Molecular and biochemical studies of fragrance biosynthesis in rose

Sun, P.

Citation for published version (APA):
Chapter 1

General introduction

Second part of this chapter has been published in:
Past and present of rose fragrance

Roses have attracted people’s attentions since ancient times. The earliest cultivation of roses can be found in China, western Asia and northern Africa and the history of cultivation can be traced back to 5000 years ago (Gudin, 2003). However, the development of cultivation was slow due to the limitation of local rose varieties and breeding methods. In the 18th century, the Chinese roses, such as *Rosa x odorata*, *R. rugosa* and *R. chinensis* cv. ‘Old Blush’ (also called Parsons’ Pink China), were introduced into Europe and brought the trait of remontancy to the European roses (Wang, 2003). Thanks to this introduction, starting from late 19th century, extensive hybridization among Chinese, European as well as Middle-Eastern roses resulted in thousands of modern rose cultivars, referred to as *Rosa x hybrida* (Channelière et al., 2002, Cherri-Martin et al., 2007, Gudin, 1999, Marriott, 2003).

There are approximately 150 naturally occurring rose species but only a few of them were selected for rose hybridization (Wang, 2003, Zlesak, 2006). Rose cultivars were selected mainly based on their floral shape, petal numbers, colour, recurrence of flowering, disease resistance, long vase life and fragrance. Most of these roses are well known as cut flowers and garden ornamental plants around the world. Besides, these cultivated roses are also economically important in cosmetic and perfumery industries. For instance, *Rosa x damascena* and *R. x centifolia* are the two predominantly cultivated roses for the productions of essential oil, rose water, rose concrete and rose absolute (Huang et al., 2009a, Rusanov et al., 2011). Rose products also have valuable applications in medicinal industry due to their pharmaceutical properties (antibacterial and antioxidant) (Boskabady et al., 2011). Although the fragrance of roses has an economical interest, the scent trait was the least important during rose breeding due to the various factors, which can affect the expression of scent, such as environment, maturity of the bloom and human perception differences (Zlesak, 2006). Among all currently known roses species, only about 20% of them are considered as “fragrant”, while the rest are either lightly-scented (about 50%) or non-scented (Schulz, 2003). However, due to the increasing interest in scented cut roses, various approaches in rose fragrance production have been started in order to reintroduce scented roses into the floral market.

Rose volatile profiles

Due to its complex and long history of cultivation, there is no single and easy classification system for roses. Roses can be classified according to their usage, for example, garden roses and cut roses. According to the horticultural rose classification developed by the World Federation of Roses and the American Rose Society, roses can be divided into three groups: species roses, old garden roses (before 1867) and modern roses (after 1867) (Zlesak, 2006). Species roses are also known as wild or botanical roses, they represent the group of roses that naturally occur, such as *R. gallica* and *R. gigantea*. The list of old garden roses and modern roses were presented in Zlesak, 2006. Some examples of old garden roses groups are Alba, Bourbon, Centifolia, Damask, Noisette, Portland and Tea; hybrid tea roses and miniature roses are examples of modern roses (Zlesak, 2006). Among these three main groups of roses, many of them were studied due to their appealing floral scent. More than 400 volatile compounds have been identified by analysing various roses with gas chromatography-mass spectrometer (GC-MS) (Flament et al., 1993, Schulz, 2003). Table 1 summarizes the major volatile compounds of some roses, among which three horticultural groups with their flower volatiles were collected by either solvent extraction or headspace (Antonelli et al., 1997, Joichi et al., 2005, Magnard et al., 2015). According to Table 1, most commonly volatiles found in roses are monoterpenoids (especially geraniol, citronellol and nerol) and phenylpropanoids (especially 2-phenylethanol), which originated from ancient European roses (Scalliet et al., 2002). However, the Chinese-origin roses (most *R. chinensis* roses; *R. gigantea*) have their characteristic scents: 1,3,5-trimethoxybenzene (TMB) and DMT (3,5-dimethoxytoluene), which bring the “tea”
aroma (Scalliet et al., 2002) and fatty acid derived volatiles (e.g. Z-3-Hexenol and Z-3-Hexenyl acetate) to many of their rose descendants.

Table 1. Several major volatile compounds present in individual roses that belong to one of the three groups: species rose, old garden rose and modern roses. Volatile data originated from Antonelli et al., 1997, Joichi et al., 2005 and Magnard et al., 2015.

<table>
<thead>
<tr>
<th>Name</th>
<th>Relative abundance of volatile compound*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Species roses</strong></td>
<td></td>
</tr>
<tr>
<td><em>Rosa gigantea</em></td>
<td>2-Phenylethanol</td>
</tr>
<tr>
<td><em>R. multiflora</em></td>
<td>DMT</td>
</tr>
<tr>
<td><strong>Old Garden roses</strong></td>
<td></td>
</tr>
<tr>
<td>*R. × alba cv. ‘Felicité Parmentier’</td>
<td>2-Phenylethanol, Benzy alcohol, Geraniol, Nerol, Geranyl acetate</td>
</tr>
<tr>
<td>*R. × alba cv. ‘Maidens Blush’</td>
<td>2-Phenylethanol, Benzy alcohol, Geraniol, Nerol, Geranyl acetate</td>
</tr>
<tr>
<td>*R. × alba cv. ‘Mme Plantier’</td>
<td>2-Phenylethanol, Benzy alcohol, Geraniol, Nerol, 2-Phenylethyl benzoate</td>
</tr>
<tr>
<td>*R. × alba cv. ‘Queen of Denmark’</td>
<td>2-Phenylethanol, Benzy alcohol, Geraniol, Nerol, Benzyl benzoate</td>
</tr>
<tr>
<td>*Bourbon R. cv. ‘Louise Odier’</td>
<td>2-Phenylethanol, Benzy alcohol, Geraniol, Nerol, Isoeugenol</td>
</tr>
<tr>
<td>*Bourbon R. cv. ‘Mme Isaac Pereire’</td>
<td>2-Phenylethanol, Geraniol, Nerol, Benzy alcohol, Geranyl acetate</td>
</tr>
<tr>
<td>*Bourbon R. cv. ‘Variegata di Bologna’</td>
<td>2-Phenylethanol, Geraniol, Nerol, Benzy alcohol, Benzyl benzoate</td>
</tr>
<tr>
<td>*R. x centifolia’ cv. ‘Blanche Moreau’</td>
<td>2-Phenylethanol, Geraniol, Geranyl acetate, Nerol, Isoeugenol</td>
</tr>
<tr>
<td>*R. x centifolia’ cv. ‘Old Pink Moss’</td>
<td>2-Phenylethanol, Benzy alcohol, Nerol, Geraniol, E,E-farnesol</td>
</tr>
<tr>
<td>*R. x centifolia’ cv. ‘Souvert et Notting’</td>
<td>2-Phenylethanol, Geraniol, Benzy alcohol, Nerol, Geranyl acetate</td>
</tr>
<tr>
<td>*R. chinensis cv. ‘Semperflorens’</td>
<td>Z-3-Hexenyl acetate, Z-3-Hexenol, 2-Phenylethanol, Dihydro-β-ionol, TMB</td>
</tr>
<tr>
<td>*R. chinensis var. spontanea</td>
<td>TMB</td>
</tr>
<tr>
<td>*R. chinensis cv. ‘Single Pink’</td>
<td>Z-3-Hexenyl acetate, Z-3-Hexenol, Dihydro-β-ionol</td>
</tr>
<tr>
<td>*R. chinensis cv. ‘Viridiflora’</td>
<td>Z-3-Hexenyl acetate, Z-3-Hexenol, TMB</td>
</tr>
<tr>
<td>*R. chinensis cv. ‘Mutabilis’</td>
<td>Geraniol, Geraniol, Germacrene D, E-2-Hexenal, Z-3-Hexenyl acetate, &amp; Linalool</td>
</tr>
<tr>
<td>*R. chinensis cv. ‘Miss Lowe’</td>
<td>Z-3-Hexenyl acetate, Z-3-Hexenol, 2-Phenylethanol, TMB, Dihydro-β-ionol</td>
</tr>
<tr>
<td>*R. chinensis cv. ‘Hermosa’</td>
<td>Z-3-Hexenyl acetate, Z-3-Hexenol, 2-Phenylethanol, TMB</td>
</tr>
<tr>
<td>*R. chinensis cv. ‘Old Blush’</td>
<td>Geraniol, Geraniol, Germacrene D, E-2-Hexenal, Linalool</td>
</tr>
<tr>
<td>*Rosa. x damascena cv. ‘Kazanlik’</td>
<td>Geraniol, Geraniol, Linalool, Benzy alcohol, Geranyl acetate</td>
</tr>
<tr>
<td>*R. x damascena cv. ‘Marie Louise’</td>
<td>Geraniol, Geraniol, Benzy alcohol, Isoeugenol</td>
</tr>
<tr>
<td>*R. x damascena cv. ‘Pompon des Prince’</td>
<td>Geraniol, Geraniol, Benzy alcohol, Isoeugenol</td>
</tr>
<tr>
<td>*R. x eglanteria cv. ‘Manning’s Blush’</td>
<td>Benzy alcohol, Geraniol, 2-Phenylethanol, Nerol, Benzyl benzoate</td>
</tr>
<tr>
<td>*R. x gallica cv. ‘Belle Isis’</td>
<td>Benzy alcohol, Geraniol, Benzy alcohol, Nerol, Geranyl acetate</td>
</tr>
<tr>
<td>*R. x gallica cv. ‘Charles de Mills’</td>
<td>Benzy alcohol, Geraniol, Benzy alcohol, Nerol, Geranyl acetate</td>
</tr>
<tr>
<td>*R. x gallica cv. ‘Jenny Duval’</td>
<td>Geraniol, Geraniol, Benzy alcohol, Benzy alcohol, Nerol</td>
</tr>
<tr>
<td>*R. x gallica cv. ‘Rose du Maître d’Ecole’</td>
<td>Geraniol, Geraniol, Benzy alcohol, Benzy benzoate</td>
</tr>
<tr>
<td>*R. x gallica cv. ‘Tour de Malakoff’</td>
<td>Geraniol, Benzy alcohol, E,E-farnesol</td>
</tr>
<tr>
<td>*Noisette R. cv. ‘Aimée Vibert’</td>
<td>Benzy alcohol, Nonanal + Z-rose oxide, Geraniol &amp; Benzy benzoate</td>
</tr>
<tr>
<td>*Noisette R. cv. ‘Alister Stella Grey’</td>
<td>Benzy alcohol, Geraniol, Linalool, Benzy benzoate</td>
</tr>
</tbody>
</table>

Table 1 continued...

1 Chapter 1 General introduction
Table 1. (Continued)

<table>
<thead>
<tr>
<th>Name</th>
<th>Relative abundance of volatile compound*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (Most abundant)</td>
</tr>
<tr>
<td>Old Garden roses</td>
<td></td>
</tr>
<tr>
<td>Perpetual R. cv. ‘Ferdinand Pichard’</td>
<td>2-Phenylethanol</td>
</tr>
<tr>
<td>Portland R. cv. ‘Compte de Chambord’</td>
<td>2-Phenylethanol</td>
</tr>
<tr>
<td>R. x rugosa cv. ‘Roserarie de l’Hay’</td>
<td>Geraniol</td>
</tr>
<tr>
<td>Tea R. cv. ‘Lady Hillingdon’</td>
<td>DMT</td>
</tr>
<tr>
<td>Modern roses</td>
<td></td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Diorama’</td>
<td>DMT</td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Grand Mogul’</td>
<td>DMT</td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Rouge Meilland’</td>
<td>Geraniol</td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Hacienda’</td>
<td>DMT</td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Pariser Charme’</td>
<td>Geraniol</td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Baccara’</td>
<td>Benzyl alcohol</td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Blue Moon’</td>
<td>Geraniol</td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Anna’</td>
<td>Germacrene D</td>
</tr>
<tr>
<td>R. x hybrida cv. ‘Black Baccara’</td>
<td>Benzyl alcohol</td>
</tr>
</tbody>
</table>

Note: * 1-5: indicates relative abundance of the indicated compound within the volatile blend, 1 indicates the most abundant and 5 indicates the least abundant. Abbreviation: TMB, 1,3,5-trimethoxybenzene; DMT, 3,5-dimethoxytoluene.

Uses of rose volatiles

As described above, just like many other flowers, roses are capable of producing various pleasant fragrance compounds, which originally served reproduction purposes. By comparing the volatile profiles of entomophilous and anemophilous plants, it was found that entomophilous plants emitted more floral volatile compounds and faster than the anemophilous plants, indicating fast emission of a complex blend of floral volatile compounds is required to facilitate the biotic pollination (Farré-Armengol et al., 2015). A scent-baited experiment using water-bowl traps demonstrated that some floral compounds, like phenylacetaldehyde, were strong attractants to pollinators but also attracted their unwelcome enemies, the florivores (Theis, 2006).

Apart from the general biological applications of floral volatiles, rose volatiles also have significant impacts in industry. For example, both geraniol and 2-phenylethanol give sweet floral rose-like odour (Burdock, 2002a, Burdock, 2002b), therefore they are in large demand in the flavour and fragrance markets. According to Schwab et al., until 2006, the annual consumption of both compounds (and their esters derivatives) reached more than 10,000 tonnes (Schwab et al., 2008). For the production of these scent compounds, roses are still widely used in perfumery industry. However, due to the high cost or lack of natural materials for flavour extraction, most of the commercial fragrance compounds are chemically synthesized via environmentally unfriendly processes (e.g. using heavy metal catalysts) (Guentert, 2007). Using bioproduction, including extraction from natural materials and conversion of natural precursors using micro-organisms or isolated enzymes is more desirable from a sustainable point of view (Schwab et al., 2008). Therefore, in the recent several years, the volatile compounds that have particular interests to the industry, like certain monoterpenoids and phenylpropanoids, are the most important target compounds for scientific research. As roses are valuable natural resources for flavour and fragrance industries, they draw significant attentions in order to find an economically efficient way to produce large quantities of desired compounds.
Biosynthesis of rose volatile compounds

So far, in rose, enzymes that are potentially involved in the biosynthesis of rose volatiles have been identified (Chen et al., 2011, Farhi et al., 2010, Guterman et al., 2002, Hirata et al., 2012, Hirata et al., 2016, Huang et al., 2009a, Huang et al., 2009b, Lavid et al., 2002, Magnard et al., 2015, Sakai et al., 2007, Scalliet et al., 2002, Scalliet et al., 2006, Shalit et al., 2003, Wang et al., 2012, Wu et al., 2004, Wu et al., 2003). Known biosynthesis pathways of several rose scent compounds are summarized in Figure 1.

Phenolic compounds

Two scent compounds: 1,3,5-trimethoxybenzene (TMB) and 3,5-dimethoxytoluene (DMT) are important scent components of many rose varieties nowadays, especially Tea roses and Hybrid Tea roses (Flament et al., 1993, Joichi et al., 2005, Nakamura, 1987, Scalliet et al., 2002, Scalliet et al., 2008). The “Tea” in their names is given by the fact that both compounds, especially DMT, give characteristic light, earthy and spicy notes sometimes said to be reminiscent of black tea smell. Because of their unique fragrance and widely existence in various rose varieties, TMB and DMT are among the very first compounds that were intensively studied. A family of enzymes, O-methyltransferases (OMT), appears to play an important role in the biosynthesis of both TMB and DMT. Two similar OMTs from *R. chinensis* cv. ‘Old Blush’ and *R. × hybrida* cv. ‘Lady Hillingdon’ can use orcinol and 3,5-dihydroxyanisole as substrates and produce DMT or TMB, depending on the availability of the substrates (Scalliet et al., 2002). At the same time, another group of scientists also found two highly homologous orcinol OMTs (OOMT) in *R. × hybrida* cv. ‘Fragrant Cloud’ and *R. × hybrida* cv. ‘Golden Gate’, which were able to methylate the precursor orcinol (methylated by OOMT1), generating an intermediate product: 3-hydroxy-5-methoxytoluene, which was subsequently methylated by OOMT2 and resulted in the final product, DMT (Lavid et al., 2002). In the biosynthetic pathway of TMB, this OMT enzyme family is not only involved in the final two steps of the pathway, in fact, it is involved in the last three successive steps. Another OMT (POMT) found in *R. chinensis* var. *spontanea* used phloroglucinol as initial precursor and methylated one of the three hydroxyl groups to generate 3,5-dihydroxyanisole. This compound can be further methylated by two OMTs to synthetize TMB (Wu et al., 2004). Apart from the biosynthetic pathways towards DMT or TMB, two other OMTs can also methylate eugenol and isoeugenol and yield methylcugenol and isomethylcugenol, respectively, in *R. chinensis* var. *spontanea* (Wu et al., 2003). So far, the subcellular locations of most of the identified OMT enzymes remain unknown, except for the OOMTs that are responsible for the biosynthesis of DMT. They were located in cytosol using a green-fluorescent-protein (GFP) tag and predominantly in both adaxial and abaxial epidermal cells of rose petals (Bergougnoux et al., 2007, Scalliet et al., 2006).

Besides the special “Tea” fragrance compounds, the biosynthetic pathways of two sweet floral fragrance compounds: 2-phenylacetaldehyde (PAld) and 2-phenylethanol (2PE) have also been intensively studied. The conversion of phenylalanine (Phe) to PAld via phenylacetaldehyde synthase (PAAS) was first discovered in *Petunia hybrida* cv. ‘Mitchell’ and in rose (Kaminaga et al., 2006). The conversion from Phe to PAld and then to 2PE via aromatic amino acid decarboxylase (AADC, identical to PAAS) and phenylacetaldehyde reductase (PAR) was also described by Sakai and colleagues, using feeding experiments of 8-deuterium-substituted 2PE (Sakai et al., 2007). Furthermore, the rose PAAS was characterized using a yeast (*Saccharomyces cerevisiae*) mutant that lacked phenylpyruvate decarboxylase activity (controlled by a gene called *aro10*) to further verify the biosynthetic pathway from Phe to 2PE via PAld (Farhi et al., 2010). A *aro10*-knockout (KO) yeast strain produced up to eight times less PAld and 2PE than the wild type (WT) and expression of *RhPAAS* could complement the KO mutant and increased the PAld production, confirming the involvement of *RhPAAS* in the biosynthesis of PAld and 2PE. The rose phenylacetaldehyde reductase (PAR) was functionally
characterized using recombinant PAR protein for enzymatic assay, which showed that its preferable substrate was PAlD (Chen et al., 2011). Interestingly, there is an alternative pathway from Phe to 2PE in rose. An aromatic amino acid aminotransferase (AAAT) can convert Phe into phenylpyruvate (PPA)

Figure 1 Summary of known biosynthesis pathways of volatile compounds in cytosol and plastid in rose. Terpenes precursors are normally synthesized via the cytosolic mevalonic acid (MVA) pathway and the plastidial methylerythritol phosphate (MEP) pathway. Abbreviations: ACC: acetyl-CoA carboxylase; AAAT: aromatic amino acid aminotransferase; AAT: alcohol acetyltransferase; CCD: carotenoid cleavage dioxygenase; DMAPP: dimethylallyl diphosphate; DMT: 3,5-dimethoxytoluene; EGS: eugenol synthase; FPP: farnesyl diphosphate; FPS: 1,6-FPP synthase; GDS: germacrene D synthase; GGPP: geranylgeranyl diphosphate; GGPS: GGPP synthase; GPP: geranyl diphosphate; GPS: GPP synthase; IDI: isopentenyl diphosphate isomerase; IGS: isoeugenol synthase; IPP: isopentenyl diphosphate; LIS: linalool synthase; NUDX: NUDIX hydrolase; OOMT: orcinol O-methyltransferase; OMT: O-methyltransferase; PAAS: phenylacetaldehyde synthase; PAR: phenylacetaldehyde reductase; POMT: pholoroglucinol O-methyltransferase; PPDC: phenylpyruvate decarboxylase; TMB: 1,3,5-trimethoxybenzene. *, The location and the function of these proteins are unproven. Double black arrow indicates multiple enzymatic reactions; single purple arrow indicates transportation.
(Hirata et al., 2012), which can be subsequently converted into PAld by a phenylpyruvic acid decarboxylase (PPDC) (Hirata et al., 2016). This pathway seems to be seasonal specific because the expression of PPDC and the production of 2PE were highly induced by high temperatures in summer (Hirata et al., 2016). As the study of biosynthesis of phenylpropanoids is still relatively recent, the subcellular locations of all these identified enzymes are still not revealed, although PAAS is predicted to be in the cytosol while PAR is predicted to be in the plastids according to the results of the submission of their amino acid sequences to WoLF PSORT (http://www.genscript.com/wolf-psort.html). In addition, a eugenol synthase (EGS) was discovered in rose (Wang et al., 2012, Yan et al., 2012) but neither their substrates nor their subcellular location are clear, although their substrates may very likely be coniferyl acetate based on the results obtained from the EGSs from sweet basil and petunia (Koeduka et al., 2006).

**Terpenoids and their derivatives**

Compared to phenolic volatile compounds, the biosynthesis of terpenoids in rose remains largely unexplored. Till today, only six enzymes have been identified that are involved in the biosynthesis of terpenoids in rose. The first terpene synthase (TPS) in rose, a germacrene D synthase (GDS) was discovered using genomic approaches by comparing the transcripts (expressed sequence tag library) from scented and unscented cultivars of roses. The biochemical function of GDS was determined using recombinant protein to react with substrate farnesyl diphosphate (FPP) and resulted in the production of germacrene D (Guterman et al., 2002). Later, a putative linalool synthase (RrLIS) was also isolated in rose using RT-PCR and rapid amplification of cDNA ends (RACE) technique (Feng et al., 2014). They found that the expression level of RrLIS was roughly correlated with the linalool content found in the rose, that is, when the expression of RrLIS gene reached the highest level in full opening period of the flower development stage, the linalool content also reached the highest. However, no biochemical experiment was conducted to verify the enzymatic activity of this protein. Therefore, it is still unclear whether this RrLIS gene is coding a linalool synthase or not. Recently, a Nudix hydrolase, which is not a terpene synthase, was found in rose and it was shown to be responsible for the biosynthesis of monoterpenoids, especially geraniol, in rose (Magnard et al., 2015). The studies of this protein are described in detail in Chapter 2.

Short chain terpenoids can not only be synthesized by TPS but can also come from degradation of long-chain terpenoids. Irregular terpenoids, i.e. C_{13}-norisoprenoids, like β-ionone, are present in rose petals in small quantities, but due to their low odor threshold, they may contribute significantly to the rose scent. These compounds are produced via degradation of carotenoids by carotenoid cleavage (di-)oxygenases (CCDs) in rose (Huang et al., 2009a, Huang et al., 2009b). The CCD1 protein can accept several different carotenoids as substrates in vitro and generate different C_{13} products (Huang et al., 2009a) while CCD4 protein accepted smaller range of substrates compared to CCD1 protein and only produced β-ionone (Huang et al., 2009b).

In addition to terpenoid-biosynthesis related proteins, an acetyltransferase in rose was responsible of the conversion of geraniol and citronellol into geranyl acetate and citronellyl acetate, respectively (Shalit et al., 2003). Furthermore, this acetyl-coenzyme A: geraniol acetyltransferase (AAT) coding gene was flower specific and reached the maximum transcript level at stage 4 of flower development, where the volatile emission reached at its peak.
References


Following the discovery of an alternative pathway for the biosynthesis of terpenoids in rose (see Chapter 2), we found that an increasing number of studies show that plants can employ different strategies for the biosynthesis of volatile molecules. The article presented here, published in Trends in Plant Science (October 2016, Vol.21, No. 10), summarizes our knowledge on these new alternative pathways.

**My way: Non-canonical biosynthesis pathways for plant volatiles**

Pulu Sun¹,², Robert C. Schuurink², Jean-Claude Caissard¹, Philippe Hugueney³ and Sylvie Baudino¹,*

1. Université de Lyon, UJM-Saint-Etienne, CNRS, BPam FRE 3727, F-42023, Saint-Etienne, France
2. Department of Plant Physiology, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands
3. SVQV, INRA, Université de Strasbourg, 68000 Colmar, France

*Correspondence: Sylvie.baudino@univ-st-etienne.fr (S. Baudino)

**Keywords:** biosynthesis, terpenes, benzenoids, phenylpropanoids

**Trends:**

Isopentenyl phosphate kinase (IPK) recycles the isopentenyl phosphate/dimethylallyl phosphate (IP/DMAP) pool that is likely to be formed by the action of phosphatases on IPP/DMAPP and contributes to the formation of both mevalonic acid (MVA) and methylyerythritol (MEP) pathway-derived terpenes.

Cytosolic monoterpene synthases (LdGES, FaNES1) and a diphosphohydrolase (RhNUDX1) have been linked to the biosynthesis of acyclic monoterpenes.

The complete iridoid biosynthesis pathway has been identified recently, indicating a complex biosynthesis network of monoterpenes in the cytosol. Z-prenyl diphosphates (NPP and Z,Z-FPP) can also act as substrates for mono- and sesquiterpene synthases, respectively.

There are three different routes for the biosynthesis of phenylacetaldehyde from phenylalanine.

**Glossary:**

Archaea: a microbe that has no cell nucleus or any other membrane-bound organelles in its cell (prokaryote), comprising a domain or kingdom of single-celled microorganisms, like bacteria. Archaea have isoprenoids as side chains in their membrane lipids instead of fatty acids as in bacteria.

Glandular trichomes: specialized “hairs” on the surface of plants that can produce and secrete metabolites such as terpenes.

(Plant) direct defense: characteristics of a plant that affect the physiology or behavior of herbivores negatively. Direct defense can be classified as: mechanical defense like thorns or thick cuticle; or chemical defense like toxic specialized metabolites.

(Plant) indirect defense: herbivore induced-characteristics of a plant that can enhance the performance of natural enemies of herbivores, using additional resources such as alternative food (extrafloral nectar, pollen) or herbivore-induced plant volatiles.

Polyphyletic group: a polyphyletic group is characterized by one or more similar phenotypes due to convergence or reversion from different ancestors.
Plant volatiles are crucial for various interactions with other organisms and their surrounding environment. A large number of these volatiles belong to the terpenoid and benzenoid/phenylpropanoid classes, which have long been considered to be exclusively synthesized from a few canonical pathways. However, several alternative pathways producing these plant volatiles have been discovered recently. This review summarizes the current knowledge about new pathways for these two major groups of plant volatiles, which open new perspectives for applications in metabolic engineering.

**Noncanonical metabolic pathways**

Plants produce various volatile organic compounds for different purposes like plant growth and development (e.g., ethylene, nitric oxide). Other volatiles can play a major role in interactions with their surrounding environment, including direct/indirect defense (see Glossary) against herbivores and pathogens, attraction of pollinators, and plant-plant interactions [1–6]. Up to 5-10% of assimilated carbon may be converted into plant volatiles and even more when plants are under stress [7,8]; in floral scents alone, more than 1700 different volatiles have been identified [9]. According to their biosynthetic origin and chemical structure, plant volatiles can be grouped into several classes, including: terpenoids, benzenoids/phenylpropanoids, fatty acid derivatives, amino acid derivatives, and carbohydrate derivatives such as furanones [2,3,9,10]. For decades it was believed that nearly all of these compounds were exclusively manufactured through a limited number of pathways shared by all plants. However, recently several publications indicate that different plants have evolved alternative pathways to produce the same compounds. These alternative pathways often occur through convergent evolution, which has been extensively discussed in [11].

In this review, we focus on the biosynthesis pathways of two major groups of plant volatiles: low-molecular-weight terpenoids (including monoterpenes, sesquiterpenes and homoterpenes) and benzenoids/phenylpropanoids. This review complements the latest reviews on volatile terpenoids [3,4,12] and benzenoids/phenylpropanoids [2,3,13] biosynthesis and specifically deal with the implications of recently discovered alternative metabolic pathways.

**New players in plant volatile biosynthesis**

Here we describe the recent and unexpected characterization of proteins belonging to families that had never been associated with volatiles before.

**Involvement of isopentenyl kinase in terpene biosynthesis**

Terpenes are one of the largest and most structurally diverse groups of natural products, which include more than 40,000 compounds [4]. All terpenes are constructed from two types of five-carbon molecules: isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) [14]. Sequential condensation of IPP and DMAPP leads to the formation of prenyl diphosphates, the precursors of most terpenes, i.e. geranyl diphosphate (GPP, C\(_{10}\)), farnesyl diphosphate (FPP, C\(_{15}\)), geranylgeranyl diphosphate (GGPP, C\(_{20}\)) and geranylfarnesyl diphosphate (GFPP, C\(_{25}\)). Generally, GPP and FPP are the precursors of monoterpenes (C\(_{10}\)) and sesquiterpenes (C\(_{15}\)) respectively, while GGPP is the precursor of diterpenes (C\(_{20}\)). GFPP is the newly discovered precursor of sesterterpenes (C\(_{25}\), [15]). Sequential condensation of FPP and GGPP results in precursors of triterpenes (C\(_{30}\)) and tetraterpenes (C\(_{40}\)), respectively. Further catalytic reactions (e.g., cyclization of an acyclic precursor followed or not by Wagner-Meerwein rearrangements, hydroxylation, dehydroxylation, reduction, oxidation, or glycosylation) on the parental skeletons give rise to a great diversity of various terpenoid-derived products.
There are two well-established pathways generating IPP and DMAPP: the cytosolic mevalonic acid (MVA) pathway and the plastidic mevalonolylthiotritol phosphate (MEP) pathway (reviewed in [16]). In the MVA pathway, IPP is generated by decarboxylation, from mevalonate-5-diphosphate by mevalonate diphosphate decarboxylase (MVD or MDD) [17]. In the MEP pathway, IPP is generated from 4-hydroxy-3-methylbut-2-enyl-diphosphate (HMBPP) by HMBPP reductase [18]. In the MEP pathway, both IPP and DMAPP can be produced by HMBPP reductase [19]. In the MVA pathway, IPP is converted into DMAPP by IPP isomerase (IDI) [20], which also operates in plastids and mitochondria [21].

However, recent studies have shown that the genes encoding the final two enzymes of the MVA pathway, namely phosphomevalonate kinase and MDD, are absent in Archaea. These organisms use an alternative route via isopentenyl phosphate kinase (IPK) for the formation of IPP [22]. Dellas and colleagues found that IPK also exists in all green plants whose genome has been sequenced. Fully functional IPKs are distributed across all three domains of life; thus, this does not seem to be a case of convergent evolution. In Arabidopsis thaliana, it was shown that IPK is a new member of the plant terpenoid metabolic network [23]. This cytosolic IPK is coexpressed with various genes involved in the MVA pathway and downstream terpene biosynthesis pathways, including MDD, farnesyl diphosphate synthase, squalene synthase and C-8,7 sterol isomerase [23]. Using A. thaliana ipk and ipk/mdd1 double mutants, it was found that IPK functions in parallel to the MVA pathway. In addition, overexpression of the cytosolic AtIPK in tobacco increased both the production of FPP- and GPP-derived terpenes in the cytosol and plastids, suggesting a communication between these two compartments. The proposed function of IPK in A. thaliana is then to recycle the isopentenyl phosphate/dimethylallyl phosphate (IP/DMAP) pool that is likely formed by the action of phosphatases on IPP/DMAPP (Figure 1).

Involvement of a Nudix hydrolase in monoterpane biosynthesis

Monoterpenes are well known as components of floral scent [24] and also as constituents of essential oils of aromatics plants that are widely used in the cosmetic, perfume, food [25], and pharmaceutical industries [26]. Because of their high economic relevance, monoterpane biosynthesis has been extensively studied. The biosynthesis of monoterpenes usually requires the catalytic reaction triggered by monoterpane synthases, converting C_{10} prenyl diphosphates to various cyclic and acyclic products [27]. No monoterpane synthase has so far been functionally characterized in rose (Rosa sp.), although sesquiterpene synthases have been associated with the biosynthesis of sesquiterpenes in this genus [28]. Recently, a terpene synthase-independent pathway to monoterpenes has been reported in rose. By comparing the volatile profiles and differential gene expression of scented and unscented rose cultivars, an unexpected enzyme, a Nudix hydrolase, RhNUDX1 was found to be responsible for the formation of geraniol [29]. The function of RhNUDX1 was elucidated by using recombinant protein on several potential substrates, by co-localizing the gene with a major QTL for geraniol production, and by analyzing RNAi-RhNUDX1 transgenic plants. The results indicated that RhNUDX1 used GPP as the substrate, as classical terpene synthases, but hydrolyzed one phosphate resulting in geranyl monophosphate (Figure 1). Thus, a further catalytic reaction with the assistance of an unidentified phosphatase to form geraniol is required. This mode of action differs from all known monoterpane synthases that remove two phosphates in one reaction [29]. In addition, RhNUDX1 enzyme is located in the cytosol, indicating that its substrate, GPP, was obtained by either transport from plastids or via cytosolic generation, but this remains to be solved. The discovery of such a unique pathway has provided a potential molecular marker for scented-rose breeding. Meanwhile, it also has raised a lot of questions on, for example, how this pathway evolved and how is it distributed in the plant kingdom [30].
Iridoids represent a broad class of monoterpenoids derived from 10-oxogeranial that have interesting pharmaceutical and antibacterial properties [31]. Madagascar periwinkle (Catharanthus roseus) has been used as a model plant for extensive studies of the iridoid pathway. The monoterpane branch of this pathway starts with the formation of geraniol by a classical plastidial geraniol synthase [32]. Geraniol 10-hydroxylase (G10H), a cytochrome P450 enzyme, which is anchored to the endoplasmic reticulum (ER), converts geraniol into 10-hydroxygeraniol [33]. Since the downstream protein G10H is located in the cytosol [34], this indicates that in C. roseus, geraniol is transported from plastids to cytosol perhaps through stromules, as they are closely associated with the ER [35]. The product from the first reaction, 10-hydroxygeraniol, is then converted by a
cytosolic 10-hydroxygeraniol dehydrogenase (Cr10HGO, also called 8HGO) into 10-oxogeranial (also 8-oxogeranial) [31, 36]. Subsequently, 10-oxogeranial serves as substrate for the newly discovered iridoid synthases, which show high similarity to progesterone-5β-reductases [37]. This unusual cyclization step leads to the formation of cyclic monoterpenoids (e.g. Z-E-nepetalactol), which are in turn converted to iridoids [31, 37-40]. Thus this is another example of a new non-canonical player involved in terpene biosynthesis.

**Biosynthesis of vanillin**

The most important phenylpropanoid volatile as seen from an economic perspective is certainly vanillin. It was recently shown that a single hydratase/lyase type enzyme designated vanillin synthase (VpVAN) catalyses direct conversion of ferulic acid or its glucoside into vanillin or vanillin glucoside, respectively [41]. The conversion of ferulic acid to vanillin catalysed by VpVAN is thought to occur via an initial hydration addition reaction followed by a retro-aldol elimination reaction. The VpVAN protein shows high sequence similarity to cysteine proteinases, including some important conserved residues. The cysteine protease family encompasses a large group of enzymes with versatile physiological functions. The involvement of such proteins in volatile biosynthesis is unexpected and their characterization offers new opportunities for the vanilla-pod based industries.

**New substrates for old enzymes**

As already mentioned, classical substrates for monoterpane synthases and sesquiterpene synthases are GPP and E,E-FPP, respectively. However, it has been shown that some enzymes use alternative substrates.

First, some plastidial monoterpane synthases use the Z isomer form of GPP, neryl diphosphate (NPP), to form monoterpenes. A plastidial Z-prenyltransferase, neryl diphosphate synthase 1 (NDPS1/SICPT1) has been recently identified in the glandular trichomes of cultivated tomato (Solanum lycopersicum) [42, 43]. Recombinant NDPS1/SICPT1 protein was able to use both IPP and DMAPP to generate NPP [42, 43]. By analyzing transgenic tomato plants with reduced NDPS1/SICPT1 transcript levels, a decrease of in the production of the monoterpane β-phellandrene was observed [43]. Both results indicate that NDPS1/SICPT1 is providing NPP substrate for the biosynthesis of monoterpenes.

Alternative substrates such as NPP are not routinely tested in functional characterization of plant terpene synthases. So far, few identified plastidial monoterpane synthases have shown preference for NPP rather than GPP as substrates. Well-described examples are: i) a phellandrene synthase (SIPHS1) [42]; ii) a limonene synthase (ShLMS) [44]; iii) a pinene synthase (ShPIS) [44]; and iv) a nerol synthase (GmNES) [45]. Recombinant SIPHS1 protein, which was highly expressed in the glandular trichomes of S. lycopersicum, converted NPP mainly to β-phellandrene and, to a lesser extent, to four other monoterpenes: δ-2-carene, α-phellandrene, limonene and γ-terpinene [42]. A similar result was found in another *in vitro* study of the recombinant PHS1 protein of wild tomato (S. habrochaites) (ShPHS1) that exhibited high specificity to NPP and generated β-phellandrene and δ-2-carene [44]. The NPP preference characteristic of PHS1 could be further confirmed by the overexpression of NDPS1 and PHS1 in S. lycopersicum fruits, which resulted in a monoterpane blend [46]. Another two monoterpane synthases: ShLMS and ShPIS found in the glandular trichomes of S. habrochaites solely used NPP as substrates and mainly generated limonene and α-pinene, respectively [44]. In soybean (*Glycine max*), a nerol synthase (GmNES) exclusively converted NPP to nerol [45].

Sesquiterpenes are a group of C₁₅ terpenes that play important roles in various physiological and ecological processes of plants. Apart from being constituents of floral scents, sesquiterpenes are also
known as repellents of insects (e.g. 7-epizingiberene and R-curcumene) [47] and as antimicrobial compounds (e.g. E-β-caryophyllene) [48]. It is generally agreed that E,E-FPP is the precursor for the biosynthesis of most sesquiterpenes and their corresponding sesquiterpene synthases are located in the cytosol. However, in glandular trichomes of wild tomatoes, sesquiterpenes can also be synthesized from Z,Z-FPP. Sallaud et al. [49] discovered a Z,Z-farnesyl diphosphate synthase (ZFPS) in S. habrochaites that catalyzed the formation of Z,Z-FPP from IPP and DMAPP, and a sesquiterpene synthase, santalene and bergamotene synthase (SBS), which used Z,Z-FPP as substrate to generate a mixture of these two sesquiterpenes. In a later study, additional Z-prenyltransferases were found in S. lycopersicum, with SICPT6 that was able to convert NPP into Z,Z-FPP in roots and red fruits [43]. More sesquiterpene synthases that use Z,Z-FPP, for example zingiberene synthase (ShZIS) – to generate 7-epizingiberene in glandular trichomes of S. habrochaites have also been identified [44, 50].

The puzzling localization of enzymes

Almost all of the identified monoterpene synthases have been located inside plastids, and, with a few exemptions inside mitochondria [51, 52], either based on predicted plastid transit peptides or by visualization using fusions with green fluorescent protein (GFP). However, monoterpene biosynthesis may also be present in the cytosol (Figure 1). The idea that monoterpene can be made in the cytosol was brought back to the synthesis of shikonin [53]. The first discovered cytosolic monoterpene synthase is actually a dual-function terpene synthase, nerolidol synthase 1 in Fragaria x ananassa (FaNES1), which can react with both GPP and FPP to produce linalool or nerolidol, respectively. Its localization in the cytosol was demonstrated by a GFP fusion of the FaNES1 protein [51]. In the same study, FvPINS, a genuine monoterpene synthase was also localized in the cytosol. Moreover, (S)-linalool and (−)-α-pinene in the fruit of F. x ananassa and F. vesca, respectively, were found, using deuterium (²H) labeling, to be generated exclusively via cytosolic MVA pathway [54]. In raspberry fruits, feeding experiments also demonstrated that (S)-linalool and (−)-α-pinene were exclusively synthetized via the cytosolic MVA pathway [55]. These studies in Rosaceae revealed that monoterpene can be also synthesized in the cytosol rather than only in plastids and that precursors probably come from the MVA pathway. Cytosolic monoterpene synthases were also discovered in other species. In Lippia dulcis, a cytosolic geraniol synthase (LdGES) has been characterized [56]. Both transient and stable transformation of tobacco confirmed monoterpene synthase activity of LdGES, resulting in the production of geraniol-derived products. Studies with GFP fusions of LdGES by transient transformation assays in Nicotiana benthamiana indicated its cytosolic localization [56]. Ocimum basilicum α-zingiberene synthase (ZIS) is a cytosolic bifunctional terpene synthase, which not only takes FPP but also GPP as substrate to produce both sesqui- and monoterpene [57].

The presence of these cytosolic monoterpene synthases or bifunctional enzymes suggests that there is a complex biosynthetic network of monoterpene in cytosol for some species. This also raises the question as to the origin of their substrates. Are they exported from plastids or are they synthesized in the cytosol? Several lines of evidence indicate that molecules such as GPP can be exported from plastids to cytosol. Co-expression in tomato fruits of the snapdragon small catalytic unit of GPP synthases (GPPSSSU, plastidial) and the cytosolic bifunctional ZIS from O. basilicum [57], resulted in increased accumulation of ZIS-derived monoterpene, demonstrating the transport of GPP from plastid to cytosol [58]. The hypothesis of GPP export from plastids is also supported by a kinetic study of the export of prenyl diphosphates in isolated chloroplast envelope membrane vesicles of spinach (Spinacea oleracea) [59]. They showed that the export rate of IPP and GPP is relatively higher than that of DMAPP and FPP. Finally, ectopic expression of GES from Valeriana officinalis in three different cell compartments (plastids, cytosol and mitochondria) showed that there is considerable (b)directional GPP exchange between these compartments [60]. In addition, no cytosolic GPP synthase (GPPS) has been found so far, with the possible exception of the one from Lithospernum erythrorhizon, which has not been characterized [61]. Although these experimental
results indicates that plastid-generated GPP could support a cytosolic monoterpene biosynthesis, no transporter of IPP or GPP has been characterized to date. Another possibility, albeit less probable, is that GPP could be generated in the cytosol as a byproduct of FPPS activity. Indeed, as GPP is an intermediate in the synthesis of FPP, it could be that part of it is released and used for the biosynthesis of monoterpenes in the cytosol. However, the amount of GPP available might not be sufficient to support the biosynthesis of a large quantity of monoterpenes.

Sesquiterpenes were first thought to be synthesized exclusively in the cytosol, using precursors generated via the MVA pathway. However there seems to be some exceptions to this rule. Labeling experiments in snapdragon flowers have shown that IPP and DMAPP, used for the biosynthesis of sesquiterpenes, were produced in plastids via the MEP pathway [62]. Since snapdragon sesquiterpene synthases are localized in the cytosol [63] this indicates a transport of IPP/DMAPP from plastids to cytosol. More recently it has become clear that all identified Z-prenyltransferases, related to mono- and sesquiterpene synthases, are located inside plastids [43, 45, 49], suggesting that the major pool of Z-prenyl diphosphates may be mainly produced in plastids. Concomitantly, several sesquiterpene synthases able to use these Z-prenyl diphosphates substrates in the plastids have been identified in wild tomato [50]. Both the MVA and MEP pathways contribute to the biosynthesis of sesquiterpenes at equal rates in the glandular trichomes of Stevia rebaudiana [64] and both routes are also utilized in grape berries [65]. Thus, there are multiple examples of “metabolic crosstalk” in the biosynthesis of sesquiterpenes, as summarized in [66].

Some diterpenes are also made in the cytosol. Indeed, while diterpene synthases are primarily located in the plastid, geranyllinalool synthase from A. thaliana, which is involved in the biosynthesis of (E,E)-geranyllinalool, the precursor of 4,8,12-trimethyltrideca-1,3,7,11-tetraene (TMTT), resides in the cytosol or the ER [67]. It is thus likely that its substrate GGPP is present in these compartments.

Different pathways for the same compounds

Biosynthesis of homoterpenes

Homoterpenes (more precisely tetra-nor-polyterpenes, but in this field commonly referred to as homoterpenes) belong to a special class of terpenes that are generated through degradation of regular terpenes. They are known to be floral constituents of night-scented flowers like Orchidaceae and Liliaceae as well as damaged-induced leaf volatiles that may attract natural enemies of herbivore insects [68-72]. The two most studied homoterpenes are the C\textsubscript{11} homoterpene 4,8-dimethylnona-1,3,7-triene (DMNT) and C\textsubscript{16} homoterpene TMTT, which were thought to be derived from sesquiterpenoids (e.g. nerolidol) and diterpenoids (e.g. geranyllinalool) by oxidative degradation, respectively [68]. In A. thaliana it was shown that one P450 enzyme (CYP82G1) was able to catalyze the formation of both compounds [70]. However, a recent study showed that in A. thaliana, the precursor for biosynthesis of DMNT is not limited to nerolidol only [73]. Very interestingly, DMNT can be generated by degradation of a C\textsubscript{30} triterpenoid, arabidol, in A. thaliana roots. This degradation reaction was catalyzed by another cytochrome P450 monoxygenase CYP705A1 (At4g15330), probably located in the ER [74]. The reaction is induced by a root oomycete pathogen, Pythium irregulare, and contributes to the plant defense. Thus, even within one species, in this case A. thaliana, multiple pathways lead to the same homoterpene, albeit in different tissues.

Biosynthesis of geranial

As described above, geraniol can be produced by two different pathways. This is also true for the corresponding aldehyde geranial. In sweet basil and perilla, geranial is formed from geraniol by alcohol dehydrogenases [75-77]. However, monoterpenes can also be derived from long-chain terpenoids. Carotenoids cleavage dioxygenase 1 (CCD1) from rice (Oryza sativa var. japonica cv. TP309) uses the carotenoid (C\textsubscript{40}) lycopene as substrate to generate geranial, and some C\textsubscript{6}/C\textsubscript{13}
ketones by oxidative cleavage of C7-C8/C7’-C8’ double bonds of carotenoid backbones [78]. CCDs also catalyze the formation of a large number of tomato fruit volatiles, including cis-pseudoionone, neral, geranial, and farnesylacetone [79].

**Biosynthesis of 2-phenylethanol (2-PE)**

In terms of diversity, benzenoids and phenylpropanoids are considered the second largest class of plant volatiles after terpenoids [2]. They are also involved in various interactions with the surrounding environment [2, 3]. This class of compounds is mainly derived from L-phenylalanine (Phe), one of the three aromatic amino acids synthesized via the shikimate pathway [80]. Comprehensive reviews on the biosynthesis of benzenoids [13] and phenylpropanoids [81] are available. Here we only shortly emphasize the alternative routes that plants have developed to make the same compounds.

Briefly, benzenoids are C_6-C_4 compounds, having an aromatic six-carbon ring with one carbon attached. Their formation is initialized by L-phenylalanine ammonia lyase (PAL), which de-aminates Phe to t-cinnamic acid (CA) (Figure 2). The subsequent route from CA to benzenoid-related products proceeds via the β-oxidative pathway (in peroxisomes), the non β-oxidative pathway (in cytosol) or a combination of both pathways, as nicely reviewed recently by Widhalm and Dudareva [13].

Other volatiles derived from Phe – for example phenylacetaldehyde (PAld) and 2-PE – are important scent compounds in numerous flowers such as rose and petunia. The genes and enzymes responsible for the biosynthesis of these phenylpropanoid-related compounds (C_6-C_2 aromatic six-carbon ring with two carbons attached) have been identified and characterized. Interestingly, plants have developed different pathways to convert Phe to PAld, all located in the cytosol (Figure 2). In petunia (P. hybrida cv. Mitchell), a bifunctional phenylacetaldehyde synthase (PAAS) can achieve both decarboxylation and oxidation reactions and convert Phe to PAld [82]. In contrast, in tomato (S. lycopersicum cv. M82) fruits, Phe is first converted to phenethylamine (PEA) by an aromatic L-amino acid decarboxylase (AADC) and then subjected to oxidation by an, yet to be identified oxidase [83]. The conversion from PAld to 2-PE in tomato is catalyzed by a phenylacetaldehyde reductase (PAR) [84]. In roses two enzymes have been identified, RhPAAS of Rosa x hybrida cv. Fragrant Cloud [82, 85] and Rose-PAR of R. x damascena Mill. [86], which possess similar functions as PAAS in petunia and PAR in tomato, respectively. Surprisingly, another pathway leading to PAld biosynthesis from Phe also exists in rose. In this pathway, Phe is first converted into phenylpyruvate (PPA) by an aromatic amino acid aminotransferase (RyAAAT3) that then undergoes decarboxylation to form PAld by a phenylpyruvic acid decarboxylase [87, 88]. This is another example, as in homoterpene biosynthesis in A. thaliana, of two different routes used in one species to produce the same compound. The difference with A. thaliana being that here, the two alternative pathways are supposed to function in parallel in the same petal tissue. Finally, an additional route was discovered recently in the fruit of melon (Cucumis melo), in which Phe is converted by an amino acid transaminase (CmArAT1) into its corresponding α-keto acids, which are efficient precursors for subsequent modifications to synthesize various aromatic compounds, including PAld [89].

**Concluding remarks and future perspectives**

In recent years, numerous studies have revealed that plants have evolved complex routes for terpene biosynthesis. For example, monoterpenes, which can be synthesized in both plastid and cytosol, can be generated from GPP, NPP and even from long-chain terpenoids, and can be synthesized by canonical monoterpenene synthases or by a terpene synthase-independent pathway (RhNUDX1 in roses). Similar plasticity for the biosynthesis of benzenoids and phenylpropanoids is also observed. The diversity of these volatile biosynthesis pathways is often an example of convergent evolution: multiple pathways giving rise to similar or identical products have evolved.
Figure 2: Known and novel plant volatile benzenoid and phenylpropanoid biosynthesis pathways. For a Figure360 author presentation of Figure2, see http://dx.doi.org/10.1016/j.tplants.2016.07.007#mmc1. Abbreviations: 4CL, 4-coumarate CoA-ligase; AAAT, amino acid aminotransferase; AADC, aromatic amino acid decarboxylase; AOs, aldehyde oxidases; ArAT, amino acid transaminase; BPBT, benzoyl-CoA:benzylalcohol/2-phenylethanol benzoyltransferase; BSMT, benzoic acid/salicylic acid carboxyl methyltransferase; BZ01, homolog gene of cinnamoyl CoA ligase in A. thaliana; C4H, cinnamic acid 4-hydroxylase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CFAT, coniferylalcohol acetyltransferase; CHD, cinnamoyl-CoA hydratase/dehydrogenase; CNL, cinnamoyl-CoA ligase; EGS1, eugenol synthase 1; IGS1, isoeugenol synthase 1; KAT, 3-ketoacyl-CoA thiolase; PAL, phenylalanine ammonial yase; PAAS, phenylacetaldehyde synthase; PAR, phenylacetaldehyde reductase; pCAT, plastidial cationic amino acid transporter; PEA, phenethylamine; PPDC, phenylpyruvate decarboxylase; PXA1, peroxisomal ATP-binding cassette (ABC) transporter1.

Independently across various taxa, perhaps as the result of different adaptative responses to similar environments or ecological niches [11]. In rose, the alternative pathway identified for the production of 2-phenylethanol could be activated in response to seasonal changes in temperature [88]. Examples of convergent evolution occur also in the biosynthesis of other, non-volatile, specialized metabolites. For example, the N-methyltransferases (NMTs) genes, that are involved in caffeine formation in Coffea canephora, nest within a gene clade distinct from those of cacao (Theobroma cacao) and tea (Camellia sinensis), which suggests a polypheletic origin of NMT activity [90]. Thus, due to the biodiversity in land plants, it seems very probable that there exist several, yet to be discovered, alternative pathways for the biosynthesis of identified plant volatile compounds. As for Nudix in rose, previously characterized enzymes may possess completely different functions in other plants. It is possible that the capacity of Nudix phosphohydrolases has been recruited for other functions in different plants. For example, another Nudix might be able to use IPP as substrate to produce IP, which seems to limit the biosynthesis of terpenes in the cytosol [23]. Moreover, other
novel enzymes could be discovered over time during extensive studies of plant volatile biosynthesis (see outstanding questions). The use of RNA sequencing as well as available plant and microbial genomes will facilitate the future identification of these proteins [91-93].

**Outstanding Questions**

How are the prenyl diphosphates, precursors of terpenes, transported between subcellular compartments? What are the proteins responsible for such transport?

Could geranyl diphosphate be synthesized in the cytosol? If yes, what are the enzymes responsible for such biosynthesis?

Are there other distinct enzymes (besides NUDX1 and terpenesynthases) involved in the biosynthesis of terpenes?

When two pathways operate in the same species for the production of one compound, what are the molecular mechanisms that regulate the two pathways?

What are the environmental parameters that favor the evolution of alternative pathways for the production of a given compound?

**Acknowledgments**

S.B., J.C.C., and P.S. were supported by funding from the Région Rhône-Alpes and CNRS (GDR MediatEC 3658 and PEPS ExoMod), France. The authors thank Christian Dumas (ENS Lyon, France) for his support and scientific input.
References

10. Song, C., et al. (2016) Glucosylation of 4-Hydroxy-2,5-Dimethyl-3(2H)-Furanone, the key strawberry flavor compound in strawberry fruit. *Plant Physiol.* 171, 139-151
65. May, B., et al. (2013) Biosynthesis of sesquiterpenes in grape berry exocarp of Vitis vinifera L.: evidence for a transport of farnesyl diphosphate precursors from plastids to the cytosol. Phytochemistry 95, 135-144
68. Boland, W., et al. (1998) Biosynthesis of C\textsubscript{11} and C\textsubscript{16} homoterpenes in higher plants; stereochemistry of the C-C-bond cleavage reaction. Tetrahedron 54, 14725-14736


79. Ilg, A., et al. (2014) Tomato carotenoid cleavage dioxygenases 1A and 1B: Relaxed double bond specificity leads to a plenitude of dialdehydes, mono-apocarotenoids and isoprenoid volatiles. *FEBS Open Bio* 4, 584-593


Research objectives and outlines

In order to advance our knowledge of the biosynthesis of scent volatiles in roses, this thesis project aimed mainly to identify the genes and their related proteins, and to study their biological functions regarding to the biosynthesis of interesting scent compounds produced by roses. At the beginning of this project, a Nudix hydrolase, RhNUDX1 had been found in our group to correlate with the production of monoterpenoids in rose. My project is the continuation of this project. To be more precise, my objectives were to answer the following questions:

1. How many NUDX1 genes exist in rose and what are the biological functions of these NUDX1 proteins?
2. How are these NUDX1 genes transcriptionally regulated?
3. What information can RNA-sequencing and transcriptome data analysis provide for other rose scent-related genes?

Chapter 2 describes a series of experiments in which the rose Nudix hydrolase 1 (NUDX1) was identified and characterised. Genetic, molecular biology and biochemistry approaches were applied for the gene identification and the protein characterization; rose volatiles were analysed using gas-chromatography mass spectrometry (GC-MS) and correlation with the expression of rose NUDX1 gene was investigated in order to gain insight in its biological function. These results are presented in the form of an article, published in Science (Magnard et al., 2015).

Chapter 3, presented in the form of an article to be submitted, illustrates the identification and characterization of another NUDX1 genes in rose, NUDX1-2. Consequently, the first isolated gene was renamed NUDX1-1. NUDX1-2 function was studied using molecular biology and biochemistry approaches similar to the approaches described in chapter 2. In addition, the functions of all identified NUDX1 proteins were compared, regarding to their biochemical activities and their protein structures. Transient expression studies were conducted in order to study the function of NUDX1 proteins in planta.

Chapter 4 shows two pilot experiments on the promoters of three rose NUDX1 genes and two transcription factor (TF) candidates. Several tests were performed to determine the activity of each rose promoter. Moreover, each TF candidate was tested for its ability to activate each of the three NUDX1 promoters.

Chapter 5 contains the bioinformatic study on RNA-sequencing results of four rose samples within a rose progeny, including both parents (R. chinensis cv. ‘Old Blush’ and R. x wichurana) and two of their progenies OW9035 and OW9047. They were selected based on their specific scent profiles that R. chinensis cv. ‘Old Blush’ and OW9047 produce geraniol but not farnesol, while R. x wichurana and OW9035 produce farnesol but not geraniol. The RNA-sequencing quality was analysed and the general annotation was performed for the whole dataset. In addition, using the volatile profiles of those four rose samples, other scent-related genes were identified as well as their potential roles in the secondary metabolites biosynthesis pathways. Furthermore, all annotated TFs were also listed for further study.

Finally, chapter 6 summarises the findings in relation to the research objectives formulated above. Research directions and perspectives are also discussed.