Polyketide reductases in defense-related parasorboside biosynthesis in *Gerbera hybrida* share processing strategies with microbial polyketide synthase systems

Lingping Zhu, Milla Pietiäinen, Juha Kontturi, Anna Turkkelin, Paula Elomaa and Teemu H. Teeri

Department of Agricultural Sciences, Viikki Plant Science Centre, University of Helsinki, Helsinki, 00014 UH, Finland

Summary

- Plant polyketides are well-known for their crucial functions in plants and their importance in the context of human health. They are synthesized by type III polyketide synthases (PKSs) and their final functional diversity is determined by post-PKS tailoring enzymes. *Gerbera hybrida* is rich in two defense-related polyketides: gerberin and parasorboside. Their synthesis is known to be initiated by GERBERA 2-PYRONE SYNTHASE 1 (G2PS1), but the polyketide reductases (PKRs) that determine their final structure have not yet been identified.
- We identified two PKR candidates in the pathway, GERBERA REDUCTASE 1 (GRED1) and GRED2. Gene expression and metabolite analysis of different gerbera tissues, cultivars, and transgenic gerbera plants, and *in vitro* enzyme assays, were performed for functional characterization of the enzymes.
- GRED1 and GRED2 catalyze the second reduction step in parasorboside biosynthesis. They reduce the proximal keto domain of the linear CoA-bound intermediate before lactonization.
- We identified a crucial tailoring step in an important gerbera PKS pathway and show that plant polyketide biosynthesis shares processing strategies with fungi and bacteria. The two tailoring enzymes are recruited from the ancient sporopollenin biosynthetic pathway to a defense-related PKS pathway in gerbera. Our data provide an example of how plants recruit conserved genes to new functions in secondary metabolism that are important for environmental adaptation.

Introduction

Plant polyketides represent one of the largest groups of structurally and functionally diverse secondary metabolites. They exhibit a wide range of bioactivities and play important roles in a variety of biological and physiological functions, including pigmentation of flowers and fruits (anthocyanins), UV photoprotection (flavonoids), defense against pathogens and herbivores (e.g. stilbenes and 2-pyrones), and pollen development (sporopollenin; Austin & Noel, 2003; Grienenberger et al., 2009; Neves et al., 2021). Many of them are regarded as important medicinal compounds or potential novel therapeutics, and their final functional diversity is determined by post-PKS tailoring enzymes (Shimizu et al., 2013; Lim et al., 2016; Abe, 2020; Bisht et al., 2021).

The core structures of the vast range of plant polyketides are biosynthesized by polyketide synthases (PKSs), which catalyze sequential decarboxylative condensations of an acyl-CoA starter with a malonyl-CoA extender, which is usually terminated by intramolecular cyclization (Austin & Noel, 2003; Morita et al., 2019). Polyketide synthases are multifunctional enzymes derived from fatty acid synthase (FAS), and they fall into three groups based on their architectural configuration (Austin & Noel, 2003). Fungal and bacterial polyketide biosynthesis relies mainly on large multi-domain type I and multi-enzyme type II PKSs. Besides essential functional domains or modules like acyl carrier proteins (ACP) and keto synthases (KS) for conducting starter loading and chain extension, they also contain optional tailoring domains or modules such as ketoreductases (KR), dehydratases (DH), enoyl reductases (ER) and cyclases (CYC; Hertweck, 2009; Neves et al., 2021). Selective manipulations of growing polyketide chains performed by these tailoring proteins are one of the main factors contributing to the diversity of fungal and bacterial polyketides (Hertweck, 2009; Neves et al., 2021). Type III PKSs, however, are simple homodimeric KSs that use a single active site to employ starter units and to perform iteratively decarboxylative condensations. The highly oxidated intermediates subsequently undergo cyclization by Claisen or aldol condensation, or by lactonization (Austin & Noel, 2003; Morita et al., 2019). The processing of polyketide backbones to achieve the final structure and bioactivity is performed by post-PKS tailoring enzymes (Shimizu et al., 2017; Morita et al., 2019). Type III PKSs are among the best-described enzymes in plants. So far, over 30 different plant PKS functions have been
characterized (Shimizu et al., 2017). Diverse polyketide backbones are synthesized by using different starters, catalyzing different lengths of extension reactions, and terminating via different modes of intramolecular cyclization (Stewart et al., 2013; Morita et al., 2019). However, the post-PKS tailoring enzymes that determine the final functional diversity of plant polyketides are largely unknown.

Polyketide reductases (PKRs) are important accessory enzymes in many plant polyketide biosynthesis pathways, but only a few cases have been investigated in any detail. Chalcone reductase from Fabaceae, the three-dimensional structure of which has been established, specifically reduces a keto group of the nonaromatized coumaryl-trione intermediate, the immediate Claisen cyclized product of the chalcone synthase catalyzed coumaryltetraketide (Bomati et al., 2005). In addition, the raspberry ketone/zingerone synthase (RZS) was found to reduce the PKS-synthesized diketide intermediate to raspberry ketone (Koeduka et al., 2011). Polyketide reductases are always involved in the lactone-forming PKS pathways. The lactone-forming PKSs synthesize unreduced 2-pyrones in in vitro reactions, while in planta they cooperate with PKRs to make reduced lactones (e.g. gerberin and parasorboside in gerbera) or catalyze different types of cyclization (e.g. to naphthoquinones; Eckermann et al., 1998; Jindaprasert et al., 2008; Grienenberger et al., 2010; Pietiäinen et al., 2016). For example, the Droophyllum lusitanicum PKS (DlHKS) was shown to produce a hexaketide 2-pyrene in vitro, and it has been suggested that it cooperates with an unknown PKR in plumbagin biosynthesis (Jindaprasert et al., 2008). To date, the only two identified PKRs in the plant lactone-forming PKS pathways are the Arabidopsis TETRAKETIDE α-PYRONE REDUCTASE 1 (TKPR1) and TKPR2. They catalyze the reduction of the carbonyl group of the PKS-synthesized tetraketide intermediates to hydroxylated 2-pyrene compounds, important precursors of sporopollenin (Grienenberger et al., 2010).

The ornamental plant gerbera (Gerbera hybrida) is rich in the polyketide-derived lactones gerberin, parasorboside and 4-hydroxy-5-methylcoumarin (HMC), which are mainly found in the tribe Mutisieae of Asteraceae (where Gerbera is located) and only sporadically in few other taxa (Murray, 1997; Koskela et al., 2011). These compounds have been shown to play key role in defense against pathogens (Koskela et al., 2011). They were also reported to be important bioactive compounds in some gerbera relatives (Nagumo et al., 1989; He et al., 2014), and they probably represent a major adaptive advantage of this branch of the Asteraceae family with respect to microbial attack and insect herbivores (Koskela et al., 2011). The GERBERA 2-PYRONE SYNTHASE 1 (G2PS1) was originally identified due to its sequence similarity to chalcone synthases (Helariutta et al., 1996). However, G2PS1 did not accept the bulky chalcone synthase starter molecule p-coumaroyl-CoA and it proved to be the first plant PKS described to use acetyl-CoA as a starter to synthesize a proposed intermediate, triacetolactone (TAL, 4-hydroxy-6-methyl-2-pyrene), in gerberin and parasorboside biosynthesis (Fig. 1), thus revealing a novel pathway in plant polyketide biosynthesis (Eckermann et al., 1998). Later, we identified two homologues of G2PS1, G2PS2 and G2PS3, which initiate the biosynthesis of HMC by using the same substrates as G2PS1 but undergoing a longer chain extension (Pietiäinen et al., 2016). More recently, we identified the anther-specific chalcone-synthesizing enzymes GASCL1 and GASCL2 from gerbera (Kontturi et al.,
They are orthologues of the Arabidopsis PKSA and PKSB, respectively, and likewise utilize medium- or long-chain acyl-CoA starters to synthesize tetraketide 2-pyrene intermediates in sporopollenin biosynthesis (Kim et al., 2010; Kontturi et al., 2017). One or two thus-far-unknown PKRs have been proposed to reduce the 2-pyrene ring of TAL to get the glycone form of gerberin and parasorboside (Fig. 1; Eckermann et al., 1998). Similarly, it is thought that an unknown PKR processes the unreduced intermediate 4,7-dihydroxy-5-methylcoumarin into HMC in gerbera (Piettiäinen et al., 2016).

In this study, we aimed to characterize the unknown PKRs in the PKS pathway leading to the synthesis of gerberin and parasorboside. We identified two TKPR2-like reductases encoded by **GERBERA REDUCTASE 1** (**GRED1**) and **GRED2** genes in gerbera. **GRED1** is co-expressed with **G2PS1**, and **GRED2** shows a high degree of sequence similarity with **GRED1**. We showed, by gene expression and metabolite analysis of different gerbera tissues, cultivars, and transgenic lines, and by *in vitro* enzyme assays, that the two reductases catalyze the second reduction step in parasorboside biosynthesis. However, **GRED1** and **GRED2** were shown to reduce not TAL, but a linear triketide intermediate before its lactonization, similar to the tailoring reactions in fungal and bacterial type I and type II PKS systems, indicating that the previously proposed pathway must be updated. Combined with our previous results (Zhu et al., 2021), our data shows that these two enzymes were recruited from the ancient sporopollenin biosynthetic pathway to a defense-related PKS pathway. This demonstrates how plants recruit conserved genes to perform new functions within secondary metabolism, contributing to their environmental adaptation.

**Materials and Methods**

**Plant material**

**Gerbera hybrida** (*Gerbera jamesonii* Bolus ex Adlam × *Gerbera viridifolia* Schultz-Bip) cultivars ‘Regina’, ‘President’, ‘Grizzly’ and ‘Pingpong’ were obtained from Terra Nigra BV (De Kwakel, the Netherlands) and grown under conditions described previously (Ruokolainen et al., 2010). Wild-type and transgenic gerbera plants were grown and multiplied in vitro (Helariutta et al., 2001). Developmental stages of gerbera inflorescences have been described previously by Helariutta et al. (1993).

**Identification and cloning of **GRED1** and **GRED2**

**GERBERA REDUCTASE 1** (**GRED1**) was originally identified as the transcript with the highest expression correlation to **G2PS1** and was annotated as a ‘reductase’ in gerbera microarray data (Laitinen et al., 2005). **GRED2** was further identified from the transcriptome data as a highly similar transcript to **GRED1**. **GRED1** and **GRED2** sequences (MW842919 and MW842920, respectively) were recovered in full length, and their coding sequences were amplified from the cultivar ‘Regina’ petal cDNA of development stage 6 (about half of the petals pigmented) with gene-specific primers (Supporting Information Table S1). The polymerase chain reaction (PCR) products containing the full-length open reading frames of **GRED1** and **GRED2** with *attB* flanking sequences were cloned into the pDONR221 vector through the Gateway BP reaction (Katzen, 2007). The created entry clones were verified by sequencing.

**RNAi and overexpression vector construction and gerbera transformation**

The gene silencing constructs for **GRED1** and **GRED2** were generated by LR recombination between the entry clones and the Gateway binary vector pK7GWIWG2D(II) (Karimi et al., 2002). The **GRED1** overexpression construct, under the control of the CaMV 35S promoter, was generated by LR recombination between the entry clone and the Gateway binary vector pK2GW7 (Karimi et al., 2002). These expression constructs were electroporated into *Agrobacterium tumefaciens* strain C58Cl(pGV2260) (Deblaere et al., 1985) and subsequently transformed into gerbera cultivar ‘Regina’ as described previously by Elomaa & Teeri (2001).

**Gene expression analysis with quantitative real-time polymerase chain reaction (qRT-PCR)**

Gerbera tissue samples were collected from the cultivar ‘Regina’ for the determination of expression patterns of **G2PS1**, **GRED1** and **GRED2**. Petal, carpel, ovary, papus, anther, receptacle, bract, and scape (the leafless inflorescence stem) samples were pooled samples corresponding to inflorescence development stages 2, 4, 6 and 8 (from emerging petals to fully open petals, see Helariutta et al., 1993). Young leaves of 8–9 cm in length and fully expanded mature leaves of 37–38 cm in length were sampled separately. *In vitro* leaves were from 4–5 cm tall *in vitro* plants growing on rooting medium. For the determination of expression levels of **GRED1** and **GRED2** in different gerbera cultivars, fully expanded mature leaves were collected from the gerbera cultivars ‘Regina’, ‘President’, ‘Pingpong’ and ‘Grizzly’. For the determination of **GRED1** and **GRED2** expression levels in gerbera transgenic lines, *in vitro* leaves were sampled from **GRED1** and **GRED2** RNAi transgenic lines, and mature leaf samples were collected from **GRED1** overexpression transgenic lines. Three biological replicates with three technical replicates from each sample were used for the expression analysis. Total RNA isolation was performed using a modified cetyltrimethylammonium bromide (CTAB) method, as described by Chang et al. (1993), followed by DNAase treatment using a NucleoSpin RNA Clean-up Mini Kit (740948.50; Macherey-Nagel, Düren, Germany) and verification in agarose gel that no DNA contamination remained. RNA quality and concentration was assessed using a Nanodrop Spectrophotometer (Thermo Scientific, Wilmington, NC, USA). One microgram of total RNA of each sample was applied in the synthesis of the first-strand cDNA, using a SuperScript III Reverse Transcriptase Kit (18 080 044; Invitrogen). Quantitative real-time polymerase chain reaction experiments were performed as described previously (Kontturi et al., 2017) with primers listed in Table S1; melting curves of the
PCR products are shown in Fig. S1. Relative expression values were calculated using the $2^{-\Delta\Delta C_T}$ method (Livak & Schmittgen, 2001), and the gerbera housekeeping gene GGAPDH was applied as a reference.

Gerberin and parasorboside content analysis
Fresh gerbera tissues corresponding to the samples used for gene expression analysis were collected from ‘Regina’, ‘Pingpong’, ‘President’, ‘Grizzly’, transgenic GRED1 and GRED2 RNAi and overexpression lines, with two biological replicates for each sample. Fresh samples were ground into powder in liquid nitrogen and freeze-dried. Metabolites were extracted from the dried powders using a 20X volume of 100% methanol (20 μl methanol for each 1 mg dry sample) overnight at 4°C. Extracts were centrifuged (17,000 g, 5 min) and the supernatants were collected for thin-layer chromatography (TLC) analysis with ethyl acetate : formic acid : acetic acid : water (91:10:10:23, v/v) as mobile phase (Yrjönä et al., 2002). Gerberin and parasorboside were visualized on the plates (TLC silica gel 60 F254; Merck, Darmstadt, Germany) after dipping in sulfuric acid : methanol (10:85, v/v) followed by heat treatment (120°C, 10 min; Aina-soja et al., 2008). Photographs of the TLC plates were converted to 8-bit grayscale images with Fiji IMAGEJ v.1.51, and the intensity of the spots corresponding to gerberin and parasorboside was measured to represent their amount. The percentage values of gerberin and parasorboside compared to the total amounts of the two compounds in each sample were calculated to represent the distribution between the two compounds.

Production of recombinant proteins
The recombinant G2PS1 was prepared in an Escherichia coli expression system, constructed as described previously (Pietiäinen et al., 2016). Recombinant GRED1 and GRED2 were prepared in a plant expression system. The coding sequences of GRED1 and GRED2 were cloned into the plant expression vector pEAQ-HT-DEST2 (Sainsbury et al., 2009) and electroporated into the A. tumefaciens strain C58C1 (pGV2260) (Deblaere et al., 1995). The recombinant proteins were transiently expressed in leaves of 6-wk-old Nicotiana benthamiana plants by agroinfiltration as described previously (Bashandy et al., 2015). The infiltrated leaves were sampled 3 d later, and soluble proteins were extracted from the leaf samples using cold extraction buffer (degassed 100 mM Tris-HCl, pH 7.5) supplemented with 0.2% β-mercaptoethanol and complete Mini EDTA free protease inhibitor cocktail (9693185901; Roche), homogenized on ice, then centrifuged for 17,000 g, 10 min at 4°C. Small molecules were removed with PD MiniTrap G-25 columns (28,918 007; GE Healthcare, Freiburg, Germany) before performing enzyme assays.

Enzyme assays
Enzyme assays for the reductases were conducted in a 100 μl reaction volume containing 100 mM HEPES-KOH buffer (pH 6.0), 1 mM NADPH (N2385; Sigma), 15–20 μg of tobacco protein extract containing recombinant GRED1 or GRED2, and 10 μl of substrate, incubated at 30°C for 1 h and terminated by adding 20 μl acetic acid. As candidate substrates for GRED1 and GRED2, TAL (H43415; Sigma) and gerberin aglycone (309 680; Sigma) were prepared as 1 mg ml⁻¹ stock in 100 mM HEPES-KOH buffer (pH 6.0). In addition to TAL and gerberin aglycone, a metabolite mixture was prepared from methanol extracts of gerbera leaves. Among other compounds, these extracts are rich in gerberin and parasorboside (Eckermann et al., 1998). Two 100 μl samples of methanol extracts were evaporated to dryness. One sample was dissolved in 40 μl HEPES-KOH buffer (pH 6.0) and was directly applied as substrate in the enzyme assays. The other sample was treated with β-glucosidase to produce aglucosyl substrates for the enzymatic assays. For this purpose, the sample was dissolved in 300 μl of 100 mM Na-acetate buffer (pH 5.0) containing 2 mg β-glucosidase (49 290; Sigma) and incubated for 3 h at 37°C. The reaction was stopped with 10 μl 4 M HCl and extracted twice with 350 μl ethyl acetate, evaporated to dryness in a vacuum centrifuge, and then dissolved in 40 μl 100 mM HEPES-KOH buffer (pH 6.0).

The products of enzyme assays with the candidate substrates were extracted, after stopping, twice with 200 μl ethyl acetate and evaporated to dryness. The dry products were dissolved in 10 μl methanol and analyzed with silica TLC with methanol : dichloromethane (15:85, v/v) as mobile phase. Consumption of gerberin and parasorboside compared to the total amounts of the two compounds in each sample were calculated to represent the distribution between the two compounds.

Results
Identification and cloning of GRED1 and GRED2
Previously, G2PS1 was identified and shown to encode the committed PKS which catalyzes the first step in gerberin and...
parasorboside biosynthesis in *G. hybrida* (Helariutta et al., 1996; Eckermann et al., 1998). To identify the accessory tailoring enzymes in the pathway, we searched for reductase-encoding genes that share a similar expression pattern with *G2PS1* from our early transcriptome data (Laitinen et al., 2005). By conducting Pearson correlation analysis, we identified a transcript containing the full-length coding sequence of a short-chain dehydrogenase/reductase (SDR) that was co-expressed with *G2PS1*, and we named it *GERBERA REDUCTASE 1* (GRED1).

In the transcriptome data, we discovered another transcript encoding a similar SDR gene that shared 85% nucleotide sequence identity with GRED1, and we named it GRED2. The two genes were considered as potential candidate genes encoding reductases in the gerberin and parasorboside biosynthetic pathway. The full-length coding sequences of GRED1 and GRED2 were amplified from gerbera petal cDNA. Both GRED1 and GRED2 contain a 960-bp open reading frame encoding polypeptides of 320 amino acids, sharing 84% amino acid sequence identity with each other. They are both similar (67% and 65% amino acid sequence identity, respectively) to the Arabidopsis tetraketide α-pyrone reductase encoding gene TKPR2 (Zhu et al., 2021).

**GRED1 and GRED2 show diverse expression patterns in gerbera**

In Arabidopsis, AtTKPR2 is a single-copy gene and is expressed specifically in anthers (Grienengerber et al., 2010). Preliminary analysis of the transcriptome data indicated that GRED1 and GRED2 are not anther specific and GRED1 is co-expressed with the widely expressed *G2PS1* gene. We determined the spatial expression profiles of GRED1, GRED2, and *G2PS1* using qRT-PCR. In 12 selected vegetative and reproductive tissues of the cultivar ‘Regina’, our results showed that GRED1 and GRED2 have diverse expression patterns, and neither was restricted to the anthers (Fig. 2a). The GRED1 expression profile across the tissues is highly similar to that of *G2PS1*, and it shows the highest expression in *in vitro* leaf, carpel, and ovary tissues, and somewhat lower expression in anthers and roots (Fig. 2a). GRED2 expression, however, is highest in the receptacle, scape, and *in vitro* leaves (Fig. 2a). The expression of both GRED1 and GRED2 was relatively high in *in vitro* leaves but was lower in young and mature leaves (Fig. 2a).

**The distribution of gerberin and parasorboside in gerbera tissues**

To explore whether the expression levels of candidate genes GRED1 and GRED2 are correlated with the biosynthesis of gerberin and parasorboside, we analyzed the spatial distribution of the two compounds in the gerbera tissues corresponding to those used for the gene expression analysis. Using TLC followed by sulfuric acid staining to analyze the metabolites (Yrjönen et al., 2002), gerberin and parasorboside could be separated and visualized. The grayscale intensities of spots corresponding to the two compounds in photographed TLC plates were determined to represent the distribution of the compounds in the samples in a semiquantitative way. The metabolite analysis showed that gerberin and parasorboside accumulated in nearly all gerbera tissues; however, they were observed to be present at very low concentrations in anthers and roots, where *G2PS1* (and GRED1) also showed the lowest expression levels (Fig. 2b). Moreover, the relative distribution of gerberin and parasorboside showed variation between the tissues (Fig. 2b, c). Parasorboside was preferentially accumulated in the young and *in vitro* leaves, receptacle, root, and scape tissues, while gerberin accumulated preferentially in bract, carpel, anther, pappus, ovary, and petal tissues (Fig. 2b, c). Distinct gerberin and parasorboside distribution patterns were observed between leaves of different stages (Fig. 2b, c). The *in vitro* leaves were shown to almost exclusively synthesize parasorboside, young leaves to produce more parasorboside than gerberin, and the mature leaves to accumulate more gerberin than parasorboside (Fig. 2b, c).

The distribution of gerberin and parasorboside in gerbera cultivars

We determined in an earlier study (Ainasoja et al., 2008) that while all gerbera cultivars produce gerberin and parasorboside in leaves and flowers, the ratio of these metabolites varies. To further explore the relationship between GRED1 and GRED2 expression and gerberin and parasorboside accumulation, we investigated the expression levels of GRED1 and GRED2 and the contents of the two compounds in mature leaves from gerbera cultivars ‘Regina’, ‘President’, ‘Pingpong’ and ‘Grizzly’. The expression levels of GRED1 and GRED2 varied between the cultivars (Fig. 2d). Compared with ‘Regina’, GRED1 was more strongly expressed in ‘Pingpong’ and ‘Grizzly’, while GRED2 showed higher expression levels in ‘President’ and ‘Pingpong’ (Fig. 2d). The gerberin and parasorboside content in mature leaves from these cultivars were analyzed with TLC. Cultivars ‘President’, ‘Pingpong’ and ‘Grizzly’ accumulated parasorboside preferentially, in contrast to ‘Regina’ that accumulated gerberin preferentially (Fig. 2e, f).

**Overexpression of GRED1 in gerbera results in an increase in parasorboside accumulation at the expense of gerberin in transgenic gerbera leaves**

To investigate the role of GRED1 and GRED2 in gerberin/parasorboside biosynthesis, we tried to produce transgenic gerbera lines overexpressing GRED1 and GRED2. We successfully obtained three transgenic lines transformed with the GRED1 overexpression construct but failed to get correct lines with the GRED2 construct. Two of the three transgenic lines, TR2 and TR3, showed significant overexpression of GRED1 in leaves, while GRED1 expression was reduced in TR1, a likely co-suppression effect (Fig. 3a). In all three lines, expression of GRED2 was lower than in the wild-type (Fig. 3a). We next analyzed the changes in gerberin and parasorboside distribution in these transgenic lines. Compared with the gerberin and parasorboside content in the mature leaves of the wild-type, there was an
Fig. 2 Expression patterns of G2PS1, GRED1 and GRED2, and distribution of gerberin and parasorboside in different gerbera tissues and cultivars. (a) Quantitative real-time polymerase chain reaction analysis of relative G2PS1, GRED1 and GRED2 expression levels in a gerbera tissue series. Expression was calculated using the $2^{-\Delta\Delta CT}$ method and is represented relative to expression levels in ‘Scape’, which was set at 1. Error bars (SE) were calculated from three biological replicates. Tissue series were from cv. ‘Regina’. Leaf (in vitro), leaves from 4–5 cm tall in vitro plants; Leaf (L), fully expanded mature leaves; Leaf (S), young leaves of 8–9 cm in length. (b) Thin-layer chromatography (TLC) analysis of methanol extracts from gerbera tissue series. G, gerberin; P, parasorboside. (c) Distribution of gerberin and parasorboside in the gerbera tissue series. Grayscale values for the spots corresponding to gerberin and parasorboside on the TLC image were measured to represent their amounts, and their proportions were calculated to represent the distribution of the two compounds. Error bars (SE) are calculated from two biological replicates. Red dashes represent the 50% level. (d) Quantitative real-time polymerase chain reaction analysis of relative GRED1 and GRED2 expression levels in fully expanded mature leaves of gerbera cultivars. Expression was calculated as described for (a) and is represented relative to expression levels in ‘Regina’, which was set at 1. Error bars (SE) were calculated from three biological replicates. (e) Thin-layer chromatography analysis of methanol extracts from fully expanded mature leaves of gerbera cultivars. (f) Distribution of gerberin and parasorboside in the cultivars. Error bars (SE) are calculated from two biological replicates.
increase in parasorboside content and a decrease in gerberin content in the leaves of transformants TR2 and TR3, where *GRED1* was overexpressed (Fig. 3a–c).

### Decreased parasorboside accumulation with increased gerberin accumulation in gerbera *GRED1* and *GRED2* downregulation lines

To further clarify the functions of *GRED1* and *GRED2* in gerberin and parasorboside biosynthesis, we applied a double-stranded RNA interference (RNAi) vector to produce *GRED1* and *GRED2* downregulated gerbera transgenic lines. The high degree of nucleotide sequence similarity of *GRED1* and *GRED2* resulted in cross-downregulation of both genes in many transgenic plants (Fig. 4a). The expression of *GRED1* and *GRED2* in *in vitro* leaves of these transgenic lines was analyzed, and we identified several lines that displayed suppression of *GRED1* and/or *GRED2* expression (Fig. 4a).

The *in vitro* leaves of nontransgenic cultivar ‘Regina’ accumulate parasorboside almost exclusively (Fig. 4b,c). However, the two severely *GRED1/2* double downregulated lines (*GRED2* RNAi TR8 and TR12) produced gerberin almost exclusively (Fig. 4b,c). This indicates that either *GRED1* or *GRED2* expression is necessary for the biosynthesis to proceed towards parasorboside. When both are missing, gerberin accumulates instead.

### Enzymatic activity of *GRED1* and *GRED2*

According to the previously proposed pathway (Eckermann et al., 1998; Fig. 1), TAL could be the potential substrate for the first PKR, and gerberin aglycone (or gerberin) the potential substrate for the second PKR on the pathway to the glucosides gerberin and parasorboside. To investigate the enzymatic activity of *GRED1* and *GRED2*, we conducted *in vitro* enzyme assays to test these potential substrates. Besides the commercially available TAL and gerberin aglycone, we also tested gerbera native metabolites prepared from leaf methanol extracts and their deglycosylated forms prepared by β-glucosidase treatment. Thin-layer chromatography analysis of the enzyme assay products showed that there is no detectable consumption of TAL, gerberin, or gerberin aglycone in reactions containing the cofactor NADPH and heterologously expressed *GRED1* or *GRED2*, compared with control reactions that contained *N. benthamiana* crude proteins only (Fig. 5a). This suggests that the substrates for *GRED1* or *GRED2* might not be the previously proposed TAL, gerberin aglycone or gerberin. Instead, the true substrate could be a thus-far-unrecognized intermediate in the pathway.

To explore this possibility, we reconstructed the biosynthetic pathway *in vitro*, by providing acetoacetyl-CoA, [2-¹⁴C]-labeled malonyl-CoA, and heterologously expressed G2PS1 in the enzyme assays to produce radiolabeled potential intermediate substrates, and then adding NADPH and *GRED1* or *GRED2*. The generated products were extracted with ethyl acetate and fractionated by TLC, and the radiolabeled products were visualized by exposure to photographic film. The result showed that G2PS1 synthesizes TAL, consistent with previously reported results (Eckermann et al., 1998). However, a new compound running slower than TAL was generated at the expense of TAL when providing the cofactor NADPH and *GRED1* or *GRED2* separately (or together) in the assays (Fig. 5b). The new
compound was not formed without NADPH or either of the two reductases (Fig. 5b), indicating that it was a reduced compound made by GRED1 and GRED2. Both reductases functioned with equal efficiencies and produced a compound with identical mobility in TLC (Fig. 5b).

As the product of the radiometric enzyme assay remained unknown, we tested whether methyl 3,5-dioxohexanoate could be reduced by GRED1 or GRED2. This molecule is a methyl ester of the linear triketide, resembling the CoA thioester of the G2PS1 product before its lactonization. Both GRED1 and GRED2 showed NADPH consumption when this substrate was provided (Table 1; Fig. S2).

**Discussion**

GRED1 and GRED2 are the PKRs necessary for parasorboside biosynthesis

Two closely related glucosidic lactones, gerberin and parasorboside, are known to be synthesized through a shared PKS pathway initiated by G2PS1, but the PKRs determining their final structure have remained unknown. In this work, we identified two PKRs that are important for parasorboside biosynthesis, the SDR superfamily reductases GRED1 and GRED2. GRED1 was identified by its co-expression pattern with G2PS1, and GRED2 was
distinguished through its high nucleotide sequence identity with GRED1. Expression and metabolite analysis indicated that GRED1 and GRED2 might be associated with the differences in distribution between gerberin and parasorboside in gerbera. We showed that the pathway preferentially synthesizes parasorboside in tissues and cultivars exhibiting high GRED1 or GRED2 expression, and tend to produce more gerberin in tissues and transgenic lines exhibiting low GRED1 and GRED2 expression. The strongest support came from transgenic gerbera lines in which downregulation of both GRED1 and GRED2 together abolished parasorboside accumulation, and only gerberin was made. These data suggest that GRED1 and GRED2 are the key branch-point PKRs in gerberin/parasorboside biosynthesis. In particular, they catalyze the second reduction step in the pathway, since no correlated changes in the total content of the two compounds were observed in the transgenic (upregulated or downregulated) lines.

GRED1 and GRED2 reduce the proximal keto domain of the linear triketide intermediate before lactonization

In the previously proposed gerberin/parasorboside biosynthetic pathway, it was proposed that one or two reductions in the 2-pyrone ring of TAL, performed by one or two missing PKRs, were required for the formation of gerberin and parasorboside (Eckermann et al., 1998; Fig. 1). However, TAL, gerberin aglycone or gerberin could not be reduced by GRED1 or GRED2 in our in vitro enzyme assays. This suggested that none of these lactones is the substrate for GRED1 or GRED2, and our previously proposed pathway needs an update. Another possibility is that the PKRs act on linear intermediates before their lactonization. Although the unreduced diketide acetoacetyl-CoA is readily extended by G2PS1, 3-hydroxybutryl-CoA and crotonyl-CoA (reduced diketides) are not (Eckermann et al., 1998). The remaining linear intermediate is the triketide 3,5-dioxohexanoic-CoA, where the distal keto group is first reduced to produce 5-hydroxy-3-oxohexanoic-CoA. The proximal group is then reduced by the second PKR (GRED1 or GRED2) to generate 3,5-dihydroxyhexanoic-CoA (Fig. 6a). These reduced linear intermediates are subsequently lactonized to gerberin aglycone and parasorboside aglycone, respectively (Fig. 6a). In this scenario, TAL is a derailment product rather than an intermediate in gerberin and parasorboside biosynthesis.

Similar reactions have been reported in other PKS pathways. An unidentified KR was shown to be involved in 6-

### Table 1 Reduction of methyl 3,5-dioxohexanoate by GRED1 and GRED2.

|                  | NADPH consumption | Relative activity |
|------------------|-------------------|------------------|
| Blank (no extract) | 0.051             | –                |
| Control (tobacco extract) | 0.091             | 1                |
| Tobacco extract containing GRED1 | 0.154             | 2.6              |
| Tobacco extract containing GRED2 | 0.250             | 5.0              |

NADPH consumption was measured by following the decrease in absorbance at 340 nm for 30 min.
hydroxymellein biosynthesis in carrot, reducing the distal keto group of a triketide intermediate to form a 5-hydroxy-3-oxohexanoic-CoA intermediate, and TAL was produced as a derailment product when NADPH was absent (Kurosaki et al., 2002). This KR catalyzes a similar reaction to the proposed first (as yet uncharacterized) PKR catalyzed reaction in gerberin biosynthesis (Fig. 6a). In 6-methylsalicylic acid (6-MSA) biosynthesis in fungus, the triketide-CoA was lactonized to TAL instead of the 3-hydroxy intermediate, when the KR domain of the type I PKS was mutated or blocked by the absence of NADPH (Scott et al., 1974; Campuzano & Shoolingin-Jordan, 1998). This KR catalyzes the same reaction as GRED1 and GRED2 in the updated parasorboside biosynthesis scheme and, furthermore, GRED1 and GRED2 readily reduce the linear triketide substrate analog methyl 3,5-dioxohexanoate (Fig. 6). Recently, 5-hydroxyhexanoic acid 3-O-β-D-glucoside was identified in and isolated from gerbera leaves, and it was suggested that it is synthesized through the shared gerberin/parasorboside biosynthetic pathway (Mascellani et al., 2021). This finding supports our updated pathway and suggests that PKRs do indeed act on the linear triketide before lactonization, and the reduced triketide either lactonizes to parasorboside aglycone or is released and

Fig. 6 Updated in vivo and in vitro biosynthetic pathways to gerberin and parasorboside. (a) Our results indicate that reduction of the triketide backbone leading to gerberin and parasorboside takes place before lactonization; this idea is supported by the presence of the reduced linear form as a glucoside in gerbera tissues. (b, c) Consequently, the in vitro products of GRED1 and GRED2 would be 3-hydroxy derivatives of the substrates. G2PS1, gerbera 2-pyrone synthase 1; GA, gerberin aglycone; PA, parasorboside aglycone; PKR, polyketide reductase; TAA-CoA, triacetic acid-CoA; TAL, triacetolactone. Polyketide reductases and proposed reduced substituents are marked in red. Dashed arrows refer to speculated reactions.
glucosylated, converting to the 5-hydroxyhexanoic acid 3-O-β-D-glucoside (Fig. 6a). Many gerbera cultivars are rich in both gerberin and parasorboside, and both gerberin aglycone and parasorbose aglycone accumulate in the plant under stress (Koskela et al., 2011), but neither TAL nor its glucoside has ever been detected in gerbera.

The updated pathway suggests that all PKS pathways share similar processing strategies

In bacteria and fungi, functional modules like KR, DH, ER and CYC are commonly present in type I and II PKS systems. They act on the linear intermediates and are required for the correct region-specific cyclization of the growing chain (Fischbach & Walsh, 2006; Hertweck, 2009; Neves et al., 2021). The production of lactones as derailment products has been observed in many type I and II PKS initiated pathways that the KR was missing or blocked, such as in the type I PKS catalyzed 6-MSA biosynthesis (Campuzano & Shoolingin-Jordan, 1998) or type II PKS initiated actinorhodin biosynthesis, in which an SDR type I PKS initiated many type I and II PKS initiated pathways when the KR was not involved (Hadfield et al., 2004). The lactonization of linear polyketide thioesters into 2-pyrones is in fact a spontaneous reaction (Light et al., 1966; Springob et al., 2007). This explains why a lactone is frequently produced as a derailment product in PKS pathways when the system is absent of tailoring enzyme functions. Polyketide reductases acting on linear intermediates before cyclization in the updated gerberin/parasorboside pathway highly resemble the tailoring reactions performed by optional KR subunits in type I and II PKS systems. Similar reactions are thought to occur in many other lactone-forming type III PKSs. For instance, unknown PKRs are supposed to be involved in the Plumbago PKS or Drosophyllum HKS initiated natural polyketide biosynthesis pathways (Springob et al., 2007; Jindaprasert et al., 2008). These plant PKSs were again shown to make lactones in vitro instead of producing the expected intermediates in the original plants. Clearly, the accessory PKRs in the lactone forming PKS initiated pathways play a crucial function in stabilizing the linear intermediates and preventing them from unspecific cyclization during their biosynthesis. The similarity in tailoring reactions among the type I, II and III PKS pathways is also supported by observations of the first plant polyketide CYC, olivetolic cyclase (OAC), which is involved in cannabinoid biosynthesis in Cannabis sativa (Gagne et al., 2012). Olivetolic cyclase is structurally similar to type II PKS associated CYC and catalyzes cyclization of the linear tetraketide intermediate to olivetolic acid, instead of the derailment products olivetol and 2-pyrones, a similar reaction to that which occurs in type II PKS system catalyzed pathways (Gagne et al., 2012).

GRED1 and GRED2 were recruited from the conserved sporopollenin biosynthesis pathway to defense-related polyketide biosynthesis

GRED1 and GRED2 are both close orthologues of the Arabidopsis TKPR2, which encodes one of the two anther-specific PKRs involved in sporopollenin biosynthesis. As a major structural element of pollen walls, sporopollenin is crucial for the reproductive success of land plants (Wallace et al., 2011; Quilichini et al., 2015). Sporopollenin has been shown to be synthesized through an ancient PKS pathway constructed from a set of functionally conserved enzymes (Kim & Douglas, 2013). In Arabidopsis, two PKRs, the predominant AtTKPR1 and the minor AtTKPR2, reduce the PKS synthesized tetraketide 2-pyrones into hydroxylated 2-pyrones, which are important sporopollenin precursors (Grienenberger et al., 2010). Gerbera has a single TKPR1 orthologue (GTKPR1) that is anther-specific (Zhu et al., 2021), and two non-anther-specific AtTKPR2 orthologues, GRED1 and GRED2, which are described here. In our previous work we showed that while GTKPR1 is the predominant PKR in sporopollenin biosynthesis (as AtTKPR1 is in Arabidopsis), the TKPR2 orthologues GRED1 and GRED2 have maintained not only expression in the tapetum of anthers but also involvement in pollen wall biosynthesis to some extent (Zhu et al., 2021). However, the dramatically expanded expression patterns of GRED1 and GRED2 indicated that they have probably been recruited to novel functions out of sporopollenin biosynthesis, and in this study we show what these functions are. During evolution, GRED1 and GRED2 have been recruited from the ancient sporopollenin biosynthesis pathway to the defense-related polyketide biosynthesis pathway in gerbera.

Plants have adapted to changing environments and biotic challenges by producing a wide range of secondary metabolites. However, secondary metabolism associated protein folds are restricted (Weng et al., 2012). Gene duplication is believed to be the driver of metabolic innovation (Moghe & Last, 2015). The diversification of plant secondary metabolism is believed to be an evolutionary result of gene duplication events and subsequently functional divergence events (Moghe & Last, 2015). This mechanism is well reflected in the evolution of plant polyketide biosynthesis. Plant PKSs have shared protein folds but synthesize diverse polyketide backbones, relying on slight variations of their catalytic cavity. They have been shown to be derived from FAS and have acquired functional diversity through gene duplication and mutation events (Jenke-Kodama et al., 2008; Xie et al., 2016; Yonekura-Sakakibara et al., 2019). The sporopollenin biosynthesis associated PKSA/B clade is one of the two early clades of plant PKSs that emerged before the divergence of Bryophytes and Tracheophytes, as shown by a kingdom-wide evolution analysis of plant PKSs (Naake et al., 2021). TKPR1s and TKPR2s, on the other hand, form two clades of plant SDR superfamily enzymes. They belong to the SDR108E subfamily that harbors many plant secondary metabolite biosynthesis associated reductases, like the dihydroflavonol 4-reductase and anthocyanidin reductase in flavonoid biosynthesis, and cinnamoyl-CoA reductase in lignin biosynthesis (Moummou et al., 2012). In evolution, TKPR2 appeared later than TKPR1, but earlier than many defense-related polyketide biosynthesis associated reductases (Moummou et al., 2012), which supports the possibility that TKPR2 has been recruited from the ancient sporopollenin biosynthesis pathway to other PKS pathways not only in gerbera, but maybe also in other plants (Zhu et al., 2021). Our data show how plants recruit conserved genes to perform new functions.
within secondary metabolism, contributing to their environmental adaptation. This is evolution in front of our eyes.

Acknowledgements

We gratefully acknowledge Anu Rokkanen and Eija Takala for excellent technical assistance, the glasshouse team led by Sanna Peltola for excellent care of the plants, and Nina Sipari from the Viikki Metabolomics Unit (Helsinki Institute of Life Science, University of Helsinki and Biocenter Finland) for her expertise with the liquid chromatography–mass spectroscopy analyses. We would like to thank Teng Zhang for providing helpful discussion and encouragement in performing this research, and Antti Lehtinen for participating in gerbera tissue sampling. This research was supported by the Academy of Finland (grant 139513 for THT). We would also like to thank the China Scholarship Council for providing a grant for LZ’s PhD work.

Author contributions

LZ and THT designed the research; LZ, MP, JK and AT carried out the experiments; LZ and THT analyzed the results; LZ, THT and PE wrote the manuscript with input from all authors. All authors read and approved the final manuscript.

ORCID

Paula Elomaa https://orcid.org/0000-0001-6512-0810
Milla Pietiäinen https://orcid.org/0000-0002-4875-0682
Teemu H. Teeri https://orcid.org/0000-0002-3812-7213
Lingping Zhu https://orcid.org/0000-0001-6959-7755

Data availability

The nucleotide sequences of GRED1 and GRED2 have been deposited to GenBank with accession numbers MW842919 and MW842920, respectively.

References

Abe I. 2020. Biosynthesis of medically important plant metabolites by unusual type III polyketide synthases. *Journal of Natural Medicines* 74: 639–646.
Ainasoo MM, Pohjala LL, Tammela PS, Somervuo PJ, Vuorela PM, Teeri TH. 2008. Comparison of transgenic *Gerbera hybridoides* lines and traditional varieties shows no differences in cytotoxicity or metabolic fingerprints. *Transgenic Research* 17: 793–803.
Austin MB, Noel JP. 2003. The chalcone synthase superfamily of type III polyketide synthases. *Natural Product Reports* 20: 79–110.
Bashandy H, Jalkanen S, Teeri TH. 2015. Within leaf variation is the largest source of variation in agroinfiltration of *Nicotiana benthamiana*. *Plant Methods* 11: 1–7.
Bish R, Bhattacharrya A, Shrivastava A, Saxena P. 2021. An overview of the medicinally important plant type III PKS derived polyketides. *Frontiers in Plant Science* 12: 2155.
Bomati EK, Austin MB, Bowman ME, Dixon RA, Noel JP. 2005. Structural elucidation of chalcone reductase and implications for deoxychalcone biosynthesis. *Journal of Biological Chemistry* 280: 30496–30503.
Campuzano LDG, Shoolingin-Jordan PM. 1998. Incubation of 6-methylsalicylic acid synthase with alternative starter units in the absence of NADPH and the identification of the resulting triacetic acid lactones. *Biochemical Society Transactions* 26: S284.
Chang S, Puryear J, Cairney J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11: 113–116.
Deblaere R, Bytebier B, De Greve H, Deboeck F, Schell J, Van Montagu M, Leemans J. 1985. Efficient octopine Ti plasmid-derived vectors for *Agrobacterium*-mediated gene transfer to plants. *Nucleic Acids Research* 13: 4777–4788.
Deng X, Bashandy H, Ainasoo M, Kontturi J, Pietiäinen M, Laitinen RAE, Albert VA, Valkonen JPT, Elomaa P, Teeri TH. 2014. Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of *Gerbera hybridoides*. *New Phytologist* 201: 1469–1483.
Eckermann S, Schröder G, Schmidt J, Strack D, Edrada RA, Helariutta Y, Elomaa P, Kotilainen M, Kilpeläinen I, Proksch P. 1998. New pathway to polyketides in plants. *Nature* 396: 387–389.
Elomaa P, Teeri TH. 2001. Transgenic *Gerbera*: In: Bajaj YPS, ed. *Biotechnology in agriculture and forestry*, vol. 48. Berlin, Germany: Springer-Verlag, 139–154.
Fischbach MA, Walsh CT. 2006. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic machinery, and mechanisms. *Chemical Reviews* 106: 3468–3496.
Gagne SJ, Stout JM, Liu E, Boubakir Z, Clark SM, Page JE. 2012. Identification of olivetolic acid cyclase from *Cannabis sativa* reveals a unique catalytic route to plant polyketides. *Proceedings of the National Academy of Sciences, USA* 109: 12811–12816.
Grienenberger E, Kim SS, Lallemand B, Geoffroy P, Heintz D, de Azevedo SC, Heitz T, Douglas CJ, Legrand M. 2010. Analysis of TETRaketide α-PYrone Reductase function in Arabidopsis thaliana reveals a previously unknown, but conserved, biochemical pathway in sporopollenin monomer biosynthesis. *Plant Cell* 22: 4067–4083.
Hadfield AT, Limpkin C, Teartasin W, Simpson TJ, Crosby J, Crump MP. 2004. The crystal structure of the actIII actinorhodin polyketide reductase: proposed mechanism for ACP and polyketide binding. *Structure* 12: 1865–1875.
He F, Wang M, Gao M, Zhao M, Bai Y, Zhao C. 2014. Chemical composition and biological activities of *Gerbera ananthera*. *Molecules* 19: 4046–4057.
Helariutta Y, Elomaa P, Kotilainen M, Seppäläinen P, Teeri TH. 1993. Cloning of cDNA coding for dihydrodihydroflavon-4-reductase (DFR) and characterization of dfr expression in the corollas of *Gerbera hybridoides* var. Regina (Compositae). *Plant Molecular Biology* 22: 183–193.
Hiratake J, Watanabe B, Suzuki S, Hiratake J, Mano J, Yazaki K. 2011. Diversity of duplicated chalcone synthase genes in anthocyanin biosynthesis. *Frontiers in Plant Biology* 5: 1–8.
Hertweck C. 2009. The biosynthetic logic of polyketide diversity. *Angewandte Chemie International Edition* 48: 4688–4716.
Jenke-Kodama H, Müller R, Dittmann E. 2008. Evolutionary mechanisms underlying secondary metabolite diversity. *Progress in Drug Research* 65: 120–140.
Jindaprasert A, Springob K, Schmidt J, De-Eknamkul W, Kutchan TM. 2008. Pyrone polyketides synthesized by a type III polyketide synthase from *Drosophilus lausitanicus*. *Plosochemistry* 69: 3043–3053.
Karimi M, Inez D, Depicker A. 2002. GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science* 7: 193–195.
Katzen F. 2007. Gateway™ recombinational cloning: a biological operating system. *Expert Opinion on Drug Discovery* 2: 571–589.
Kim SS, Douglas CJ. 2013. Sporopollenin monomer biosynthesis in arabidopsis. *Journal of Plant Biology* 56: 1–6.
Kim SS, Grienenberger E, Lallemand B, Colpitts CC, Kim SY, de Azevedo SC, Geoffroy P, Heintz D, Krahn D, de Azevedo SC, Albert VA, Valkonen JPT, Elomaa P, Teeri TH. 2014. Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of *Gerbera hybridoides*. *New Phytologist* 201: 1469–1483.
Koeduka T, Watanabe B, Suzuki S, Hiratake J, Mano J, Yazaki K. 2011. Characterization of raspberry ketone/zingiberone synthase, catalyzing the alpha, beta-hydrogenation of phenylbutenones in raspberry fruits. *Biochemical and Biophysical Research Communications* 1: 104–108.
