Polyketide Reductases in *Gerbera hybrida*

Lingping Zhu

Doctoral Program in Plant Sciences (DPPS)
Doctoral School in Environmental, Food and Biological Sciences (YEB)

Department of Agricultural Sciences
Faculty of Agriculture and Forestry
University of Helsinki

ACADEMIC DISSERTATION

To be presented for public examination, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, in Hall 108 (LSB3), Forest Sciences Building, Latokartanonkaari 7, on the 6th of May, 2022 at 12 o’clock.

Helsinki, Finland. 2022
Supervisor  
**Professor Teemu Teeri**  
Department of Agricultural Sciences  
University of Helsinki, Finland

Follow-up group  
**Principal Scientist Heiko Rischer**  
VTT Technical Research Centre of Finland

Senior Scientist Milla Pietiäinen  
VTT Technical Research Centre of Finland

Pre-examiners  
**Senior Scientist Suvi Häkkinen**  
VTT Technical Research Centre of Finland

Associate professor Katja Karppinen  
UiT, The Arctic University of Norway

Opponent  
**Professor Markku Keinänen**  
Department of Environmental and Biological Sciences  
University of Eastern Finland, Finland

Custos  
**Professor Teemu Teeri**  
Department of Agricultural Sciences  
University of Helsinki, Finland

ISBN 978-951-51-8106-0 (PRINT)  
ISBN 978-951-51-8107-7 (ONLINE)  
ISSN 2342-5423 (PRINT)  
ISSN 2342-5431 (ONLINE)

Cover image: The lovely gerbera in the greenhouse (taken by Lingping Zhu).
万里归来颜愈少，微笑，笑时犹带岭梅香。

After years of strivings, you are returning with a young heart.

— 苏轼 《定风波》

— “Calming the Waves” by Su Shi
ABSTRACT

The vast ranges of chemically diverse plant polyketides are well-known for their important functions in plants and to human health. Plant polyketides consist of the most widely distributed pollen wall structural element sporopollenin, the ubiquitous and chemically diverse flavonoids, and various taxa-specific natural products. Despite the vast diversity, plant polyketides are synthesized in a common logic. Type III polyketide synthases (PKSs) synthesize the backbones and post-PKS tailoring enzymes provide accurate modifications to determine the final chemical diversity. Compared with plant PKSs that have been intensively investigated, the important tailoring enzymes are largely unknown.

The ornamental plant Gerbera hybrida is rich in two defense-related polyketides gerberin and parasorboside, in addition to sporopollenin and flavonoids. Gerberin/parasorboside biosynthesis has been shown to be initiated by gerbera 2-pyrene synthase 1 (G2PS1), while the accessory polyketide reductases (PKRs) have been missing. In this thesis, by coexpression analysis, we identified two PKR candidates in this pathway, gerbera reductase 1 (GRED1) and GRED2. They were shown to be the second PKR required for parasorboside biosynthesis by expression and metabolite analysis of different gerbera tissues, cultivars, and transgenic gerbera plants and in vitro enzyme assays. We also showed that PKRs act on the linear triketide intermediate prior to its lactonization in gerberin/parasorboside biosynthesis. This updated pathway indicates that plant polyketide biosynthesis shares tailoring strategies with fungi and bacteria.

Sporopollenin has been uncovered to be synthesized through a conserved anther-specific PKS pathway. We showed that GRED1 and GRED2 are orthologues of Arabidopsis tetraketide α-reductase 2 (AtTKPR2), a minor PKR involved in sporopollenin biosynthesis. Their important role in parasorboside biosynthesis indicates that the duplicate gerbera TKPR2s have been recruited from the ancient sporopollenin biosynthesis to a defense-related PKS pathway. The functional diversification of gerbera TKPR2s is an example of how plants get metabolic innovation for adaption during evolution. We showed that GRED1 and GRED2 still sustained the minor role in pollen wall formation. Moreover, we identified and characterized GTKPR1, the gerbera orthologue of AtTKPR1, the predominant PKR in sporopollenin biosynthesis. GTKPR1 was shown to play a conserved predominant role in pollen wall formation.

Cultivar-specific pigmentation of gerbera has been shown to be a result of substrate preferences of dihydroflavonol 4-reductases (DFRs). Substrate specificity of DFRs from different species have been investigated for decades. However, which amino acids are responsible for it remains unknown. Two major technical challenges hindering the investigation are: difficult to make heterologous DFRs and lack of a reliable assay for the
determination of substrate preferences. In this thesis, we established a highly efficient DFR expression system in tobacco and refined the traditional BuOH-HCl assay into a reliable and robust assay for the determination of DFR substrate preferences.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>1</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>3</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>5</td>
</tr>
<tr>
<td>LIST OF PUBLICATIONS</td>
<td>6</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>7</td>
</tr>
<tr>
<td>1.1 Plant polyketides: a vast group of structurally and functionally diverse natural products</td>
<td>7</td>
</tr>
<tr>
<td>1.1.1 Sporopollenin: the most widely distributed and functionally conserved plant polyketide</td>
<td>8</td>
</tr>
<tr>
<td>1.1.2 Flavonoids: the widely distributed and structurally diverse plant polyketides</td>
<td>8</td>
</tr>
<tr>
<td>1.1.3 Taxa-specific polyketides</td>
<td>9</td>
</tr>
<tr>
<td>1.2 Polyketide biosynthesis and the logic of chemo-diversity</td>
<td>9</td>
</tr>
<tr>
<td>1.2.1 Polyketide synthase (PKS) catalyzes the synthesis of polyketide backbone</td>
<td>10</td>
</tr>
<tr>
<td>1.2.2 Post-PKS tailoring enzymes determine the functional diversity</td>
<td>12</td>
</tr>
<tr>
<td>1.2.3 Polyketide reductase (PKRs): necessary tailoring enzymes in many plant PKS pathways</td>
<td>14</td>
</tr>
<tr>
<td>1.3 Polyketide biosynthesis in <em>Gerbera hybrida</em></td>
<td>14</td>
</tr>
<tr>
<td>1.3.1 <em>Gerbera hybrida</em>: Model plant for polyketide biosynthesis</td>
<td>14</td>
</tr>
<tr>
<td>1.3.2 Gerberin and parasorboside biosynthesis: PKRs remain unknown</td>
<td>16</td>
</tr>
<tr>
<td>1.3.3 Sporopollenin biosynthesis: TKPRs – Functional conservation or diversification?</td>
<td>16</td>
</tr>
<tr>
<td>1.3.4 Anthocyanin biosynthesis: molecular mechanism underlying substrate specificity of DFRs remains unknown</td>
<td>17</td>
</tr>
<tr>
<td>2. AIMS OF THE STUDY</td>
<td>19</td>
</tr>
<tr>
<td>3. MATERIALS AND METHODS</td>
<td>20</td>
</tr>
<tr>
<td>4. RESULTS AND DISCUSSION</td>
<td>21</td>
</tr>
<tr>
<td>4.1 PKRs in gerberin/parasorboside biosynthesis in <em>Gerbera hybrida</em></td>
<td>21</td>
</tr>
<tr>
<td>4.1.1 Identification of two PKR candidates in gerberin/parasorboside biosynthesis</td>
<td>21</td>
</tr>
<tr>
<td>4.1.2 GRED1 and GRED2 are the PKRs necessary for parasorboside biosynthesis</td>
<td>21</td>
</tr>
<tr>
<td>4.1.3 PKRs act on the linear triketide prior to lactonization, suggesting an update on the pathway leading to gerberin and parasorboside</td>
<td>22</td>
</tr>
<tr>
<td>4.2 PKRs in sporopollenin biosynthesis in <em>Gerbera hybrida</em> (I, II)</td>
<td>23</td>
</tr>
</tbody>
</table>
4.2.1 Three PKRs, GTKPR1, GRED1 and GRED2, all participate in sporopollenin biosynthesis in gerbera ................................................................. 23
4.2.2 Gerbera TKPRs: functional conservation with minor differentiation .......... 24
4.2.3 Functional diversification of PKRs: an evolutionary mechanism underlying chemical diversity of plant polyketides ............................................ 26

4.3 Developing a highly efficient DFR expression system and advanced DFR assay for determination of DFR substrate preferences (III) ......................................................... 26
4.3.1 Developing a highly efficient DFR expression system ................................. 27
4.3.2 Refining the classical BuOH-HCl assay into a reliable and robust assay for determination of DFR substrate preferences ........................................ 27
4.3.3 Substrate profiles of different DFRs .......................................................... 28

5. CONCLUSION AND PROSPECTS ................................................................ 29
ACKNOWLEDGEMENTS ............................................................................... 30
REFERENCES ............................................................................................... 33
ABBREVIATIONS

ACOS: Fatty acyl-CoA synthetase
ACP: Acyl carrier protein
ANR: Anthocyanidin reductase
ANS: Anthocyanidin synthase
ASCL: Anther specific chalcone synthase
AT: Acyltransferase
CCR: Cinnamoyl-CoA reductase
CHI: Chalcone isomerase
CHS: Chalcone synthase
Cy: Cyanidin
CYC: Cyclase
CYP: Cytochrome P450-enzyme
DFR: dihydroflavonol 4-reductase
DH: Dehydratase
DHK: Dihydrokaempferol
DHM: Dihydromyricetin
DHQ: Dihydroquercetin
Dp: Delphinidin
ER: Enoyl reductase
FAS: fatty acid synthase
F3H: Flavanone 3-hydroxylase
F3’H: Flavonoid 3’-hydroxylase
F3’5’H: Flavonoid 3’5’-hydroxylase
GRED: Gerbera reductase
GT: Glucosyltransferase
HMC: 4-hydroxy-5-methylcoumarin
HPLC: High performance liquid chromatography

KS: Keto synthase
KR: Keto reductase
Mv: Malvidin
NMR: nuclear magnetic resonance
OAC: Olivetolic acid cyclase
PKR: Polyketide reductase
PKS: Polyketide synthase
PKSA: Polyketide synthase A
PKSB: Polyketide synthase B
2PS: 2-pyrene synthase 1
Pg: Pelargonidin
Pn: Peonidin
Pt: Petunidin
RNAi: RNA interference
SDR: Short-chain dehydrogenase/reductase
STS: Stilbene synthase
TAL: Triacetolactone
TAA-CoA: 3,5-dioxohexanoic-CoA
TE: Thioesterase
TLC: Thin-layer chromatography
TKPR: tetraketide α-reductase
TKS: tetraketide synthase
VIGS: Virus Induced Gene Silencing
LIST OF PUBLICATIONS

This PhD Thesis is based on the following two publications and one manuscript, which are listed following the leading roman numbers:


Author’s contributions

I. LZ designed the research together with MP, JK and THT. LZ performed the majority of experiments, including construct cloning, metabolic analysis and expression analysis of gerbera tissues, cultivars and transgenic lines, in vitro enzyme assays. LZ and THT analyzed the results and wrote the manuscript with inputs from all authors.

II. LZ designed the research together with THT. LZ performed gene identification and isolation, construct cloning, Arabidopsis complementation experiment, expression and phenotype analysis of transgenic gerbera lines and transgenic Arabidopsis lines, and phylogenetic analysis. SEM was conducted by TZ. LZ and THT analyzed the results and wrote the manuscript with inputs from all authors.

III. LZ was involved in designing the experiments with THT. LZ contributed to the expression system testing, DFR assay optimization and evaluation. LZ was involved in the data interpretation and manuscript completion.
1. INTRODUCTION

1.1 Plant polyketides: a vast group of structurally and functionally diverse natural products

During the long process of evolution, plants have acquired the ability to produce a vast array of specialized secondary metabolites that play crucial roles in helping plants successfully survive in the changing environment. Unlike primary metabolites that are common to all plants, plant secondary metabolites exhibit remarkable diversity in chemical structure and function and many of them often accumulate in a limited number of species (Afendi et al., 2012). Polyketides (Figure 1) constitute one of the most numerous and widely occurring groups of the chemically diverse secondary metabolites in the plant kingdom. They play important roles in plant life processes by exhibiting a wide range of biological and physiological functions, such as defense systems, pigments, and structural elements of the pollen wall (Stewart. et al., 2013; Lim et al., 2016). Moreover, many of them are important natural compounds presenting valuable pharmaceutical or industrial benefits to human beings (Bisht et al., 2021).

Figure 1. The diversity of plant polyketides. Typical chemical structures of three groups of plant polyketides are shown in individual frames. Their distributions are shown in bold.
1.1.1 Sporopollenin: the most widely distributed and functionally conserved plant polyketide

The most widely distributed polyketide derivative must be sporopollenin, a structurally robust polymer that functions as a structural element of pollen wall or spore wall of land plants (Jiang et al., 2013). In land plants, the formation of the sculpted pollen or spore wall is critical for protecting pollen grains and male gametophytes against various biotic and abiotic stresses, and therefore, in reproductive success (Wallace et al., 2011). Despite exhibiting a great diversity in surface morphology, land plant pollen walls share a conserved fundamental structure that consists of a pectocellulosic inner intine layer and a tough outer exine layer (Brooks & Shaw, 1978; Wallace et al., 2011). Sporopollenin is the major constituent of the exine layer, which has been considered to be the key component enabling the remarkable resistance of pollen walls to diverse stresses, due to its extreme chemical stability and physical strength (Brooks & Shaw, 1978; Wallace et al., 2011). A series of investigations by applying multiple degradation and nuclear magnetic resonance (NMR) methods have demonstrated that sporopollenin is a complex biopolymer consisting of highly cross-linked polyketide derived hydroxylated aliphatic, aromatic, and phenylpropanoid moieties (Ahlers et al., 1999; Li et al., 2019; Xue et al., 2020). Recent reverse genetic and biochemical studies have revealed that hydroxy tetraketide α-pyrone (Figure 1) synthesized by an anther-specific polyketide biosynthesis pathway are the major precursors of sporopollenin (Kim & Douglas, 2013).

1.1.2 Flavonoids: the widely distributed and structurally diverse plant polyketides

Flavonoids include over 9000 chemically diverse metabolites sharing a C₆-C₃-C₆ carbon framework (two aromatic rings connected by a heterocyclic ring) (Figure 1) and present the largest group of plant polyketides (Williams & Grayer, 2004). They are ubiquitously distributed in vascular plants and bryophytes (except hornworts) (Davies et al., 2020) and are widely involved in various plant life processes by functioning as pigments, copigments, defense systems, etc. (Panche et al., 2016). They are subdivided into flavones, isoflavones, flavonols, chalcones, and anthocyanins, based on the position of ring conjugation, degree of unsaturation, number of hydroxylation, and type of substitution (Panche et al., 2016) (Figure 1). A well-known group is anthocyanins, the major plant pigments ubiquitously distributed in the plant kingdom. They not only give flowers and fruits vivid colors to attract pollinators and seed spreaders but also have a role in protecting plants against various biotic and abiotic stresses (Landi et al., 2015; Smeriglio et al., 2016). Anthocyanins are also major health ingredients in vegetables, fruits, tea, wine, and chocolate due to their amazing antioxidant activity (Smeriglio et al., 2016).
Anthocyanins are derived from anthocyanidins by the addition of conjugated sugar and acyl moieties. Despite the diversity of flower and fruit colors, the common anthocyanins in pigmented plant tissues are derived from six prevalent anthocyanidins, pelargonidin (Pg), cyanidin (Cy), delphinidin (Dp), peonidin (Pn), petunidin (Pt), and malvidin (Mv) (Smeriglio et al., 2016; Liu et al., 2018). The latter three are the methylated forms of Cy and Dp (Smeriglio et al., 2016; Liu et al., 2018). Among them, Cy-, Dp- and Pg-derivatives are the most widely distributed anthocyanins that are present in 70% of pigmented plant tissues (Smeriglio et al., 2016). Pg, Cy and Dp are primary anthocyanidins that differ from each other by the degree of hydroxylation in the B-ring and show orange/scarlet, red/magenta, and violet/blue hues, respectively (Liu et al., 2018). Despite their wide occurrence, the content of the three primary anthocyanidin derivatives differ greatly in different pigmented tissues and are even missing in some species or some varieties of some species. For instance, roses, carnations, and gerbera lack Dp-derived anthocyanins, and petunia lacks Pg-derived anthocyanins (Tanaka & Ohmiya, 2008).

1.1.3 Taxa-specific polyketides

Unlike the widely distributed sporopollenin and flavonoids, various polyketides occur only in specific plant taxa or species and display diverse structures. They usually do not directly participate in growth or development, and mainly function as defense systems by exhibiting a range of bioactivities to protect plants against pathogens and herbivores. These bioactive polyketides provide plants of specific taxa unique survival advantages in challenging environments. To humans, many of them have been regarded as important medicinal resources or potential novel therapeutics, given their antimicrobial, immunosuppressive, and anticancer properties (Lim et al., 2016; Bisht et al., 2021). Typical examples are stilbenes from pine, grapevine and peanut (Chong et al., 2009), gerberin and parasorboside from gerbera (Koskela et al., 2011), cannabinoids from Cannabis sativa (Gagne et al., 2012), tropane alkaloids from Solanaceae and Erythroxylaceae (Kohnen-Johannsen & Kayser, 2019), and hypericin from Hypericum species (Karppinen et al., 2008) (Figure 1).

1.2 Polyketide biosynthesis and the logic of chemo-diversity

Despite the diversity in chemical structures and functions, the wide ranges of plant polyketides are assembled in a common chemical logic. Polyketide synthases (PKSs) synthesize the polyketide backbone by catalyzing decarboxylative condensation of acyl- and malonyl-CoA primary metabolites followed by intramolecular cyclization or a release as a linear product (Hertweck, 2009) (Figure 2A). The PKS synthesized core structures undergo accurate modifications catalyzed by a series of accessory tailoring enzymes to get the final structure-specific polyketide (Hertweck, 2009) (Figure 2).
1.2.1 Polyketide synthase (PKS) catalyzes the synthesis of polyketide backbone

PKSs are multifunctional enzymes derived from fatty acid synthase (FAS), and they are classified into type I, II and III according to their architectural configurations (Staunton & Weissman, 2001; Austin & Noel, 2003; Fischbach & Walsh, 2006; Neves et al., 2021). The type I PKSs are large multifunctional enzymes that contain a set of modularly or iteratively assembled polypeptides including keto synthase (KS), acyltransferase (AT), acyl carrier protein (ACP), thioesterase (TE), and optional tailoring domains such as ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) (Neves et al., 2021) (Figure 2B). The type II PKSs are multi-enzyme complexes composed of individual monofunctional enzymes that carry on a set of iterative catalyzations, which contain at least two KSs and an ACP, and optional tailoring subunits like KR and cyclase (CYC) (Fischbach & Walsh, 2006; Wang et al., 2020) (Figure 2B). Finally, the type III PKSs are simple homodimers that only contain a KS domain, which alone performs iterative condensations and cyclization within a single active site (Austin & Noel, 2003; Abe & Morita, 2010) (Figure 2B). Type I and type II PKSs are widely present in bacterial and fungal polyketide biosynthesis, while type III PKSs are mainly found in plants initiating plant polyketide biosynthesis (Yu et al., 2012).

The first plant PKS encoding gene, *chalcone synthase* (CHS), was isolated from *Petroselinum crispum* by Reimold et al. (1983) and uncovered to initiate flavonoid biosynthesis. Later, a large number of homologous enzymes were found ubiquitously present in the genomes of diverse plant lineages as well as in some bacteria and fungi, forming the type III PKS family (Shimizu et al., 2017). So far, over 30 functionally diverse plant type III PKSs have been experimentally characterized in different plant species (Shimizu et al., 2017). They have been shown to share high similarity in amino acid sequences, but to synthesize different backbones by using different starter units, catalyzing different lengths of elongation, or terminating with different types of cyclization (Abe & Morita, 2010; Shimizu et al., 2017) (Figure 2A). The catalytic mechanisms of type III PKSs have been elucidated by resolving the crystal structures and by mutagenesis analysis of several typical type III PKSs (Ferrer et al., 1999; Jez et al., 2000; Austin et al., 2004). They share a common αβαβα three-dimensional structure and maintain a conserved “Cys-His-Asn” catalytic triad and a CoA binding site (Abe & Morita, 2010; Stewart et al., 2013). Type III PKSs act as homodimers and each monomer contains a functionally independent active site that controls the selection of starters, number of extension reactions, and modes of intramolecular cyclization, resulting in diverse polyketide backbones (Abe & Morita, 2010; Stewart et al., 2013). Based on the cyclization types, plant PKSs can be divided into lactone-forming (L), aldol-cyclization (A), Claisen-cyclization (C), and no cyclization (X) types (Shimizu et al., 2017) (Figure 2A).
Figure 2. The mechanism of polyketide biosynthesis. (A) Typical plant polyketide biosynthesis scheme. “L-type”, “A-type”, “C-type” and “X-type” refer the lactone-type, aldol-type, Claisen-type and non-cyclization type PKSs. “R” refers to the acyl-group in the starter unit. (B) Structure and catalytic mechanisms of three types of PKSs. Functional domains are shown in circles and dashed circles refer optional domains.
1.2.2 Post-PKS tailoring enzymes determine the functional diversity

Selective manipulations of polyketo-thioester intermediates by the optional tailoring domains or modules in type I and II PKSs are one of the main strategies accounting for the diversity of fungal and bacterial polyketides (Hertweck, 2009; Neves et al., 2021). Unlike the type I or II PKSs that rely on these tailoring domains or modules to perform necessary processing before product release, the simple type III PKSs synthesize polyketide backbones and directly release highly oxidized intermediates (Abe & Morita, 2010; Shimizu et al., 2017). A series of accessory modifications for achieving the final structural and functional specificity is supposed to be performed by separate post-PKS tailoring enzymes (Shimizu et al., 2017; Morita et al., 2019). Compared with the well-elucidated plant PKSs, plant polyketide biosynthesis associated tailoring enzymes have been much less studied.

So far, only a few plant PKS pathways have been completely elucidated. The best-characterized plant PKS pathway is anthocyanin biosynthesis, owing to decades of investigations in diverse model plant species including Arabidopsis, maize, tomato, etc. (Tohge et al., 2017). In anthocyanin biosynthesis, the CHS synthesized naringenin chalcone undergoes further selective processing by accessory tailoring enzymes including chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3′-hydroxylase (F3′H), flavonoid 3′5′-hydroxylase (F3′5′H), dihydroflavonol 4-reductase (DFR), anthocyanidin synthase (ANS) and glucosyltransferases (GTs), resulting in the orange/scarlet Pg-glucosides, red/magenta Cy-glucosides, and violet/blue Pg-glucosides (Tohge et al., 2017) (Figure 3A). Another example is the ancient sporopollenin biosynthesis (Figure 3B). A series of enzymes involved in this ancient PKS pathway have been elucidated in Arabidopsis through a series of investigations of male-sterile mutants showing defective pollen exine (Morant et al., 2007; Tang et al., 2009; De Azevedo Souza et al., 2009; Dobritsa et al., 2009; Grienenberger et al., 2010; Kim et al., 2010). In this pathway, medium- to long-chain fatty acids are hydroxylated by cytochrome P450-enzymes 703A2 (CYP703A2) and CYP704B1 and further esterified by fatty acyl-CoA synthetase 5 (ACOS5) (Figure 3B) (Morant et al., 2007; De Azevedo Souza et al., 2009; Dobritsa et al., 2009). Polyketide synthase A (PKSA) and PKSB use the generated fatty acyl-CoA esters as starters to synthesize tetraketide α-pyrones Kim et al., 2010), which are reduced by accessory tetraketide α-pyrone reductase 1 (TKPR1) and TKPR2 into hydroxylated tetraketide α-pyrones, important sporopollenin precursors (Grienenberger et al., 2010) (Figure 3B).

In contrast to the widespread anthocyanin biosynthesis and sporopollenin biosynthesis, most taxa-specific polyketide biosynthetic pathways are restricted to occur in a few species and are rarely present in the intensively studied model plants like Arabidopsis and rice. These species usually lack comprehensive bioinformatics databases or well-established
transgenic systems or other molecular operation systems. These restrictions make the investigations of the biosynthetic mechanisms of many taxa-specific polyketides move on very slowly. As a result, post-PKS tailoring enzymes in most taxa-specific PKS pathways remain poorly known.

Figure 3. Three typical polyketide biosynthetic pathways in gerbera. (A) Anthocyanin biosynthesis. CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; F3’H, flavonoid 3’-hydroxylase; F3’5’H, flavonoid 3’,5’-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase; GT, glucosyltransferase; DHK, dihydrokaempferol; DHQ, dihydroquercetin; DHM, dihydromyricetin. Colored compounds are marked with their approximate corresponding colors. F3’5’H and the branch in the dashed frame are absent in gerbera. (B) Sporopollenin biosynthesis is catalyzed by a series of conserved enzymes that have been characterized in Arabidopsis. CYP: cytochrome P450-enzyme; ACOS5: fatty acyl-CoA synthetase 5; PKS: anther-specific polyketide synthase; TKPR, tetraketide α-pyrene reductase. (C) Originally proposed gerberin/parasorboside biosynthesis. G2PS1, gerbera 2-pyrene synthase 1; PKR, polyketide reductase; GT, glucosyltransferase; TAA-CoA, 3,5-dioxohexanoic-CoA; TAL, Triacetolactone; GA, gerberin aglycone; PA, parasorboside aglycone.
1.2.3 Polyketide reductase (PKRs): necessary tailoring enzymes in many plant PKS pathways

Reductase (e.g. KR, ER.) catalyzing carbonyl reductions are crucial manipulations in fungal and bacterial polyketide biosynthesis (Hertweck, 2009; Neves et al., 2021). Polyketide reductases (PKRs) also are an important group of tailoring enzymes required for plant polyketide biosynthesis, like DFR in anthocyanin biosynthesis (Figure 3B) and TKPR1 and TKPR2 in sporopollenin biosynthesis (Figure 3C). Specifically, PKRs have been proposed to be essential enzymes in many lactone-type and aldol-type PKSs initiated PKS pathways, like sporopollenin biosynthesis (Grienenberger et al., 2010), hypericin biosynthesis (Karppinen et al., 2008), plumbagin biosynthesis (Springob et al., 2007), gerberin/parasorboside biosynthesis (Eckermann et al., 1998). These PKSs are known to synthesize 2-pyrones in vitro, while in planta they cooperate with downstream tailoring enzymes leading to hydroxylated 2-pyrones (e.g. sporopollenin precursor) (Grienenberger et al., 2010), or 2-pyrene derivatives with reduced 2-pyrene ring (e.g. gerberin and parasorboside) (Eckermann et al., 1998), or polyketides even without pyrone structure (e.g. hypericin, plumbagin) (Karppinen et al., 2008; Springob et al., 2007). Besides the characterized TKPR1 and TKPR2 in sporopollenin biosynthesis, PKRs in these plant PKS pathways are still obscure.

1.3. Polyketide biosynthesis in *Gerbera hybrida*

1.3.1 *Gerbera hybrida*: Model plant for polyketide biosynthesis

The genus of *Gerbera*, a member from the tribe Mutisieae of the family Asteraceae, comprises around 40 species of perennial herbs natively distributed in Africa and Asia (http://www.theplantlist.org/1.1/browse/A/Compositae/Gerbera/) (Manning et al., 2016). Many of them are known as medicine plants that have been traditionally used to treat various diseases due to their abundant bioactive polyketide derivatives. For instance, *G. anandria, G. saxatilis, G. delavayi, G. piloselloides* are traditionally used for treating cough, fever, and as anti-inflammatories in China (Liu et al., 2010; Qiang et al., 2011; He et al., 2014; Li et al., 2020). The best-known plant from this genus is the ornamental plant gerbera (*Gerbera hybrida*), which is popular for bright colors and attractive capitulum and has become one of the most economically important cut-flower crops globally (Teeri et al., 2006a).

*Gerbera hybrida* originated from artificially hybridized progenies of two native South African species, *G. jamesonii* and *G. viridifolia*, thus giving rise to new color combinations and floral arrangements (Ambrosius, 2003). With decades of efforts by breeders, thousands of commercial gerbera varieties presenting impressive color patterns and diverse floral
morphologies have been developed and are popular worldwide. In addition to the attractive ornamental traits, *Gerbera hybrida* also inherits the ability to make polyketide derived 2-pyrones gerberin, parasorbose and gerberinside (Figure 4). These compounds provide gerbera tissues bitter taste and have been proven to have a role in defense against herbivores and pathogens (Koskela et al., 2010; Mascellani et al., 2022). They have previously been reported as the major bioactive compounds in many species from the same genus (Nagumo et al., 1989; Eckermann et al., 1998; Liu et al., 2010; He et al., 2014; Li et al., 2020). These compounds probably represent a major adaptive advantage of the genus of *Gerbera* with respect to microbial attack and insect herbivores.

The typical capitulum, diverse floral color patterns, and richness of important defense-related polyketides contribute to gerbera being a great plant material for research. Three decades of efforts in our laboratory have worked gerbera into a model plant of Asteraceae for flower development and secondary metabolism studies (Teeri et al., 2006a,b; Elomaa et al., 2018; Zhang et al., 2021). Extensive gerbera genome and transcriptome data and well-developed molecular tools of gerbera, including *Agrobacterium*-mediated stable transformation, virus-induced gene silencing (VIGS), heterologous protein expression, enzyme assays, etc., largely facilitate the investigation of polyketide biosynthesis in gerbera. Much progress in gerbera polyketide biosynthesis has been made, while many questions are still waiting for answering.

![Gerbera plant and its major polyketides](image)

**Figure 4. Gerbera plant and its major polyketides.** The approximate colors of the two major anthocyanins are shown in the corresponding frame backgrounds.
1.3.2 Gerberin and parasorboside biosynthesis: PKRs remain unknown

The first committed enzyme initiating gerberin and parasorboside biosynthesis is the lactone-forming PKS Gerbera 2-pyrone synthase 1 (G2PS1) (Eckermann et al., 1998). G2PS1 was first identified as a CHS-like gene during the investigation of gerbera flavonoid biosynthesis due to its high similarity with gerbera CHS (Helariutta et al., 1995a). However, G2PS1 was found expressed in many tissues where flavonoids do not accumulate, suggesting it is not a real CHS (Helariutta et al., 1995a). Absence of both gerberin and parasorboside in transgenic gerbera G2PS1 antisense lines suggest this enzyme might be involved in the biosynthesis of gerberin and parasorboside (Eckermann et al., 1998). Further in vitro enzymatic assays verified that G2PS1 uses acetyl-CoA as the starter and malonyl-CoA as the extender to synthesize triacetolactone (TAL, 4-hydroxy-6-methyl-2-pyrone) (Eckermann et al., 1998) (Figure 3C). G2PS1 is the first identified plant type III PKS using acetyl-CoA as the starter unit and synthesizing lactone (Eckermann et al., 1998). Its discovery suggested a new pathway in plant polyketide biosynthesis and shed light on the identification and characterization of many plant PKSs that followed, accounting for the biosynthesis of many important bioactive polyketides. Famous examples are the Hypericum perforatum PKS2 (HpPKS2) in hypericin biosynthesis (Karppinen et al., 2008), hexaketide synthase (HKS) in plumbagin biosynthesis (Springob et al., 2007), and gerbera G2PS2 and G2PS3, homologues of G2PS1, in gerberinside biosynthesis (Pietiäinen et al., 2016). Based on their chemical structures, one or two PKRs were proposed to reduce the pyrone ring of TAL to get the aglycone forms of gerberin and parasorboside (Eckermann et al., 1998) (Figure 3C). However, the PKRs have been unknown.

1.3.3 Sporopollenin biosynthesis: TKPRs – Functional conservation or diversification?

Sporopollenin biosynthesis has been well elucidated in Arabidopsis, which is constructed by a series of tapetum-specific enzymes that are tightly coregulated during pollen development (Figure 3B). Their orthologues have been found widespread in land plants ranging from mosses to angiosperms, with distinct male-organ specific expression patterns (Morant et al., 2007; Tang et al., 2009; De Azevedo Souza et al., 2009; Dobritsa et al., 2009; Grienenberger et al., 2010; Kim et al., 2010). Sporopollenin biosynthesis has been considered to be an ancient and conserved PKS pathway present in all land plant lineages and probably associated with the colonization success of land plants (De Azevedo Souza et al., 2009; Grienenberger et al., 2010; Kim et al., 2010). So far, orthologues of ACOS5, CYP703A, CYP704B1, and TKP R1 have been characterized in many model plant species, like rice, rapeseed, etc. and they were shown functionally conserved in sporopollenin biosynthesis (Wang et al., 2013; Qin et al., 2016; Lin et al., 2017; Zou et al., 2018; Shi et al., 2018; Xu et al., 2019). However, TKPR2 has been only characterized in Arabidopsis, and it was shown
to play a minor role in pollen wall formation in contrast to the predominant role of AtTKPR1 (Grienenberger et al., 2010).

In gerbera, the orthologues of Arabidopsis PKSA and PKSB, GASCL1 and GASCL2, have been shown to play conserved functions in sporopollenin biosynthesis (Kontturi et al., 2017). However, the remaining sporopollenin biosynthesis associated enzymes have still been waiting for characterization in gerbera. A single orthologue of AtTKPR1 and two orthologues of AtTKPR2 are present in the gerbera transcriptome, which implies that there might be functional diversification in gerbera TKPRs. This hypothesis is waiting for further verification.

1.3.4 Anthocyanin biosynthesis: molecular mechanism underlying substrate specificity of DFRs remains unknown

In anthocyanin-producing plant lineages, there must be at least one CHS, CHI, F3H, DFR, ANS, and GT enzyme constituting the core pathway, while the presence of F’3H and F3’5’H varies depending on species (Yonekura-Sakakibara et al., 2019) (Figure 3A). Lack of blue or purple flowers in some species like roses, carnations, gerbera, etc. has been shown to result from the absence of F3’5’H that is responsible for Dp-derived anthocyanins (Tanaka & Ohmiya, 2008). However, orange and scarlet antocyanins, the Pg-derivatives, also found absent in some species or specific cultivars of some species although they contain the core anthocyanin pathway responsible for the biosynthesis of DHK, the precursor of Pg. A well-known example is the absence of orange petunia in nature, which was shown to be a result of the disability in using DHK by petunia DFR A (Forkmann & Ruhnau, 1987). Scientists successfully made orange petunia by transferring the maize DFR A1 or the gerbera DFR into petunia (Meyer et al., 1987; Elomaa et al., 1995). The interesting substrate specificity of DFRs attracts attention. After decades of investigation, the question of which amino acids account for the specificity is still obscure.

Gerbera varieties present distinguished cultivar-specific pigmentation patterns, and they can be divided into Pg exclusive (orange to pink, e.g., ‘Regina’), Cy exclusive (red to fuchsia, e.g., ‘President’), both (deep red, e.g., ‘Passion’), and no anthocyanin (White or yellow, e.g., ‘Ivory’) types (Bashandy et al., 2015b). Three natural DFR alleles GDFR1-1, GDFR1-2, and GDFR1-3, and one nonfunctional allele GDFR1-2m were isolated from different gerbera cultivars, and they were shown to be responsible for the cultivar-specific pigmentation patterns of gerbera (Helariutta et al., 1995b; Bashandy et al., 2015b). These gerbera DFR alleles showed high similarity in amino acid sequence (> 98%) but presented significant differences in substrate specificity (Helariutta et al., 1995b; Bashandy et al., 2015b). However, which amino acids are responsible for their distinguish substrate specificity is still obscure.
In practice, two major technical challenges largely limit the progress of the investigation. First, heterologous expression of DFRs can be difficult due to the excessive formation of insoluble inclusion bodies in microbes, and in low yields when expressed in plants. Although attempts by fusion with soluble carrier proteins have successfully made DFRs in microbes (Martens et al., 2002; Gosch et al., 2014; Miosic et al., 2014), the processes are time-consuming and expensive. Lack of a robust and reliable assay for the determination of DFR activity to different substrates is the other challenge. The classical method for the determination of DFR activity is the BuOH-HCl assay-dependent method. This method converts the DFR products, leucoanthocyanidins, into colored anthocyanidins in HCl, and then estimates the DFR activity by quantification of the colored anthocyanidins by applying a spectrophotometer (Porter et al., 1985). Although this method is available to compare DFR activities to single substrates, it is difficult to determine DFR activities for different substrates. Differing in the extractability in ethyl acetate of leucoanthocyanidins from DHK, DHM and DHQ makes it not accurate to determine the consumption of each substrate. Overlapping absorption spectrums by butanoylated-Pg, -Cy and -Dp compounds make it difficult to quantify the generated products. A robust and reliable assay is required for the investigation of the molecular mechanism underlying DFR substrate specificity.
2. AIMS OF THE STUDY

This work used gerbera as a model to investigate the post-PKS tailoring steps in plant polyketide biosynthesis. We aimed to investigate the logic of plant polyketide diversity through identification and characterization of important PKRs in three typical gerbera PKS pathways, defense-related gerberin/parasorboside biosynthesis, sporopollenin biosynthesis, and anthocyanin biosynthesis.

The specific objectives are:

I. To identify and characterize the unknown PKRs in gerberin/parasorboside biosynthesis in gerbera

II. To investigate the functional conservation and diversification of TKPRs in the ancient sporopollenin biosynthesis

III. To establish an efficient heterologous expression system for plant reductases and a robust enzyme assay for investigation of substrate specificity of gerbera DFRs
3. MATERIALS AND METHODS

Materials and methods used in this study are described in detail in publications (I- III) and are summarized in Table 1.

Table 1. Materials and methods used in publications.

<table>
<thead>
<tr>
<th>Materials</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gerbera transcriptome data</td>
<td>I, II</td>
</tr>
<tr>
<td>Gerbera hybrida</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>II</td>
</tr>
<tr>
<td>Nicotiana benthamiana</td>
<td>I, III</td>
</tr>
<tr>
<td>Petunia hybrida</td>
<td>III</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methods</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification and cloning of genes</td>
<td>I, II, III</td>
</tr>
<tr>
<td>RNAi binary vector construction and gerbera transformation</td>
<td>I, II</td>
</tr>
<tr>
<td>Arabidopsis complementation experiment</td>
<td>II</td>
</tr>
<tr>
<td>Heterologous protein expression vector construction and Agro-infiltration</td>
<td>I, III</td>
</tr>
<tr>
<td>Quantitative real-time PCR</td>
<td>I, II</td>
</tr>
<tr>
<td>Semiquantitative RT-PCR</td>
<td>II</td>
</tr>
<tr>
<td>RNA in situ hybridization</td>
<td>II</td>
</tr>
<tr>
<td>Thin layer chromatography</td>
<td>I</td>
</tr>
<tr>
<td>Protein extraction and purification</td>
<td>I, II</td>
</tr>
<tr>
<td>Enzyme assay</td>
<td>I, III</td>
</tr>
<tr>
<td>BuOH-HCl assay</td>
<td>III</td>
</tr>
<tr>
<td>HPLC</td>
<td>III</td>
</tr>
<tr>
<td>Pollen grain viability assay</td>
<td>II</td>
</tr>
<tr>
<td>Scanning electron microscopy</td>
<td>II</td>
</tr>
<tr>
<td>Phylogenetic analysis</td>
<td>II</td>
</tr>
</tbody>
</table>
4. RESULTS AND DISCUSSION

4.1 PKRs in gerberin/parasorboside biosynthesis in *Gerbera hybrida* (I)

4.1.1 Identification of two PKR candidates in gerberin/parasorboside biosynthesis

Genes encoding enzymes of the same metabolic pathway are often co-regulated in space and time. *G2PS1* was previously identified and characterized as the first committed enzyme encoding gene in gerberin/parasorboside biosynthesis (Helariutta *et al.*, 1996; Eckermann *et al.*, 1998). To explore the missing accessory tailoring enzymes in this pathway, we conducted Pearson correlation analysis between the *G2PS1* and transcriptome data that was printed on a microarray chip (Laitinen *et al.*, 2005). A gene encoding short-chain dehydrogenase/reductase (SDR) showing high coexpression with *G2PS1* ($r^2 = 0.74$) was identified as a potential PKR in this pathway. We named it *Gerbera reductase 1* (*GRED1*). We also identified another SDR gene that shows high sequence similarity with *GRED1* and named it *GRED2*. GRED2 was the only gerbera enzyme showing marked similarity with GRED1, and it therefore was considered as another PKR candidate in gerberin/parasorboside biosynthesis.

4.1.2 GRED1 and GRED2 are the PKRs necessary for parasorboside biosynthesis

To investigate whether GRED1 and GRED2 are associated with gerberin/parasorboside biosynthesis, we conducted expression and metabolite analysis of diverse gerbera tissues and cultivars. We showed that gerberin and parasorboside are richly accumulated in nearly all gerbera tissues however much less in anthers and roots where *G2PS1* and GRED1 were least expressed (Figure 2 in I). The relative distribution of gerberin and parasorboside varied between tissues. Particularly, the high GRED2 expression tissues preferentially synthesize parasorboside while low GRED2 expression tissues tend to accumulate more gerberin (Figure 2 in I). Similar observations also took place in the gerbera cultivar series. Compared with ‘Regina’, the high GRED1 or GRED2 expressing cultivars ‘President’, ‘Pingpong’ and ‘Grizzly’ mainly accumulated parasorboside and produced less gerberin (Figure 2 in I). These observations support a hypothesis that GRED1 and GRED2 might be the second PKR in gerberin and parasorboside biosynthesis.

To directly investigate the functions of GRED1 and GRED2 in gerberin/parasorboside biosynthesis, we made transgenic gerbera lines with overexpression of *GRED1* (*GRED2* overexpression lines failed), and transgenic gerbera lines with suppression of both *GRED1* and *GRED2* (Figure 3,4 in I). We showed that the *GRED1* overexpression lines accumulated an increasing amount of parasorboside at the expense of gerberin, while the *GRED1* and
**GRED2** downregulated lines produced dramatically decreased amounts of parasorboside and correspondingly dramatically increases of gerberin (Figure 3, 4 in I). These data showed that both GRED1 and GRED2 correlate with the accumulation of parasorboside, suggesting that GRED1 and GRED2 catalyze the second reduction step necessary for the biosynthesis to proceed towards parasorboside. The overexpression or suppression of this reductase activity did not result in changes in the total content of gerberin and parasorboside (Figure 3, 4 in I).

4.1.3 PKRs act on the linear triketide prior to lactonization, suggesting an update on the pathway leading to gerberin and parasorboside

The previously proposed gerberin/parasorboside biosynthetic pathway suggests that one or two PKRs are supposed to reduce the 2-pyrene ring of TAL accounting for the formation of gerberin and parasorboside (Eckermann et al., 1998) (Figure 3C). The proposed intermediates TAL, gerberin aglycone, or gerberin are potential substrates of the PKRs (Figure 3C). To functionally characterize GRED1 and GRED2, we conducted *in vitro* enzymatic assays to test these potential substrates. However, none of the TAL, gerberin aglycone, or gerberin could be reduced by GRED1 or GRED2 in our assays (Figure 5 in I). This indicates that our previously proposed gerberin/parasorboside biosynthetic pathway needs an update. In radio-labeling experiments, both GRED1 and GRED2 were shown to make a new labeled product at expense of TAL in a reaction together with G2PS1 and [2-14C] malonyl-CoA (Figure 5 in I). These observations suggested that GRED1 and GRED2 are indeed involved in gerberin/parasorboside biosynthesis, but they are not acting with the 2-pyrene TAL or gerberin or gerberin aglycone. Instead, they might act on a linear polyketide intermediate prior to its lactonization. The reduced diketides (crotonyl-CoA or 3-hydroxybutyryl-CoA) can not be extended by G2PS1 (Eckermann et al., 1998), suggesting that the linear diketide acetoacetyl-CoA is not the substrate for the PKRs. Then the PKRs most likely act on the remaining linear triketide intermediate, 3,5-dioxohexanoic-CoA (TAA-CoA). The most possible pathway leading to gerberin/parasorboside biosynthesis is: G2PS1 extends the generation of triketide intermediate TAA-CoA, which is reduced by an unknown PKR at distal keto group to produce 5-hydroxy-3-oxohexanoic-CoA and subsequently reduced by the second PKR (GRED1 and GRED2) at the proximal group to generate the 3,5-dihydroxyhexanoic-CoA; the reduced linear intermediates are subsequently lactonized to gerberin aglycone and parasorboside aglycone (Figure 5B). TAL is much more likely a derailment product in gerberin/parasorboside biosynthesis than acting as the intermediate.

This updated pathway is strongly supported by a recently identified linear gerbera polyketide, 5-hydroxyhexanoic acid 3-O-β-D-glucoside (Mascellani et al., 2021), which shares the same backbone with the reduced intermediate 3,5-dihydroxyhexanoic-CoA in our updated
pathway. This supports the idea that PKRs indeed act on the linear triketide prior to lactonization. The updated pathway makes our previous observations that TAL or its glucoside has never been detected in any gerbera tissues become explicable.

The updated pathway suggests that plant polyketide biosynthesis share similar tailoring strategies with fungal and bacterial polyketide biosynthesis. In type I and II PKS systems, tailoring peptides act on linear polyketide intermediates prior to product release and are required for the region-specific cyclization of the growing chain (Fischbach & Walsh, 2006). The lactonization of linear polyketide thioesters into lactones is a spontaneous reaction (Light et al., 1966; Springob et al., 2007), which explains why lactones have been commonly produced as a derailment product in many type I and II PKS pathways when tailoring enzyme functions were absent (Campuzano & Shoolingin-Jordan, 1998; Hadfield et al., 2004). Similarly, many plant PKSs have been shown to produce lactones in vitro instead of making the expected non-pyrene intermediates in the original plants, like Plumbago HKS (Springob et al., 2007) and Cannabis tetraketide synthase (TKS) (Gagne et al., 2012). This suggests that tailoring enzymes are necessary for stabilizing the growing intermediates to get rid of unspecific cyclization in these type III PKS initiated pathways, similar to the tailoring polypeptides in type I and II PKSs.

4.2 PKRs in sporopollenin biosynthesis in Gerbera hybrida (I, II)

The identified GRED1 and GRED2 were shown to be two close orthologues of Arabidopsis TKPR2 (AtTKPR2). AtTKPR1 and AtTKPR2 are two PKRs involved in the ancient anther-specific sporopollenin biosynthetic pathway. We showed that GRED1 and GRED2 are not anther-specific and are involved in defense-related gerberin/parasorboside biosynthesis in gerbera. Several questions remained to be answered: Do the GRED1 and GRED2 still participate in sporopollenin biosynthesis in gerbera? Is there an orthologue of AtTKPR1 in gerbera and does it participate in sporopollenin biosynthesis?

4.2.1 Three PKRs, GTKPR1, GRED1 and GRED2, all participate in sporopollenin biosynthesis in gerbera

By conducting BLAST search using the AtTKPR1 sequence as the query, we identified a single AtTKPR1 homologue from the gerbera transcriptome data and named GTKPR1. GTKPR1 exhibits an anther-specific expression pattern in the transcriptome data, which was verified by semi-quantitative RT-PCR (Figure 3 in II). Further RNA in situ hybridization experiments showed that GTKPR1 presents the highly consistent precise expression pattern observed also for AtTKPR1 in developing anthers. Both of these genes are exclusively expressed in the tapetum, starting from the meiosis stage when the tapetum is first visible
until the microspore stage when the tapetum degenerates, showing strongest expression at the microspore stage (Figure 3 in II). The highly conserved tapetum-specific expression pattern suggests that GTKPR1 may play similar functions as AtTKPR1. To investigate its function in sporopollenin biosynthesis, we made transgenic gerbera lines with gene-specific suppression of GTKPR1. Similar male-sterile phenotypes as previously reported in Arabidopsis tkpr1 null mutants and partial mutants were observed in these GTKPR1 downregulated lines. The severely and moderately downregulated lines showed sterile anthers, while the mildly downregulated lines grew normal anthers but failed to release pollen grains normally (Figure S1 in II). By applying pollen vitality assay and SEM analysis, we showed that the severely downregulated lines produce completely aborted pollen grains with distorted pollen wall, and over 60% of pollen grains produced by the moderately and milder downregulated lines were dead and showed a degree of exine damage (Figure 4 in II). These observations support that GTKPR1 is involved in sporopollenin biosynthesis (Figure 5A), and its function is very crucial in sustaining the fertility of plants.

Although GRED1 and GRED2 are non-anther-specific orthologues of AtTKPR2, in the anther both of them are specifically expressed in the tapetum at the microspore stage (Figure 3 in II), a consistent expression pattern with AtTKPR2. No other TKPR2-like gene was found in the gerbera transcriptome data. These observations indicate that GRED1 and GRED2 may still work in sporopollenin biosynthesis. To investigate their functions in pollen wall formation, we analyzed the phenotypes of the pollen grains produced by the gerbera GRED1 and GRED2 RNAi downregulated lines. Although these downregulated lines grew normal anthers, we showed that the GRED1/GRED2 cross-downregulated lines produced 10% to 30% dead pollen grains that showed similar mild exine defects as previously reported in Arabidopsis tkpr2 mutants (Figure 4 in II). This suggests that GRED1 and GRED2 are still involved in sporopollenin biosynthesis despite that they play crucial roles in defense-related gerberin/parasorboside biosynthesis (Figure 5). The role of GRED1 and GRED2 are minor, in contrast to GTKPR1’s indispensable role in gerbera pollen wall formation, similar to the situation in Arabidopsis.

4.2.2 Gerbera TKPRs: functional conservation with minor differentiation

Consistent with the wide occurrence of sporopollenin, extensive evidence has demonstrated that sporopollenin biosynthesis is a highly evolutionally conserved metabolic pathway in land plants (De Azevedo Souza et al., 2009; Dobritsa et al., 2009; Grienenberger et al., 2010; Kim et al., 2010). Compared with that most sporopollenin biosynthesis associated enzymes have been characterized in diverse species playing conserved functions in sporopollenin biosynthesis, TKPR2 orthologues have not been characterized before in other species except in Arabidopsis.
In this work, we showed that gerbera has a single anther-specific AtTKPR1 orthologue GTKPR1 and two non-anther-specific AtTKPR2 orthologues GRED1 and GRED2. Expression of GTKPR1 in Arabidopsis tkpr1 mutants driven by the native promoter of AtTKPR1 completely restored the exine defects and fertility (Figure 5 in II), suggesting that GTKPR1 is functionally conserved. Despite that GRED1 and GRED2 were shown to catalyze a crucial step in the defense-related parasorboside biosynthesis in gerbera, we showed that they still play minor TKPR functions in gerbera sporopollenin biosynthesis. This suggests that the duplicate TKPR2s have functionally diversified in gerbera. Specifically, GRED1 and GRED2 have been recruited from the ancient sporopollenin biosynthesis to defense-related polyketide biosynthesis. We also showed that expression of AtTKPR2, GRED1 or GRED2 driven by AtTKPR1 native promoter failed to restore the fertility of Arabidopsis tkpr1 mutant (Figure 5 in II). These observations indicate that TKPR1 and TKPR2 are functionally differentiated even in Arabidopsis, although they have been proved to catalyze similar enzymatic functions in vitro (Grienenberger et al., 2010).

Figure 5. Updated sporopollenin and gerberin/parasorboside biosynthetic pathways in gerbera. (A) GTKPR1, GRED1 and GRED2 are involved in sporopollenin biosynthesis in the tapetum of gerbera anthers. (B) GRED1 and GRED2 catalyze the second reduction on the pathway to parasorboside in gerbera. GASCL, gerbera anther specific chalcone synthase; GTKPR1, gerbera tetraketide α-pyrone reductase 1; GRED, gerbera reductase; G2PS1, gerbera 2-pyrone synthase 1; PKR, unknown reductase; GA, gerberin aglycone; PA, parasorboside aglycone.
4.2.3 Functional diversification of PKRs: an evolutionary mechanism underlying chemical diversity of plant polyketides

Plants are completely sterile when losing functions of TKPR1 but still fertile without expression of TKPR2 (Grienenberger et al., 2010; Wang et al., 2013). TKPR1 is apparently much more important than TKPR2 in sporopollenin biosynthesis. By searching the latest transcriptome data of several model plants, we found that TKPR1s always present in a single copy expressed in the male reproductive organ, while TKPR2s sometimes appear as duplicate enzymes with broader expression patterns, like in soybean and polar (Figure 2 in II). This suggests that the predominant TKPR1 is functionally conserved in land plant evolution, while the minor TKPR2 could have differentiated novel functions through gene duplication.

Plants have adapted to changing environments by producing a remarkable diversity of secondary metabolites. The diversification of plant secondary metabolism is believed to be a result of gene duplication events followed by functionally differentiation events (Moghe & Last, 2015). The large plant type III PKS superfamily enzymes have been demonstrated to be derived from FAS and acquired functional diversity through gene duplication and mutation events (Jenke-Kodama et al., 2008; Xie et al., 2016; Yonekura-Sakakibara et al., 2019). TKPR1s and TKPR2s are two subfamilies of the highly variable SDR108E subfamily of the plant SDR superfamily, which harbors many characterized plant secondary metabolite biosynthesis associated reductases like the DFR and ANR in flavonoid biosynthesis, and cinnamoyl-CoA reductase (CCR) in lignin biosynthesis (Moummou et al., 2012). Apparently, similar to plant PKSs, functionally diversification of PKRs through duplication and mutation is another evolutionary mechanism underlying the diversity of plant polyketides. The PKSA/B clade, PKS in sporopollenin biosynthesis, is one of the two early clades of plant PKSs presented before the divergence of Bryophytes and Tracheophytes (Naake et al., 2021). Consistently, TKPR2 appeared later than TKPR1, but earlier than DFR, ANR, CCR, etc. in evolution history (Moummou et al., 2012). This supports the possibility that TKPR2 has been recruited from the ancient sporopollenin biosynthesis to defense-related PKS pathways not only in gerbera but maybe also in other plants. Recruitment of conserved genes to perform new functions within secondary metabolism largely contributes to the environmental adaption of plants.

4.3 Developing a highly efficient DFR expression system and advanced DFR assay for determination of DFR substrate preferences (III)

Substrate specificities of DFRs have been shown to be responsible for the cultivar-specific pigmentation of gerbera as well as the absence of natural orange petunia. However, the
underlying molecular mechanism is still uncovered. In this work, to solve the two major technical challenges in the investigation of DFR substrate specificity, we developed a highly efficient protein expression system and refined the classical BuOH-HCl assay.

4.3.1 Developing a highly efficient DFR expression system

In this work, we used the hypertranslatable vector pEAQ-HT (Sainsbury et al., 2009) to express three gerbera DFR alleles (GDFR1-1, GDFR1-2 and GDFR1-3) in N. benthamiana via the Agro-infiltration method (Bashandy et al., 2015a), and harvested remarkably high yields of heterologous His-tagged DFRs from the infiltrated tobacco leaves (Figure 9.1 in III). The produced DFRs can be easily purified from the crude extracts of the infiltrated tobacco leaves using the general His-tagged protein purification method (Figure 9.1 in III). More remarkably, we showed that even the unpurified crude protein extracts showed high DFR activities and sustained consistent substrate preferences as the purified enzymes (Figure 9.2-9.3 in III). This newly developed DFR expression system is simple and highly efficient, provides powerful tools for the investigation of DFRs’ substrate specificity. It is also a great option for the preparation of all kinds of proteins for enzymatic studies or other work. Using this system, we also have successfully made GRED1 and GRED2, two ‘trouble’ proteins that are difficult to make using the E.coli expression system, for in vitro enzyme assays (see methods in I). This largely promoted the elucidation of the reduction steps of parasorboside biosynthesis.

4.3.2 Refining the classical BuOH-HCl assay into a reliable and robust assay for determination of DFR substrate preferences

The traditional colorimetric BuOH-HCl assay is widely applied to assay of DFR activities to monomeric leucoanthocyanidins. However, with this method it is hard to accurate quantify the consumption of different substrates or the amounts of generated leucoanthocyanidins. This makes it not reliable for determination of DFR substrate preferences. Using crude extracts of three gerbera DFRs prepared from the newly highly efficient system, we revised the traditional BuOH-HCl assay into a reliable and robust assay targeting to determination of DFR substrate preferences.

We first established a standard DFR assay by optimization of buffer, pH, reaction time and dihydroflavonol concentration (Figure 9.2-9.3 in III). In order to make the BuOH-HCl become reliable for assay of different leucoanthocyanidins, we conducted a series of extraction of stopped standard assays and calculated the regression relationships of extraction efficiency against extraction steps for each leucoanthocyanidin (Figure 9.2 in III). The calculated specific correction factors for two extraction steps of each generated leucoanthocyanidin are available to compensate for the variance in extractability. Moreover, we quantified the
consumption of substrates by determining the remaining substrates with HPLC. By plotting the consumed substrates against the absorbance, we got the molar absorptivities of the butanoylated anthocyanidins from the slopes of the generated linear regression lines (Figure 9.4 in III). The absorbance values therefore were converted to consumed substrates in µM (Figure 9.4 in III). We also evaluated the regression relationships of HPLC measured sum peak areas and absorbance values of the BuOH-HCl assay corresponding to each butanoylated product, for the simplified assay provided with all three substrates simultaneously (Figure 9.5 in III). This allows commensuration of simplified mixed substrate assay by HPLC to single substrate assay by BuOH-HCl.

4.3.3 Substrate profiles of different DFRs

Distinguished substrate specificity of gerbera GDFR1-1, GDFR1-2 and GDFR1-3, petunia DFR A and maize DFR A1 are responsible for the specific color patterns of gerbera cultivar ‘Regina’ (GDFR1-1/GDFR1-2) and ‘President’ (GDFR1-3/GDFR1-3), non-transgenic petunia cultivar V26, and orange engineered petunia cultivar ‘African Sunset’. In this work, we applied the refined BuOH-HCl assay to determine the substrate consumption patterns of the five DFRs prepared from the newly developed high efficient system and petal extracts from the above gerbera and petunia cultivars. We showed that the substrate preference patterns of DFRs in petal extracts were consistent with the heterologously produced enzymes in single substrate assays. GDFR1-1, GDFR1-2 and ‘Regina’ petal extracts showed a strong preference for DHK and low activities with DHQ and DHM. GDFR1-3 and ‘President’ petal extracts showed similar high activities to all the three substrates. Petunia DFRA and V26 petal extracts seldom used DHK and preferred to use DHM than DHQ. Maize A1 shows the highest activities to DHK and relatively lower activities with DHQ and DHM, while the petal extracts of the ‘African Sunset’, transgenic orange petunia proved to be transferred with Maize A1, showed a combination of profiles of the maize A1 and petunia DFRA. However, the enzymes’ activities of using DHK were largely inhibited when the three substrates were provided simultaneously.
5. CONCLUSION AND PROSPECTS

In this thesis, we investigated the PKR catalyzed tailoring steps in three polyketide biosynthetic pathways in gerbera.

Firstly, we identified and characterized two PKRs, GRED1 and GRED2, in the defense-related gerberin/parasorboside biosynthesis in gerbera. We showed that PKRs act on linear triketide intermediate prior to lactonization, suggesting an updated pathway to the formation of gerberin and parasorboside. We showed that type I, II and III PKS initiated pathways share similar tailoring strategies. This investigation not only elucidates a crucial tailoring step in the important gerbera PKS pathway but also provides information for the investigations of unsolved tailoring steps in some other plant PKS pathways.

Secondly, we showed that GRED1 and GRED2 are the minor TKPRs and GTKPR1 is the predominant TKPR in gerbera sporopollenin biosynthesis. GRED1 and GRED2 are recruited from the ancient sporopollenin biosynthesis to defense-related parasorboside biosynthesis in gerbera. This is a typical example of how plants achieve secondary metabolism innovation through recruitment of conserved genes to perform new functions to adapt to the changing environments during evolution. We also showed the predominant TKPR1 is functionally conserved in land plant evolution, which makes TKPR1s excellent targets for engineering male-sterile plants in breeding.

Moreover, we established a highly efficient DFR expression system and refined the traditional BuOH-HCl assay into a reliable and robust assay for the determination of DFR substrate preferences. These powerful tools successfully overcame the two major technical challenges that hinder the investigation of DFR substrate specificities. The newly developed system has been being applied to investigate which amino acids are responsible for the substrate specificity of gerbera DFRs (ongoing project in our group). This simple and highly efficient protein expression system is also a great tool for the preparation of other proteins in enzymatic studies.

Taken together, our data not only benefits elucidation of the biosynthetic mechanisms of important gerbera polyketides but also provides useful information for understanding the evolutionary mechanism underlying plant secondary metabolism diversification.
ACKNOWLEDGEMENTS

When it goes to this part, I start to get the feeling that I am indeed stepping toward the end of my PhD journey! Eventually, after six and a half years of struggles, I made it, growing up into a real PhD researcher from a PhD candidate!

I started the journey in an impressive golden autumn of 2015. I still remember the moment I entered Finland, I felt like stepping into a world of a fairy tale. Here I would like to thank the China Scholar Council (CSC) for providing financial support for my PhD journey. With the opening scene of a fairy tale, the journey has been proved to be an adventurous fairy tale. There were joyful moments when I stepped into the world of polyketides and started the investigations of the lovely gerbera. There were frustrating moments when I was trapped by these complicated and mysterious polyketide pathways experiencing endless failures after multiple explorations. There were exciting moments when I suddenly got inspiration or when the project moved forward for a few steps after long times efforts. After a dark period of panic, depression, and self-doubting, I eventually overcame my inner fears and picked up some of the lost confidence. I would say these impressive experiences will be invaluable treasures in my whole life, despite I am still quite far away from being a qualified scientist.

I would like to express my sincere appreciation to all the people who helped, encouraged, and motivated me throughout the journey of adventure.

My deepest appreciation is to my supervisor Prof. Teemu Teeri. Thanks for your patient guidance, inspiring discussions, and endless encouragement throughout my PhD studies. You were always available to help and made constant efforts to address all kinds of challenges in the PKR project together with me. I got numerous courage and motivation by working with such an enthusiastic, creative, and persistent scientist like you. There are so many impressive memories that will influence my future career forever. You always inspired me to keep thinking in the process of exploration: thinking what the behind mechanisms are before using an instrument or starting an experiment rather than just following or copying; thinking about why and how instead of just accepting the conclusions when reading an article. This is probably why you are such a super problem-solver who could help so many people solving all kinds of technical problems all the time. Having dedicated yourself to research for around 40 years, you still show great curiosity about the world and keep exploring new scientific problems and expanding new knowledge all the time. There is no wonder everyone has been shocked by your profound knowledge of numerous fields. Moreover, I was shocked by your amazing information organization systems. How could you find out a pieces of note you wrote 20 years ago in a minute! The great system of scientifically labeling the samples, bacteria strains and the plasmids in a combination of detailed electronic notes makes many work of our lab move forward so efficiently. You are a person who is serious and critical
about scientific issues while easy and kind in life. Thanks for organizing the amazing ‘Crayfish party’. It has become a festival of the Gerbera Laboratory that is the most exciting and impressive moment in our minds.

I would like to express my sincere gratitude to the family of Gerbera Laboratory. I have been privileged to work with so many great people in such a creative, cohesive, and warm group. Prof. Paula Elomaa, thanks a lot for providing constructive advice in properly planning the three parts of the work of my thesis when I was struggling with the first manuscript. I cannot imagine when I could finish my PhD without your precious comments. You are such a creative and productive scientist who always makes careful arrangements of work to ensure everything proceeds smoothly. Led by you and the other great leader Teemu, it is undoubtedly why our Gerbera Laboratory has been in flourish for over 30 years. Teng Zhang, the rising star in flower development, I am so lucky to work with you and get numerous supports from you. We registered as PhD students on the same day, while you were much more like a mentor in my PhD journey. Thanks for providing so many creative and inspiring discussions during my PhD journey, as well as for kindly sharing all kinds of precious experiences ranging from useful tips in molecular experiments, how to take high-quality photos to how to efficiently write a manuscript. During my frustrating moments, you never rejected my negative complaints. You always helped me to calm down and provided constructive advice targeting improving the situation. I couldn’t have proceeded with the work so smoothly in the last two years if leaves out of your help. Kean-Jin Lim, Feng Wang and Yafei Zhao, you are like brothers and sister in my PhD journey. Thanks for providing all kinds of support in research as well as in daily life, and for being tolerant of my complaints as well. Anu Rokkanen, Eija Takala and Marja Huovila are thanked for your excellent technical assistance, and Sanna Peltola for your patiently and carefully taking care of plant materials in the Greenhouse. Your works are indispensable parts of my thesis. I would like to thank Dr. Xianbao Deng, a previous PhD graduated from Gerbera Laboratory, for introducing me to join this warm family. Tanja Paasela, Juha Kontturi, Minhazur Rahman and Steven Taniwan, you are important parts of the pleased family. I am so grateful to work with such a group of nice people. Guangxun Fan and Quan Zhou, you are younger than me while never hesitated to express your support all the time. I do appreciate it.

Thanks must go to my follow-up group members, principal scientist Heiko Rischer and senior scientist Milla Pietiläinen from VTT Technical Research Centre of Finland. Thanks a lot for your friendly guidance and advice concerning my PhD studies every year. I would like to thank the pre-examiners, senior scientist Suvi Häkkinen from VTT Technical Research Centre of Finland and associate professor Katja Karppinen from UiT, The Arctic University of Norway. Thanks for carefully reviewing my thesis and providing nice and constructive comments. Doctoral School in Environmental, Food and Biological Sciences (YEB) is thanked for providing financial grants for the completion of my thesis. Doctoral Programme
in Plant Sciences (DPPS) is thanked for organizing multiple meaningful seminars and scientific communication activities, which provided us with great opportunities for getting access to the latest advances in plant science. Nina Sipari from the Viikki Metabolomics Unit is thanked for MS analysis.

Outside Finland, special thanks must be expressed to my friends Rufang Wang and Ran Lv. I am so lucky to have you two good friends. We started our PhD studies at the same moment in different countries. The distance never blocked our friendships. We traveled around Europe together during the holidays and enjoyed many joyful moments together. We keep in touch with each other now and then via social media, sharing each other’s happiness and sadness in PhD studies and life, encouraging and supporting each other all the time. Although you have finished your degrees one by one and started your new careers in China, I have never been forgotten. No matter when and what I met, I could always hear your responses, your endless encouragement, and useful comments. Your constant supports are indispensable power in supporting me in finishing the journey. Associate Prof. Yanling Liu and associate Prof. Mei Yang from the Chinese Academy of Sciences, the supervisors of my master thesis, you are thanked for providing helpful comforts and useful advice in adjusting my mentality when I was short of confidence and motivation, which means a lot in my journey.

In the end, I must extend my heartfelt thanks to my parents. You know little about plant sciences, but always believe I could do it well. Your constant understanding and support in these years are invaluable.
REFERENCES


Dobritsa AA, Shrestha J, Morant M, Pinot F, Matsuno M, Swanson R, Möller BL, Preuss D.


