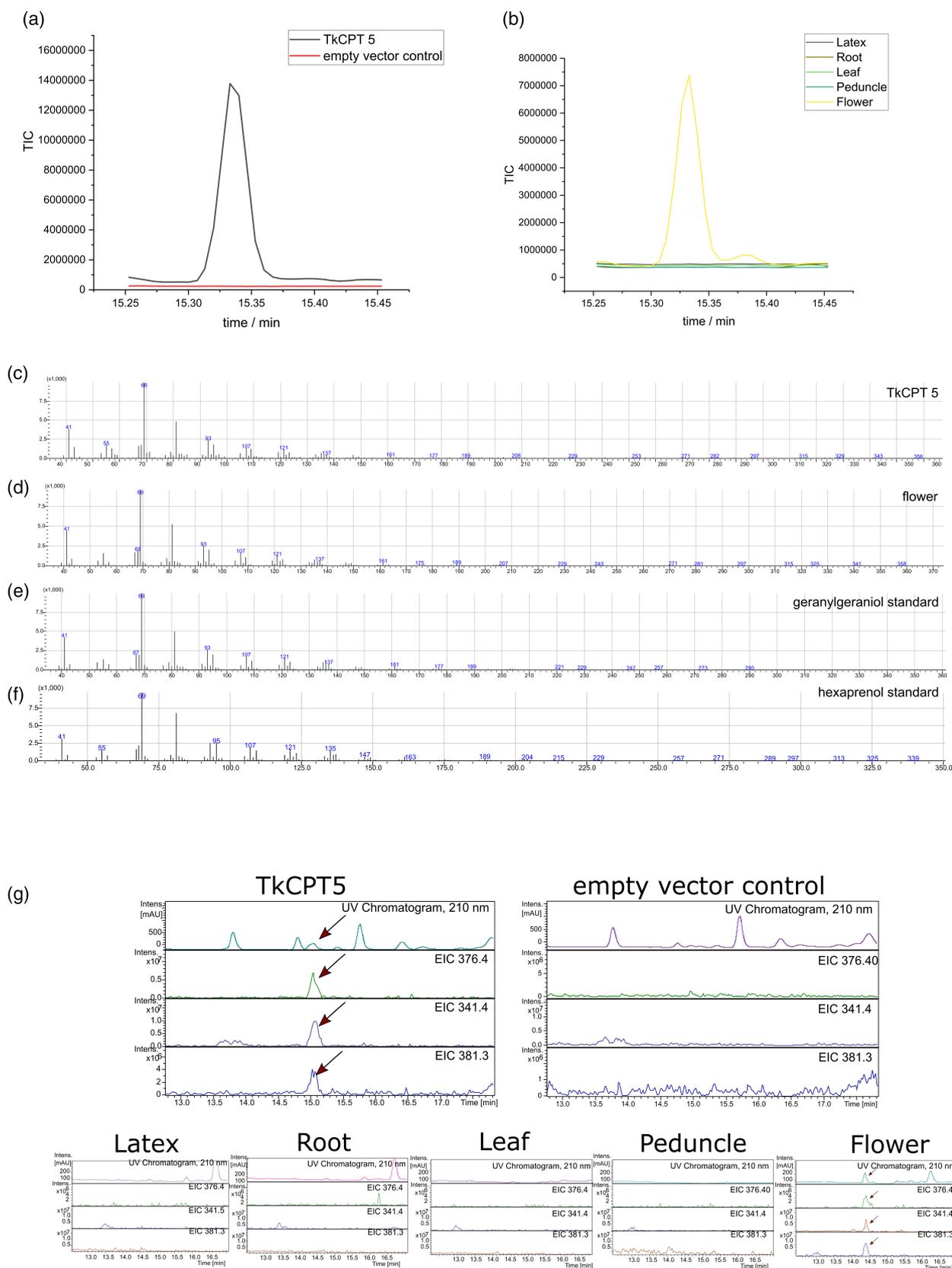
Figure 10. Pentaprenol was detected by GC–MS and LC–MS analysis in yeast expressing TkCPT5 and in *Taraxacum koksaghyz* flower tissue.

Extracts from (a) yeast expressing TkCPT5 or the vector control and (b) different *T. koksaghyz* tissues were analyzed by GC–MS, revealing an additional peak in (c) yeast expressing TkCPT5 and (d) *T. koksaghyz* flowers, with a similar mass spectrum to (e) the geranylgeraniol standard.

(f) LC–MS analysis detected in measurements with the alternate positive and negative mode the ions $m/z = 376.4$ $(M + NH_4^+)$, 341.4 $(M - OH^-)$, and 381.3 $(M + Na^+)$ that are specific for pentaprenol in the same extracts (EIC, extracted ion count; TIC, total ion count). We analyzed $n = 3$ independent expression clones.
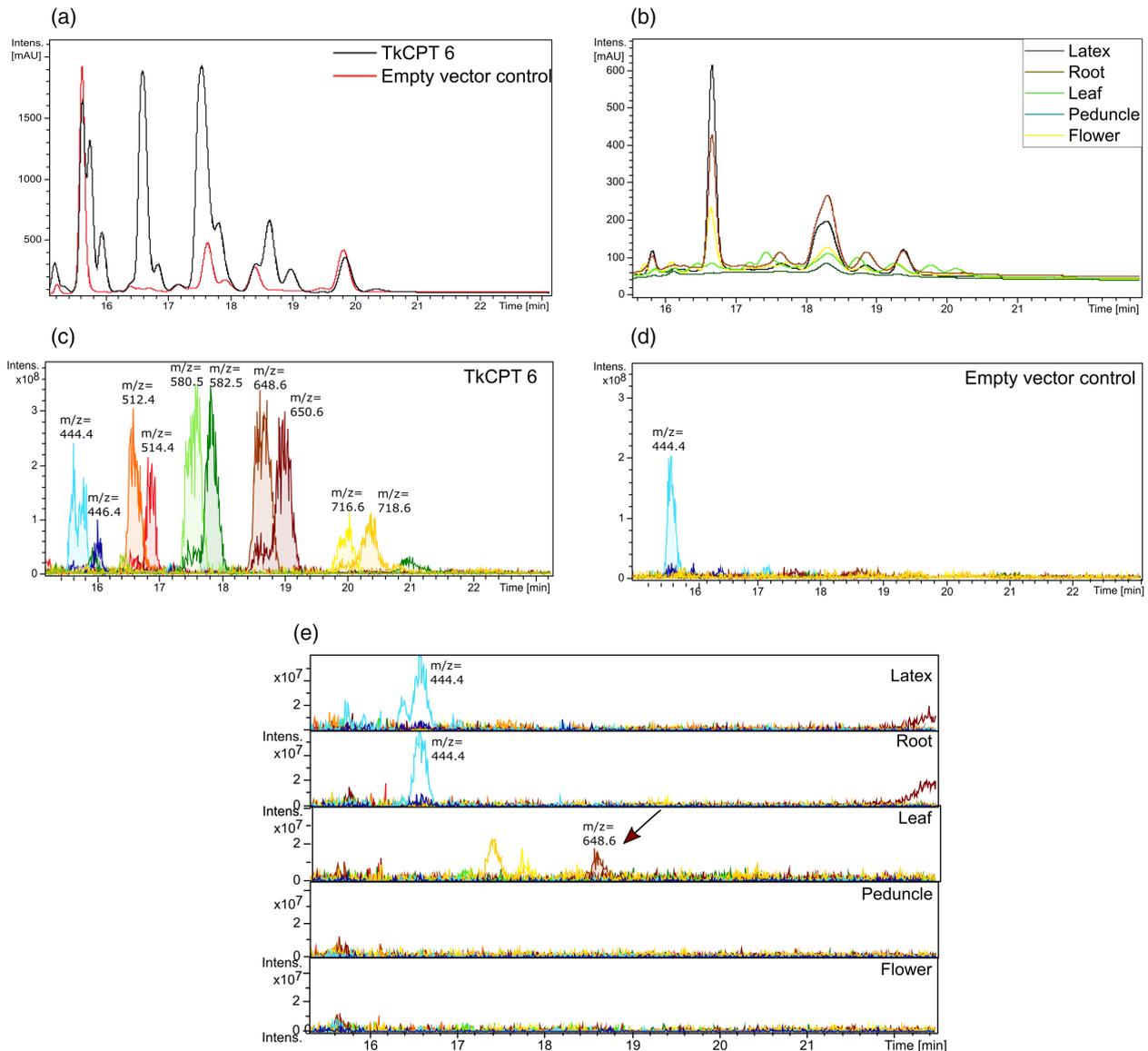
Figure 11. Polyisoprenes with a length of 6–10 isoprene units in yeast expressing TkCPT6 and nonaprenol in *Taraxacum koksaghyz* flower tissue detected by GC–MS and LC–MS analysis.

Extracts from (a, c, d) TkCPT6-expressing yeast or the vector control and (b, e) different *T. koksaghyz* tissues were analyzed by LC–MS, showing that yeast expressing TkCPT6 accumulate dolichols/polyisoprenes that (a) absorb in the UV range (UV chromatogram 210 nm) and (c) are detected as NH_4^+ ions ($m/z = 444.4$ for hexaprenol and 446 for dihydro-hexaprenol, 512.4 for heptaprenol and 514.4 for dihydro-heptaprenol, 580.5 for octaprenol and 582.5 for dihydro-octaprenol, 648.6 for nonaprenol and 650.6 for dihydro-nonaprenol, and 716.6 for decaprenol and 718.6 for dihydro-decaprenol).

(e) For the plant tissue, only nonaprenol ($m/z = 648.6$) was identified in the leaves. We analyzed $n = 3$ independent expression clones.

Phylogenetic analysis identified TkCPT3 and TkCPT4 as putative CPT subunits that form heterodimers with *cis*PT activity. These proteins were therefore grouped with TkCPT1 and TkCPT2, which form the heteromeric rubber transferase complex with TkCPTL1. In contrast, TkCPTs 5–8 were identified as putative homomeric *cis*PTs,

(a)

(b)
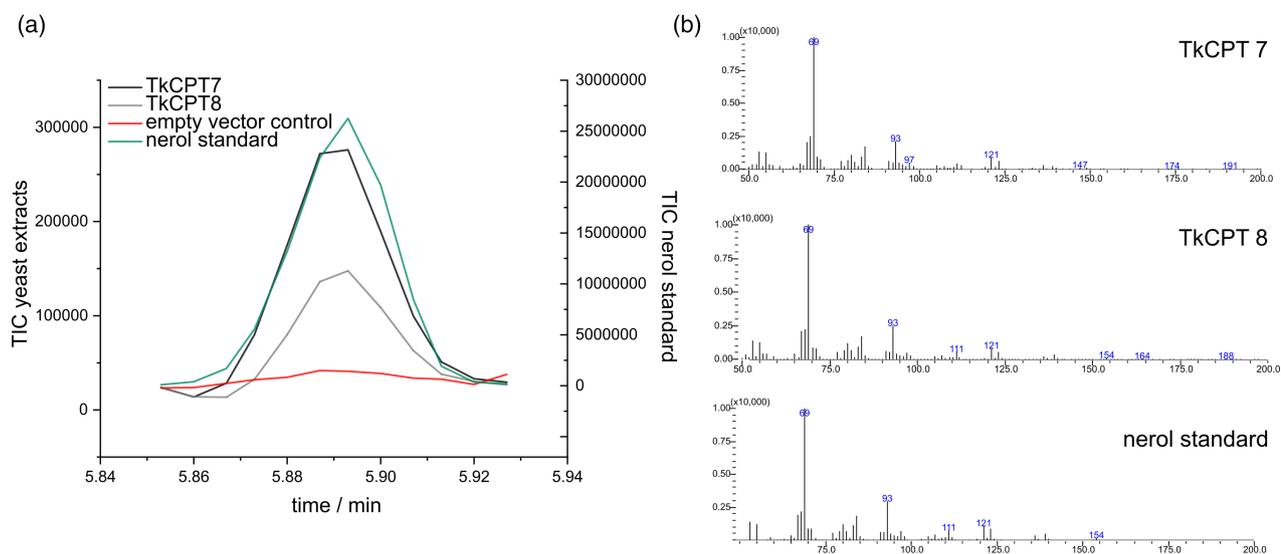
Figure 12. Nerol detected by GC–MS in yeast expressing TkCPT7 or TkCPT8.
(a) The GC profiles of the strains expressing TkCPT7 and TkCPT8 compared with the empty vector control.
(b) The corresponding mass spectra compared to the nerol standard. TIC, total ion count. We analyzed $n = 3$ independent expression clones for TkCPT7 and TkCPT8.

Table 3 Summary of the experimentally determined properties of TkCPT1–8

Name	Maximum expression	Subcellular localization when expressed in *Nicotiana benthamiana* leaf epidermis	Heteromeric complex with TkCPTL proteins	nΔrΔsΔ rescue	Product in yeast	Potential product in *Taraxacum koksaghyz*
TkCPT1	Latex	ER (and possibly cytosol)	Yes*	Yes*	Polyisoprene (approximately $C_{235\,000}$)	Natural rubber
TkCPT2	Latex	ER (and possibly cytosol)	Yes*	Yes*	Polyisoprene (approximately $C_{13\,600}$)	Natural rubber
TkCPT3	Leaf	ER (and possibly cytosol)	Yes	Yes	Dolichol/polyisoprene ($C_{55–85}$, mainly C_{70})	Dolichol/ polyisoprene
TkCPT4	Latex	ER (and possibly cytosol)	Yes	Yes	Dolichol/polyisoprene ($C_{55–85}$, mainly C_{80})	Dolichol/ polyisoprene
TkCPT5	Flower	Stroma	No	No	Pentaprenol (C_{25})	Pentaprenol
TkCPT6	Leaf	Stroma	No	Yes, hypo-glyco-sylation	Hexaprenol –decaprenol ($C_{30–50}$)	Nonaprenol
TkCPT7	Leaf	Stroma	No	No	Nerol (C_{10}, NPP)	Monoterpenes or other compounds
TkCPT8	Leaf	n.d.	No	No	Nerol (C_{10}, NPP)	Monoterpenes or other compounds

*Data from Epping et al. (2015) and Niephaus et al. (2019).

although the typical R*X*G motif in the C-terminal region is missing from TkCPT8 (Figure 1). The R*X*G motif is conserved in homomeric *cis*PTs and CPTL subunits. Mutation of the arginine and glycine residues in EcUPPS, GlUPPS, and HsNgBR reduced *cis*PT activity, suggesting that the R*X*G motif is necessary for polyisoprenyl diphosphate synthesis (Grabińska et al., 2017). The crystal structure of the human heteromeric *cis*PT complex revealed that these three conserved amino acid residues are involved in the binding of the allylic substrate (Edani et al., 2020).

As expected for the components of the rubber transferase complex, *TkCPT1* and *TkCPT2* were predominantly expressed in the latex, like *TbCPT2* and *TbCPT3* in *T. brevicorniculatum* (Schmidt et al., 2010). *TkCPT4* was uniformly expressed at low levels in all tissues, whereas *TkCPT3* and *TkCPT6–8* were strongly expressed in the leaves, indicating a tissue-specific function (Figure 2). We found that *TkCPT5* was mainly expressed in flowers, agreeing with an earlier study in which *TkCPT5* was mainly expressed in flowers and seeds (Lin et al., 2017). Our spatial expression profiles

also matched the transcriptome data for most of the other *TkCPT* genes (Lin et al., 2017). The exception was *TkCPT7*, where we observed the strongest expression in leaves but the transcriptome data indicated high *TkCPT7* expression in roots and latex. One potential explanation is the different growth stages tested in each study (our plants were cultivated for 3 months before tissues were harvested, whereas the transcriptome analysis used samples harvested after 1.5, 3, or 5 months).

The analysis of protein localization revealed that TkCPT1–4 are associated with the ER (Figure 3). These proteins did not contain an N-terminal signal peptide, but a C-terminal di-lysine motif may direct them to the ER membrane (Benghezal et al., 2000). These lysine residues usually occupy positions -3, -4 or -3, -5 relative to the C-terminus, but in TkCPT4 they were found further upstream. It would be interesting to investigate the subcellular localization of the entire heteromeric *cis*PTs given that the cytosolic HbHRT2 protein was recruited to different organelles when co-expressed with various other proteins of the NR biosynthesis machinery (Brown et al., 2017). Additionally, LsCPT3 relocates from the cytosol to the ER when co-expressed with LsCPTL2 in *N. benthamiana* leaves or *S. cerevisiae* cells (Qu et al., 2015).

The putative homomeric *cis*PTs (TkCPT5–7) were associated with the chloroplast stroma (Figure 4). Homomeric plant *cis*PTs often reside in chloroplasts but are not restricted to this organelle (Akhtar et al., 2013; Grabińska et al., 2016; Kera et al., 2012; Surmacz et al., 2014). Arabidopsis CPT4 (At5g58770) and tomato CPT5 are homomeric *cis*PTs located in the stroma. They synthesize medium-chain polyisoprenes (approximately C_{55}) that accumulate in thylakoid membranes and the plastidial envelope, influencing photosynthetic performance by controlling membrane dynamics (Akhtar et al., 2017; van Gelder et al., 2018). TkCPT5 and TkCPT6 may have a similar effect in *T. koksaghyz* by promoting the synthesis of pentaprenol (TkCPT5) and hexaprenol to nonaprenol (TkCPT6), as demonstrated in the transformed yeast strains. AtCPT3 (At2g23410) and AtCPT5 (At5g58780) are ER-resident homomeric *cis*PTs (Kera et al., 2012; Surmacz et al., 2014; Surowiecki et al., 2019), whereas AtCPT1 (At2g17570, also named AtCPT3 in other publications) is predicted to reside in mitochondria, although this has not been confirmed experimentally (Kopcsayová & Vranová, 2019). However, AtCPT1 may be associated with the ER because it interacts with the CPTL protein AtLEW1 to form a heteromeric *cis*PT complex that synthesizes dolichols, which may be important for plant stress responses (Kwon et al., 2016; Zhang et al., 2008). Recently, QTL mapping and molecular analysis identified this enzyme as a major producer of Dol-15 to Dol-17 in Arabidopsis (Gawarecka et al., 2022). Moreover, SlCPT4, SlCPT6, and SlCPT7 do not co-localize with a mitochondrial protein bearing a fluorescent protein tag (Akhtar et al., 2013).

Phylogenetic analysis suggested that TkCPT3 and TkCPT4 are CPT subunits and we confirmed their interaction with TkCPTL1 and TkCPTL2 in BiFC experiments (Figure 5). TkCPT1 and TkCPT2 were previously shown to interact with TkCPTL1 and TkCPTL2 and their identical *T. brevicorniculatum* orthologs (Epping et al., 2015; Niephaus et al., 2019). Consistently, the putative homomeric *cis*PTs TkCPT5–8 did not interact with either TkCPTL1 or TkCPL2 (Figure S1). In any case, interaction would be prevented by compartmentalization because TkCPTs 5–7 are located in the chloroplasts whereas the TkCPTL proteins are associated with the ER. TbRTA, which is identical to TkCPTL1, was also previously shown to be resident in the ER (Epping et al., 2015).

The functional analysis of *T. koksaghyz* *cis*PTs in a yeast complementation assay confirmed that TkCPT3 and TkCPT4 depend on TkCPTL proteins for enzymatic activity as part of a heteromeric complex (Figure 6). However, the CPT–CPTL interaction appears to be species-dependent, because neither TkCPT3 nor TkCPT4 was able to complement the mutant yeast strain *rer2Δsrt1Δ*. This indicates they did not interact with Nus1, the yeast homolog of CPTL, as previously reported for the Arabidopsis enzyme AtCPT1 (Kwon et al., 2016). When TkCPT3 or TkCPT4 were co-expressed with TkCPTL1 in yeast, the cells produced dolichols with 11–17 isoprene units, similar to the products synthesized in *T. koksaghyz*. However, TkCPT4 tended to produce longer polyisoprene chains than TkCPT3—the most abundant ions containing 16 and 14 isoprene units, respectively. We confirmed that TkCPT6, which functions as a homomeric enzyme in *T. koksaghyz*, is sufficient for independent *cis*PT activity (Figure 7) and therefore clusters with AtCPT3 (Surowiecki et al., 2019). The expression of TkCPT5, TkCPT7, or TkCPT8 did not enable yeast growth. This may reflect the size of the resulting prenols but could also be caused by the presence of the signal peptide, which was shown to influence the functionality of *cis*PTs in Arabidopsis (Akhtar et al., 2017). Therefore, we truncated the *cis*PTs for overexpression in a modified yeast strain. Product analysis in yeast indicated that the synthesis of polyisoprenes with 6–10 isoprene units was always accompanied by the dihydro-form, which may be sufficient to complement the missing yeast-specific dolichols. In contrast, we only detected nonaprenol (also known as solanesol) *in planta*, and this may be further processed to plastoquinone-9 or ubiquinone-9. The closest Arabidopsis homolog (AtCPT7) has a similar product range, including nonaprenol, decaprenol, and undecaprenol. The enzyme is upregulated in response to heat stress, which involves heat shock transcription factors of the HSFA1 family. Its products modulate plastid ultrastructure, the aggregation of LHCII complexes, and non-photochemical quenching (Buszewicz et al., 2021). The putative homomeric *cis*PTs TkCPT5, TkCPT7, and TkCPT8 therefore synthesized

products that did not restore the growth of the yeast deletion strains, indicating that the chain length of the polyisoprenyl diphosphates is important for functional complementation. For example, AtCPT5 synthesizes mainly C_{35} polyisoprenes/dolichols, and only partially complements the yeast strain by restoring growth but not complete glycosylation (Surmacz et al., 2014). We analyzed the glycosylation of PRC1 in yeast expressing TkCPT6 (product size C_{30-50}) and detected lower-molecular-weight bands that indicated incomplete glycosylation, whereas yeast expressing TkCPT3 or TkCPT4 (product size C_{55-85}) completed the glycosylation of PRC1. Furthermore, yeast glycosyltransferases that catalyze the transfer of sugar units such as N-acetylglucosamine to the carrier lipid dolichyl monophosphate prefer not only a polyisoprenyl chain longer than seven isoprene units ($>C_{35}$) but also a saturated α-isoprene unit (Palamarczyk et al., 1980). TkCPT5 produced pentaprenol when expressed in yeast but lacked the specific dihydro-form. The biological function of this compound, which was also detected in flower tissue, remains unknown. It may influence membrane functions in plastids. Our hypothesis that the polyisoprenyl products of TkCPT5, TkCPT7, and TkCPT8 were too short (or not sufficiently reduced to dolichols) to achieve full protein glycosylation was verified by expression in yeast because the product of TkCPT5 was pentaprenol and the product of TkCPT7 and TkCPT8 was nerol. The lower activity of TkCPT8 may reflect its missing R*X*G motif. The identification of short-chain *cis*PT products might also be of interest for NR research because it is unclear which short allylic diphosphates serve as starter molecules for the elongation reaction catalyzed by the rubber transferase complex in *T. koksaghyz*. Short *cis*-allylic compounds have received little attention as potential priming substrates.

This is the first comprehensive analysis of the localization, interactions, and products of the entire *T. koksaghyz* cis-PT family *in vivo*, not only providing insight into the regulation of isoprenoid metabolism in NR-producing species but also revealing evidence of at least one previously unknown pentaprenol found uniquely in flowers. In terms of broader impact, our work will enable the selection of targets for future experiments focusing on primary and secondary polyisoprenoid metabolites, including those of interest in the context of dandelion breeding as an NR crop. Our work indicates that the *T. koksaghyz* rubber transferase complex is a specialized *cis*PT, which should also facilitate the rational design of improved enzymes to enhance NR production.

EXPERIMENTAL PROCEDURES

Plant material and cultivation

Twelve wild-type *T. koksaghyz* plants were grown in a greenhouse for 12 weeks at 14–18°C (night) and 22–25°C (day) with a 16-h photoperiod and a daytime light intensity of 20 klx provided by high-pressure 600-W sodium lamps with an enhanced yellow and red spectrum (Greenbud, Wischhafen, Germany). After cultivation for 4–5 weeks in VM propagation substrate (Einheitserde, Sinntal-Altengronau, Germany), the plants were transferred to ED73 standard soil (Einheitserde). *N. benthamiana* plants for transient expression were cultivated in a growth chamber at 20°C under continuous light for 3–5 days.

Amplification of *TkCPT* genes

The *TkCPT3–8* full-length coding sequences were amplified using primers TkCPT3-NcoI-fwd/TkCPT3-SalI-rev, TkCPT4-NcoI-fwd/TkCPT4-SalI-rev, TkCPT5-NcoI-fwd/TkCPT5-XhoI-rev, TkCPT6-NcoI-fwd/TkCPT6-SalI-rev, TkCPT7-NcoI-fwd/TkCPT7-SalI-rev, and TkCPT8-NcoI-fwd/TkCPT8-XhoI-rev. *TkCPT3* and *TkCPT6* were amplified from leaf cDNA, *TkCPT4* and *TkCPT7* from latex cDNA, *TkCPT5* from flower cDNA, and *TkCPT8* from peduncle cDNA according to their expression profiles (Figure 2). Overlap-PCR was used to amplify *TkCPT5* without introns. *TkCPT5*-exon 1, *TkCPT5*-exon 2, and *TkCPT5*-exon 3 were amplified separately from pENTR4-TkCPT5 (carrying the full-length *TkCPT5* gene, including introns) using primers TkCPT5-NcoI-fwd/TkCPT5-Ex1OL-rev, TkCPT5-Ex2OL-fwd/TkCPT5-Ex2OL-rev, and TkCPT5-Ex3OL-fwd/TkCPT5-XhoI-rev. The three fragments with overhangs were used as the template for overlap-PCR with primers TkCPT5-NcoI-fwd and TkCPT5-XhoI-rev. All primers (Table S3) were based on the *T. koksaghyz* draft genome (Lin et al., 2017) except those used to amplify *TkCPT7*, which were based on RNA-Seq data from latex, root and leaf mRNA.

In silico analysis of TkCPT proteins

Transmembrane domains were predicted using the TMHMM Server v2.0 (http://www.cbs.dtu.dk/services/TMHMM/). Protein subcellular localization was predicted using TargetP-2.0 (Almagro Armenteros et al., 2019). Amino acid alignments were built using MUSCLE (https://www.ebi.ac.uk/Tools/msa/muscle/) and edited for visualization with Jalview v2.11.0 (Waterhouse et al., 2009). The neighbor-joining tree was constructed in MEGA6 with bootstrap values from 1000 replicates (Tamura et al., 2013). The final tree was visualized with FigTree (http://tree.bio.ed.ac.uk/software/).

Total RNA isolation and cDNA synthesis

Latex was harvested from *T. koksaghyz* roots in 90 µL rubber extraction buffer with 5 mM DTT as previously described (Schmidt et al., 2010). Root, leaf, peduncle, and flower tissue were lyophilized and ground to powder. Total RNA was then extracted using the innuPREP Plant RNA Kit (Analytik Jena, Jena, Germany) and cDNA was synthesized using PrimeScript RT Master Mix (Takara Bio Europe, Saint-Germain-en-Laye, France).

Gene expression analysis

TkCPT gene expression was analyzed by qRT-PCR in three pools, each comprising cDNA from four 12-week-old wild-type *T. koksaghyz* plants, as previously described (Laibach et al., 2015) with modifications. Gene expression was quantified using CFX Manager v3.1 (Bio-Rad Laboratories, Hercules, CA, USA). Transcript abundance was normalized to elongation factor-1α (*TkEF1α*) and ribosomal protein L27 (*TkRP*), which are suitable as reference genes (Pütter et al., 2017). Primer sequences are listed in Table S3. Primer efficiencies (66°C annealing temperature) and the use of formamide to increase specificity are described in Table S4. Primer efficiencies were determined using a cDNA dilution series representing all dandelion tissues.

Cloning strategies

The pENTR4 plasmids were constructed using the *TkCPT* sequences amplified as described above. PCR products without stop codons were prepared using corresponding reverse primers (Table S3). The gene fragments were digested with NcoI/XhoI or NcoI/SalI and were inserted into pENTR4 (Thermo Fisher Scientific, Waltham, MA, USA) at the NcoI/XhoI sites. The final plasmids were named pENTR4-TkCPT3–8 and all genes were also prepared without stop codons (pENTR4-TkCPT3–8 without stop codons).

HA-tagged TkCPT proteins were generated by digesting pENTR-HAL-MtSEO-F1 (Visser et al., 2016) with NcoI/XhoI, and the *MtSEO-F1* gene was replaced with *TkCPT* genes (NcoI/XhoI or NcoI/SalI) to yield the final vectors pENTR-HA-TkCPT3–8. *TkCPT1* was also amplified from pENTR4-TbCPT2 (Epping et al., 2015) using primers TkCPT1-PciI-fwd and TkCPT1-SalI-rev. After digestion with PciI/SalI, the product was inserted into pENTR-HA at the NcoI/XhoI sites, yielding vector pENTR-HA-TkCPT1.

The plastid stromal marker was created by amplifying the ChlTP sequence of the tobacco RuBisCO gene from leaf cDNA using primers NtRuBisCO-ChlTP-NcoI-fwd and NtRuBisCO-ChlTP-XhoI-rev (the sequence is provided in Supplementary Data S1). The PCR product was digested with NcoI and XhoI and inserted at corresponding sites in pENTR4 to yield pENTR4-NtRuBisCO-ChlTP. The *NtRuBisCO-ChlTP* sequence was also introduced into the Gateway destination vector pBatTL-ccdB-Venus (Ernst et al., 2012) using pENTR4-NtRuBisCO-ChlTP as a donor plasmid, resulting in the vector pBatTL-NtRuBisCO-ChlTP-Venus.

For co-localization studies, the eight *TkCPT* genes were inserted into the destination vector pBatTL-ccdB-Cerulean (kindly provided by David Wiedmann, Institute of Plant Biology and Biotechnology, University of Münster) using LR clonase (Thermo Fisher Scientific). The pENTR4-TkCPT vectors without stop codons were used as entry vectors. Because TkCPT1 and TkCPT2 are identical to TbCPT2 and TbCPT3, respectively, we used the existing vectors pENTR4-TbCPT2 and pENTR4-TbCPT3 without stop codons for the Gateway reaction (Epping et al., 2015). The final plasmids were named pBatTL-TkCPT1–8-Cerulean.

For the BiFC experiments, N-terminal and C-terminal split-mRFP fusions with *TkCPT3–8* were created by recombination in the Gateway destination vectors pBatTL-NmRFP-ccdB, pBatTL-CmRFP-ccdB, pBatTL-ccdB-NmRFP, and pBatTL-ccdB-CmRFP (Jach et al., 2006). The pENTR4-TkCPT vectors with or without stop codons served as donor plasmids, yielding pBatTL-NmRFP-TkCPT3–8, pBatTL-CmRFP-TkCPT3–8, pBatTL-TkCPT3–8-NmRFP, and pBatTL-TkCPT3–8-CmRFP.

For the yeast expression experiments, the *TkCPT3–8* sequences were introduced into the Gateway-compatible yeast expression vector pKG-GW2-ccdB (Park et al., 2014) with pENTR4-HA-TkCPT3–8 as donor plasmids. The final vectors were named pKG-GW2-HA-TkCPT3–8.

For the product analysis experiments in yeast, vectors pAG425_PGPD_TkCPT1–4_PTEF1_TkCPTL1 were constructed by amplifying the TEF promoter from *S. cerevisiae* genomic DNA using forward primer 5′-AAA AAG CTT GAG CTC CCC ACA CAC CAT AGC TTC-3′ and reverse primer 5′-AAA AAG CTT CTC GAG AAA CCA TGG TTT GTA ATT AAA ACT TAG ATT AG-3′. The PCR product was transferred to the HindIII sites of pACT2 (Takara Bio Europe) to generate pACT_PTEF. TkCPTL1 was amplified from *T. koksaghyz* cDNA using forward primer 5′-AAA CCA TGG CCG ATC TGTTAG ACG GACC-3′ and reverse primer 5′-AAA CTC GAG TTA TGA ACC GTA GTT CTGC-3′ and was inserted into the NcoI/XhoI sites to generate pACT2_PTEF_TkCPTL1. The TkCPTL1 expression

cassette was amplified using forward primer 5′-AAA AAG CTT GAG CTC CCC ACA CAC CAT AGC TTC-3′ and reverse primer 5′-AAA GAG CTC CAT GCC GGT AGA GGT GTG-3′ annealing to the TEF promoter and ADH terminator of pACT2_PTEF_TkCPTL1, respectively. This cassette was inserted at the SacI restriction site of pAG425_PGPD-ccdB (Alberti et al., 2007) to obtain pAG425_PGPD-ccdB-PTEF-TkCPTL1. The TkCPT1–4 genes were then introduced into pAG425_PGPD-ccdB-PTEF-TkCPTL1 by LR recombination with pENTR-TkCPT1–4. The pENTR-TkCPT5–8 vectors lacking the corresponding signal peptides were constructed by amplification with the primer pairs TkCPT5-oTP-SalI-fwd/TkCPT5-XhoI-rev, TkCPT6-oTP-SalI-fwd/TkCPT6-SalI-rev, TkCPT7-oTP-SalI-fwd/TkCPT7-SalI-rev, and TkCPT8-oTP-SalI-fwd/TkCPT8-XhoI-rev followed by insertion at the SalI sites of vector pENTR-3C and LR recombination into pAG425_PGPD-ccdB (Alberti et al., 2007). The integrity of all constructs was verified by sequencing (Sanger et al., 1977) on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Restriction enzymes were obtained from New England Biolabs (Ipswich, MA, USA).

Co-localization analysis in *N. benthamiana*

To investigate the subcellular localization of the TkCPT proteins, *Agrobacterium tumefaciens* strains carrying pBatTL-TkCPT1–8-Cerulean constructs were co-infiltrated into *N. benthamiana* leaves with strains harboring either pBatTL-NtermAtCYP51G1-mRFP (Epping et al., 2015) or pBatTL-NtRuBisCO-ChlTP-Venus as previously described (Müller et al., 2010). NtermAtCYP51G1-mRFP is an ER marker, whereas NtRuBisCO-ChlTP-Venus is a chloroplast stromal marker. Fluorescent signals were detected by confocal laser scanning microscopy (CLSM) using a TCS SP5 confocal microscope (Leica Microsystems, Wetzlar, Germany). For the combination of Cerulean-tagged and mRFP-tagged proteins, the fluorophores were excited at 458 nm (argon laser) and 543 nm (white light laser), respectively. The emission wavelengths were 469–510 and 582–650 nm, respectively. The proteins tagged with Cerulean and Venus were scanned sequentially along with chlorophyll autofluorescence to avoid interference caused by the overlapping excitation and emission spectra. Cerulean (excitation 405 nm with UV diode, emission 470–518 nm) and chlorophyll (excitation 458 nm with argon laser, emission 655–747 nm) signals were monitored simultaneously, whereas Venus (excitation 488 nm with argon laser, emission 556–615 nm) was measured separately.

BiFC analysis in *N. benthamiana*

BiFC analysis of TkCPT–TkCPTL interactions was achieved by transiently expressing N-terminal and C-terminal split-mRFP fusion proteins in *N. benthamiana* leaves by infiltration with *A. tumefaciens* (Müller et al., 2010). Combinations of *A. tumefaciens* strains carrying pBatTL-NmRFP-TkCPT constructs with pBatTL-TkCPTL1-CmRFP (Epping et al., 2015), pBatTL-TkCPTL2-CmRFP (Niephaus et al., 2019), or pBatTL-CmRFP were used for infiltrations, and mRFP fluorescence was visualized by CLSM with excitation and emission wavelengths of 534 nm (white light laser) and 564–626 nm, respectively.

Yeast complementation assay

The functionality of *T. koksaghyz cis*PT proteins was determined using a yeast complementation assay (Epping et al., 2015; Niephaus et al., 2019). The double-deletion strain KG 119 (*rer2Δsrt1Δ*) (Grabińska et al., 2010) and the triple-deletion strain KG 405

(*nus1Δrer2Δsrt1Δ*) (Park et al., 2014) were transformed with each separate TkCPT plasmid (pKG-GW2-HA-TkCPT1–8). KG 405 was also co-transformed with a combination of either pAGD425GPD-3xcMyc-TkCPLT1 (Epping et al., 2015) or pAGD425GPD-3xcMyc-TkCPLT2 (Niephaus et al., 2019) and one of the pKG-GW2-HA-TkCPT constructs.

SDS–PAGE and Western blot analysis

Yeast strains were cultivated overnight in 5 mL minimal synthetic defined (SD) dropout medium (leucine, methionine, or leucine + methionine) at 30°C. Following protein extraction (Bröker et al., 2018) and separation by SDS–PAGE on 12% polyacrylamide gels, the proteins were transferred to nitrocellulose membranes (Towbin et al., 1979) and stained with 0.1% Ponceau S in 5% acetic acid to confirm protein transfer. The blots were probed with two primary antibodies: mouse anti-hemagglutinin specific for HA-tagged proteins (Merck, Darmstadt, Germany; catalog #H3663, dilution 1:10 000) and mouse anti-cMyc specific for the 3xcMyc-tagged proteins (Bio-Rad Laboratories; catalog #MCA2200, dilution 1:1000). The goat anti-mouse IgG secondary antibody was coupled to either horseradish peroxidase (Thermo Fisher Scientific; catalog #32430, dilution 1:1500) or alkaline phosphatase (Merck; catalog #A3562-25ML, dilution 1:10 000). Bound antibodies were detected by chemiluminescence using SuperSignal West Dura Extended Duration Substrate (Thermo Fisher Scientific) followed by imaging using the G:BOX Chemi XX6 gel doc system (Syngene, Cambridge, UK).

Expression in yeast

A yeast strain optimized for terpenoid production (rox1::PGAL1-tHMGR PGAL10-ERG13 PERG9Δ::PCTR3 Bröker et al., 2020) was transformed with each separate TkCPT plasmid (pAG425_PGPD_TkCPT1–4_PTEF1_TkCPTL1, pAG425_PGPD_tTkCPT5–8). Following transformation, the yeast cells were plated on an SD medium lacking the corresponding amino acids and/or nucleotide supplements (Clontech, Mountain View, CA, USA) and incubated at 30°C. Clones were checked for integrity by colony PCR. For the expression of galactose-inducible genes, a single colony was picked, inoculated into 5 mL yeast SD medium, and cultivated overnight at 30°C on a rolling platform. From this culture, 50 mL of fresh YPD medium (containing 150 μM $CuSO_4$ when repressing the expression of ERG9) was inoculated to an OD_{600nm} of 0.2 and incubated at 30°C shaking at 140 rpm in a 250-mL Erlenmeyer flask. When the culture reached an OD_{600nm} of 0.4, the medium was changed to yeast extract peptone galactose medium to induce gene expression. For dolichol and polyisoprene analysis, yeast strains were grown at 20 and 30°C, respectively, for 72 h after induction in a 500-mL Erlenmeyer flask filled with 100 mL of galactose-containing medium (pH 4.5). Fresh medium was provided 48 h post-induction.

GC–MS analysis

Polyisoprenes and dolichols were extracted for GC–MS analysis as previously described (Rodriguez et al., 2014). Briefly, freeze-dried yeast cells were incubated at 80°C in a water bath for 5 min after adding 1 mL 6% [w/v] KOH in methanol (Carl Roth, Karlsruhe, Germany) and 100 μg cholesterol as an internal standard (Sigma–Aldrich, St. Louis, MO, USA). We then added 1 mL of *n*-hexane (Carl Roth), vortexed the mixture, and transferred the upper phase to a new vial. This was followed by two further rounds of extraction using 500 μL *n*-hexane. The solvent was removed from the pooled extracts by evaporation and the sample residues were re-solubilized in 1 mL acetone (Carl Roth). The samples were analyzed on a GC–MS-QP 2010 Ultra (Shimadzu,

Duisburg, Germany) equipped with a 30 m Rtx-5MS column. After a 1-min hold at 120°C, the temperature was increased to 330°C at 21°C min⁻¹ (pressure = 58.8 kPa) followed by a hold at 330°C for 10 min for the analysis of triterpenes and sterols. Nerol was measured with a 1-min hold at 120°C, followed by increasing the temperature to 180°C at 25°C min⁻¹. After a 5-min hold at 180°C, the temperature was increased to 300°C at 15°C min⁻¹ and was held for 3 min. The temperature was then increased to 330°C at a rate of 75°C min⁻¹ and was held for 5 min (pressure = 36.7 kPa).

Dolichol/polyisoprene extraction and analysis by LC–MS, ThFFF, and ¹H-NMR

Dolichols and polyisoprenes were extracted for LC–MS and ¹H-NMR analysis as previously described (Epping et al., 2015), with slight modifications. Freeze-dried yeast cells were extracted with 2:1 chloroform:methanol and *n*-hexane by accelerated solvent extraction using a Dionex ASE 350 (Thermo Fisher Scientific). The program included three 10-min cycles at 40°C for each solvent. All extracts were evaporated to dryness. For dolichol analysis, the dried chloroform:methanol extracts were re-solubilized in 20 mL methanol containing 6% [w/v] KOH, followed by saponification at 80°C for 2 h. The dolichols were then extracted three times with 15 mL *n*-hexane, evaporated to dryness, and re-dissolved in 5 mL methanol. For further purification, the solution was loaded onto Oasis HLB Plus LP extraction cartridges (Waters, Eschborn, Germany) after pre-equilibration of the cartridges with 5 mL methanol. The cartridges were washed with 10 mL methanol and dolichols were eluted in 5 mL *n*-hexane. The samples were evaporated to dryness and redissolved in 1 mL *n*-hexane.

LC–MS analysis was carried out using a Dionex Ultimate 3000RS UHPLC system coupled to a UV detector (Thermo Fisher Scientific), electrospray ionization (ESI) source, and amaZon speed ion trap mass spectrometer (Bruker, Bremen, Germany) or using an Acquity Premier LC system (Waters Corporation, UK) coupled to a Synapt XS 4k (Waters Corporation, UK) ion mobility time-of-flight mass spectrometer. The dolichols and polyprenols were separated on a Reprosil Pur Basic C18 column (5 μm, 4 × 250 mm; Altmann Analytik, Munich, Germany) at a flow rate of 0.8 mL min⁻¹, and the column oven temperature was set to 25°C. Samples were eluted from the column in a gradient of solvents A (60:20:20 methanol:acetonitrile:water) and B (100% ethanol) both containing 2 mM ammonium acetate. The samples were separated for 48 min using the following elution profile: 0–2 min, isocratic 100% A; 2–12 min, linear 0–100% B; 12–42 min, isocratic 100% B; followed by column re-equilibration: 42–43 min, linear 100–0% B; 43–48 min, isocratic 100% A. A 1:1 splitter (Accurate, Dionex Corporation, Sunnyvale, CA, USA) was used to spit the eluent to the UV detector (210 nm) and ESI-MS (alternating positive and negative mode with capillary voltage 4 kV, end plate offset voltage 500 V, nebulizer pressure 15 psi, dry gas flow rate 8 L min⁻¹ and dry temperature 200°C). Mass spectra were acquired over a scan range of *m/z* 50–2000 using enhanced resolution scan mode and were analyzed using Data Analysis v4.1 (Bruker, Bremen, Germany).

For ¹H-NMR-analysis, evaporated *n*-hexane extracts were dissolved in deuterated dichloromethane (CD_2Cl_2) using the same batch of solvent for each sample, followed by ¹H-NMR spectroscopy at 500 MHz as previously described (Post et al., 2014). Polymers from cells overexpressing TkCPT1/TkCPTL1 or TkCPT2/TkCPTL1 were fractionated using a TF2000 model ThFFF system coupled online to a model PN 3621 MALS detector and a model PN3510 ELSD (Postnova Analytics, Landsberg am Lech, Germany). The refractive index increment (dn/dc) for the polyisoprene samples was 0.104 mL g⁻¹ in toluene. We pumped toluene

as a carrier liquid through the ThFFF channel (45.6 cm × 2.0 cm × 250 μm) at a flow rate of 0.2 mL min^{-1}. Samples (approximately 1.6 mg mL^{-1}) were introduced into the system using a PN 5300 autosampler (Postnova Analytics) with a 100-μL sample loop and the injected volume was approximately 100 μL. The samples were run using a temperature gradient with a starting difference of 70 K, a focusing time of 5 min, and a total elution time of 90 min. The ELSD was calibrated by injecting 20 μL of a 470 000-Da standard (PSS Polymer Standards Service, Mainz, Germany) at a concentration of 2 mg mL^{-1} in toluene, a temperature difference of 5 K with no focusing time, and a 30 min elution time. The MALS detector was calibrated by injecting 50 μL of a 47 300-Da standard (PSS Polymer Standards Service) at a concentration of 10 mg mL^{-1} in toluene, a constant temperature difference of 70 K for 50 min, a 15 min linear gradient, and a focusing time of 5 min.

AUTHOR CONTRIBUTIONS

BM, EN, and CSG conceived and designed the experiments. BM and EN performed the experiments and collected data, BM, EN, and CSG validated and carried out formal data analysis. BM and EN wrote the paper and visualized the data CSG, DP, and RMT critically revised and edited the manuscript. JNB cultivated yeast and prepared extracts for NMR spectroscopy and WE performed NMR spectroscopy. CSG and DP supervised the project. The LC–MS system from Waters was funded by DFG grant GZ: INST 211/1037-1 LAGG. Part of the work was funded by the Federal Ministry for Education and Research (grant number: 031B0059C).

ACKNOWLEDGMENTS

CONFLICT OF INTEREST STATEMENT

Author RMT was employed by the company TRM Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

The authors certify that all data presented in this report are available either in the supplemental tables, figures, or data or in publicly available databases.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. BiFC analysis of homomeric TkCPTs.

Figure S2. Thermal field flow fractionation of extracts from yeast cells expressing TkCPT1/TkCPTL1 and TkCPT2/TkCPTL1.

Figure S3. Analysis of *n*-hexane extracts from yeast expressing TkCPT1/TkCPTL1 and control strains by NMR spectroscopy.

Figure S4. LC–MS chromatogram with dolichol standard.

Figure S5. Geranylgeraniol standard, hexaprenol standard, and TkCPT5 product detected by GC–MS and LC–MS analysis.

Table S1. List of proteins described in this study.

Table S2. Codon-harmonized sequences for TkCPT2 and TkCPTL1.

Table S3. List of primers used in this study.

Table S4. Primer efficiencies and formamide application for qRT-PCR.

Data S1. Chloroplast transit peptide sequence of the tobacco RuBisCO protein (NtRuBisCO-ChlTP).

REFERENCES

Alberti, S., **Gitler, A.D.** & **Lindquist, S.** (2007) A suite of gateway cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. *Yeast*, **10**, 913–919. Available from: https://doi.org/10.1002/yea.1502

Akhtar, T.A., **Matsuba, Y.**, **Schauvinhold, I.**, **Yu, G.**, **Lees, H.A.**, **Klein, S.E.** *et al.* (2013) The tomato *cis*-prenyltransferase gene family. *The Plant Journal*, **73**, 640–652.

Akhtar, T.A., **Surowiecki, P.**, **Siekierska, H.**, **Kania, M.**, **van Gelder, K.**, **Rea, K.A.** *et al.* (2017) Polyprenols are synthesized by a plastidial *cis*-prenyltransferase and influence photosynthetic performance. *The Plant Cell*, **29**, 1709–1725.

Almagro Armenteros, J.J., **Salvatore, M.**, **Emanuelsson, O.**, **Winther, O.**, **von Heijne, G.**, **Elofsson, A.** *et al.* (2019) Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance*, **2**, e2010900429.

Asawatreratanakul, K., **Zhang, Y.-W.**, **Wititsuwannakul, D.**, **Wititsuwannakul, R.**, **Takahashi, S.**, **Rattanapittayaporn, A.** *et al.* (2003) Molecular cloning, expression and characterization of cDNA encoding *cis*-prenyltransferases from *Hevea brasiliensis*—a key factor participating in natural rubber biosynthesis. *European Journal of Biochemistry*, **270**, 4671–4680.

Benghezal, M., **Wasteneys, G.O.** & **Jones, D.A.** (2000) The C-terminal dilysine motif confers endoplasmic reticulum localization to type I membrane proteins in plants. *The Plant Cell*, **12**, 1179–1201.

Brasher, M.I., **Surmacz, L.**, **Leong, B.**, **Pitcher, J.**, **Swiezewska, E.**, **Pichersky, E.** *et al.* (2015) A two-component enzyme complex is required for dolichol biosynthesis in tomato. *The Plant Journal*, **82**, 903–914.

Bröker, J.N., **Laibach, N.**, **Müller, B.**, **Prüfer, D.** & **Schulze Gronover, C.** (2018) *Taraxacum brevicorniculatum* rubber elongation factor TbREF associates with lipid droplets and affects lipid turn-over in yeast. *Biotechnology Reports*, **20**, e00290.

Bröker, J.N., **Müller, B.**, **Prüfer, D.** & **Schulze Gronover, C.** (2020) Combinatorial metabolic engineering in *Saccharomyces cerevisiae* for the enhanced production of the FPP-derived sesquiterpene Germacrene. *Bioengineering*, **7**, 135. Available from: https://doi.org/10.3390/bioengineering7040135

Brown, D., **Feeney, M.**, **Ahmadi, M.**, **Lonoce, C.**, **Sajari, R.**, **Di Cola, A.** *et al.* (2017) Subcellular localization and interactions among rubber particle proteins from *Hevea brasiliensis*. *Journal of Experimental Botany*, **68**, 5045–5055.

Buszewicz, D., **Kowalewska, Ł.**, **Mazur, R.** *et al.* (2021) HSFA1 proteins mediate heat-induced accumulation of CPT7-derived polyprenols affecting thylakoid organization. *bioRxiv*. Available from: https://doi.org/10.1101/2021.12.22.473876

Chen, C.-C., **Zhang, L.**, **Yu, X.**, **Ma, L.**, **Ko, T.-P.** & **Guo, R.-T.** (2020) Versatile *cis*-isoprenyl diphosphate synthase superfamily members in catalyzing carbon–carbon bond formation. *ACS Catalysis*, **10**, 4717–4725.

Edani, B., **Grabińska, K.A.**, **Zhang, R.**, **Park, E.J.**, **Siciliano, B.**, **Surmacz, L.** *et al.* (2020) Structural elucidation of the heterodimeric cis-prenyltransferase NgBR/DHDDS complex reveals novel insights in regulation of protein glycosylation. *bioRxiv*. 2020.06.03. Available from: https://doi.org/10.1101/2020.06.03.132209

Epping, J., **van Deenen, N.**, **Niephaus, E.**, **Stolze, A.**, **Fricke, J.**, **Huber, C.** *et al.* (2015) A rubber transferase activator is necessary for natural rubber biosynthesis in dandelion. *Nature Plants*, **1**, 15048.

Ernst, A.M., **Jekat, S.B.**, **Zielonka, S.**, **Müller, B.**, **Neumann, U.**, **Rüping, B.** *et al.* (2012) Sieve element occlusion (SEO) genes encode structural phloem proteins involved in wound sealing of the phloem. *Proceedings of the National Academy of Sciences*, **109**, E1980–E1989.

Gawarecka, K., **Siwinska, J.**, **Poznanski, J.**, **Onysk, A.**, **Surowiecki, P.**, **Sztompka, K.** *et al.* (2022) cis-prenyltransferase 3 and α/β-hydrolase are new determinants of dolichol accumulation in Arabidopsis. *Plant, Cell &*

Environment, **45**, 479–495. Available from: https://doi.org/10.1111/pce.14223

Grabińska, K.A., Cui, J., Chatterjee, A., Guan, Z., Raetz, C.R.H., Robbins, P.W. et al. (2010) Molecular characterization of the *cis*-prenyltransferase of *Giardia lamblia*. *Glycobiology*, **20**, 824–832.

Grabińska, K.A., Edani, B.H., Park, E.J., Kraehling, J.R. & Sessa, W.C. (2017) A conserved C-terminal RXG motif in the NgBR subunit of *cis*-prenyltransferase is critical for prenyltransferase activity. *The Journal of Biological Chemistry*, **292**, 17351–17361.

Grabińska, K.A., Park, E.J. & Sessa, W.C. (2016) *cis*-Prenyltransferase: new insights into protein glycosylation, rubber synthesis, and human diseases. *The Journal of Biological Chemistry*, **291**, 18582–18590.

Harrison, K.D., Park, E.J., Gao, N., Kuo, A., Rush, J.S., Waechter, C.J. et al. (2011) Nogo-B receptor is necessary for cellular dolichol biosynthesis and protein *N*-glycosylation. *The EMBO Journal*, **30**, 2490–2500.

Hofmann, N.R. (2017) The who, what, and where of plant polyprenol biosynthesis point to thylakoid membranes and photosynthetic performance. *The Plant Cell*, **29**, 1552–1553.

Ignea, C., Raadam, M.H., Motawia, M.S., Makris, A.M., Vickers, C.E. & Kampranis, S.C. (2019) Orthogonal monoterpenoid biosynthesis in yeast constructed on an isomeric substrate. *Nature Communications*, **10**(1), 3799. Available from: https://doi.org/10.1038/s41467-019-11290-x

Jach, G., Pesch, M., Richter, K., Frings, S. & Uhrig, J.F. (2006) An improved mRFP1 adds red to bimolecular fluorescence complementation. *Nature Methods*, **3**, 597–600.

Kera, K., Takahashi, S., Sutoh, T., Koyama, T. & Nakayama, T. (2012) Identification and characterization of a *cis*,*trans*-mixed heptaprenyl diphosphate synthase from *Arabidopsis thaliana*. *The FEBS Journal*, **279**, 3813–3827.

Kharel, Y. & Koyama, T. (2003) Molecular analysis of *cis*-prenyl chain elongating enzymes. *Natural Product Reports*, **20**, 111–118.

Kopcsayová, D. & Vranová, E. (2019) Functional gene network of prenyltransferases in *Arabidopsis thaliana*. *Molecules*, **24**, 4556.

Kruk, J. & Szymańska, R. (2023) Synthesis of natural polyprenols for the production of biological prenylquinones and tocochromanols. *RSC Advances*, **13**, 23122–23129. Available from: https://doi.org/10.1039/d3ra02872k

Kutsukawa, R., Imaizumi, R., Suenaga-Hiromori, M., Takeshita, K., Sakai, N., Misawa, S. et al. (2022) Structure-based engineering of a short-chain cis-prenyltransferase to biosynthesize nonnatural all-cis-polyisoprenoids: molecular mechanisms for primer substrate recognition and ultimate product chain-length determination. *The FEBS Journal*, **289**(15), 4602–4621. Available from: https://doi.org/10.1111/febs.16392

Kwon, M., Kwon, E.J.G. & Ro, D.K. (2016) *Cis*-Prenyltransferase and polymer analysis from a natural rubber perspective. *Methods in Enzymology*, **576**, 121–145.

Laibach, N., Hillebrand, A., Twyman, R.M., Prüfer, D. & Schulze Gronover, C. (2015) Identification of a *Taraxacum brevicorniculatum* rubber elongation factor protein that is localized on rubber particles and promotes rubber biosynthesis. *The Plant Journal*, **82**, 609–620.

Lakusta, A.M., Kwon, M., Kwon, E.-J.G., Stonebloom, S., Scheller, H.V. & Ro, D.-K. (2019) Molecular studies of the protein complexes involving *cis*-prenyltransferase in Guayule (*Parthenium argentatum*), an alternative rubber-producing plant. *Frontiers in Plant Science*, **10**, 165.

Lin, T., Xu, X., Ruan, J., Liu, S., Wu, S., Shao, X. et al. (2017) Genome analysis of *Taraxacum kok-saghyz* Rodin provides new insights into rubber biosynthesis. *National Science Review*, **5**, 78–87.

Müller, B., Noll, G.A., Ernst, A.M., Rüping, B., Groscurth, S., Twyman, R.M. et al. (2010) Recombinant artificial forisomes provide ample quantities of smart biomaterials for use in technical devices. *Applied Microbiology and Biotechnology*, **88**, 689–698.

Niephaus, E., Müller, B., van Deenen, N., Lassowskat, I., Bonin, M., Finkemeier, I. et al. (2019) Uncovering mechanisms of rubber biosynthesis in *Taraxacum koksaghyz*—role of *cis*-prenyltransferase-like 1 protein. *The Plant Journal*, **100**, 591–609.

Palamarczyk, G., Lehle, L., Mankowski, T., Chojnacki, T. & Tanner, W. (1980) Specificity of solubilized yeast glycosyl transferases for polyprenyl derivatives. *European Journal of Biochemistry*, **105**, 517–523.

Park, E.J., Grabińska, K.A., Guan, Z., Stránecký, V., Hartmannová, H., Hodaňová, K. et al. (2014) Mutation of Nogo-B receptor, a subunit of *cis*-prenyltransferase, causes a congenital disorder of glycosylation. *Cell Metabolism*, **20**, 448–457.

Post, J., van Deenen, N., Fricke, J., Kowalski, N., Wurbs, D., Schaller, H. et al. (2012) Laticifer-specific *cis*-prenyltransferase silencing affects the rubber, triterpene, and inulin content of *Taraxacum brevicorniculatum*. *Plant Physiology*, **158**, 1406–1417.

Pütter, K.M., van Deenen, N., Unland, K., Prüfer, D. & Schulze Gronover, C. (2017) Isoprenoid biosynthesis in dandelion latex is enhanced by the overexpression of three key enzymes involved in the mevalonate pathway. *BMC Plant Biology*, **17**, 88.

Qu, Y., Chakrabarty, R., Tran, H.T., Kwon, E.-J.G., Kwon, M., Nguyen, T.-D. et al. (2015) A lettuce (*Lactuca sativa*) homolog of human Nogo-B receptor interacts with *cis*-prenyltransferase and is necessary for natural rubber biosynthesis. *The Journal of Biological Chemistry*, **290**, 1898–1914.

Rodriguez, S., Kirby, J., Denby, C.M. & Keasling, J.D. (2014) Production and quantification of sesquiterpenes in *Saccharomyces cerevisiae*, including extraction, detection and quantification of terpene products and key related metabolites. *Nature Protocols*, **9**, 1980–1996. Available from: https://doi.org/10.1038/nprot.2014.132

Sallaud, C., Rontein, D., Onillon, S., Jabès, F., Duffé, P., Giacalone, C. et al. (2009) A novel pathway for sesquiterpene biosynthesis from *Z,Z*-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. *The Plant Cell*, **21**, 301–317.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, **12**, 5463–5467. Available from: https://doi.org/10.1073/pnas.74.12.5463

Schilmiller, A.L., Schauvinhold, I., Larson, M., Xu, R., Charbonneau, A.L., Schmidt, A. et al. (2009) Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. *Proceedings of the National Academy of Sciences*, **106**, 10865–10870.

Schmidt, T., Hillebrand, A., Wurbs, D., Wahler, D., Lenders, M., Schulze Gronover, C. et al. (2010) Molecular cloning and characterization of rubber biosynthetic genes from *Taraxacum koksaghyz*. *Plant Molecular Biology Reporter*, **28**, 277–284.

Surmacz, L., Plochocka, D., Kania, M., Danikiewicz, W. & Swiezewska, E. (2014) *cis*-Prenyltransferase AtCPT6 produces a family of very short-chain polyisoprenoids *in planta*. *Biochimica et Biophysica Acta*, **1841**, 240–250.

Surmacz, L. & Swiezewska, E. (2011) Polyisoprenoids—secondary metabolites or physiologically important superlipids? *Biochemical and Biophysical Research Communications*, **407**, 627–632.

Surowiecki, P., Onysk, A., Manko, K., Swiezewska, E. & Surmacz, L. (2019) Long-chain polyisoprenoids are synthesized by AtCPT1 in *Arabidopsis thaliana*. *Molecules*, **24**, 2789.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, **30**, 2725–2729.

Towbin, H., Staehelin, T. & Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences*, **76**, 4350–4354.

van Gelder, K., Rea, K.A., Virta, L.K.A., Whitnell, K.L., Osborn, M., Vatta, M. et al. (2018) Medium-chain polyprenols influence chloroplast membrane dynamics in *Solanum lycopersicum*. *Plant & Cell Physiology*, **59**, 2350–2365.

Visser, F., Müller, B., Rose, J., Prüfer, D. & Noll, G.A. (2016) Forizymes—functionalised artificial forisomes as a platform for the production and immobilisation of single enzymes and multi-enzyme complexes. *Scientific Reports*, **6**, 30839.

Vranová, E., Coman, D. & Gruissem, W. (2013) Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annual Review of Plant Biology*, **64**, 665–700.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. & Barton, G.J. (2009) Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.

Yamashita, S., Yamaguchi, H., Waki, T., Aoki, Y., Mizuno, M., Yanbe, F. et al. (2016) Identification and reconstitution of the rubber biosynthetic machinery on rubber particles from *Hevea brasiliensis*. *eLife*, **5**, e19022.

Zhang, H., Ohyama, K., Boudet, J., Chen, Z., Yang, J., Zhang, M. et al. (2008) Dolichol biosynthesis and its effects on the unfolded protein response and abiotic stress resistance in *Arabidopsis*. *The Plant Cell*, **20**, 1879–1898.