GRAPHICAL ABSTRACT

Lilium floral fragrance: a biochemical and genetic resource for aroma and flavor

Timothy S. Johnson, Michael L. Schwieterman, Joo Young Kim, Keun H. Cho, David G. Clark,
and Thomas A. Colquhoun

Floral volatile characterization of five *Lilium* cultivars reveals specific and diverse chemical profiles, which uncovers a simple platform to perform comparative transcriptomics for volatile biosynthesis gene discovery.

Color figure is for online reproduction on the web.

1
2
3
4 **TITLE PAGE**
5
6
7
8

9 *Lilium* floral fragrance: a biochemical and genetic resource for aroma and flavor
10
11
12
13

14 Timothy S. Johnson^{a,c}, Michael L. Schwieterman^{a,b,c}, Joo Young Kim^{b,c}, Keun H. Cho^{b,c}, David
15
16 G. Clark^{a,b,c}, and Thomas A. Colquhoun^{a,b,c,*}
17
18
19
20

21 ^aPlant Molecular and Cellular Biology Program, University of Florida, Gainesville, Florida,
22
23 32611, USA
24
25
26
27

28 ^bEnvironmental Horticulture Department, Institute of Food and Agricultural Sciences, University
29
30 of Florida, Gainesville, Florida, 32611, US
31
32
33
34
35

36 ^cPlant Innovation Center, University of Florida, Gainesville, Florida, 32611, USA
37
38
39
40

41 *Thomas A. Colquhoun, 1523 Fifield Hall, Gainesville, FL. 32611 U.S.A., office: 352-273-
42
43 4577, Fax: 352-392-3870, Email: ucntcmel@ufl.edu
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

ABSTRACT

Hybrid *Lilium* (common name lily) cultivars are among the top produced domestic fresh cut flowers and potted plants in the US today. Many hybrid *Lilium* cultivars produce large and showy flowers that emit copious amounts of volatile molecules, which can negatively affect a consumer's appreciation or limit use of the plant product. There are few publications focused on the biochemistry, genetics, and/or molecular regulation of floral volatile biosynthesis for *Lilium* cultivars. In an initial pursuit to provide breeders with molecular markers for floral volatile biosynthesis, a total of five commercially available oriental and oriental-trumpet hybrid *Lilium* cultivars were selected for analytical characterization of floral volatile emission. In total, 66 volatile molecules were qualified and quantitated among all cultivars. Chemical classes of identified volatiles include monoterpene hydrocarbons, monoterpene alcohols and aldehydes, phenylpropanoids, benzenoids, fatty-acid-derived, nitrogen-containing, and amino-acid-derived compounds. In general, the floral volatile profiles of the three oriental-trumpet hybrids were dominated by monoterpene hydrocarbons, monoterpene alcohols and aldehydes, while the two oriental hybrids were dominated by monoterpene alcohols and aldehydes and phenylpropanoids, respectively. Tepal tissues (two petal whirls) emitted the vast majority of total volatile molecules compared to the reproductive organs of the flowers. Tepal volatile profiles were cultivar specific with a high degree of distinction, which indicates the five cultivars chosen will provide an excellent differential genetic environment for gene discovery through comparative transcriptomics in the future. Cloning and assaying transcript accumulation from four floral volatile biosynthetic candidates provided few immediate or obvious trends with floral volatile emission.

1
2
3
4
5
6
7 **KEYWORDS**
8
9

10
11 Interspecific lily hybrids; *Lilium* section *Archelirion*; Liliaceae; oriental lily hybrids, oriental-
12 trumpet lily hybrids; floral volatile profiling and quantitation; GC-MS; metabolic profiling;
13
14
15
16 tepals; terpenes; phenylpropanoids
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1. INTRODUCTION

Many angiosperm species synthesise and emit volatile organic molecules from floral tissues. These floral bouquets of volatile compounds can be as simple as a few chemically related volatile molecules to a mixture of a hundred volatile molecules produced from divergent metabolic pathways (Dudareva et al., 2013). One of the most well studied floral volatile functions is the role in sexual reproduction through pollinator attraction (Kessler et al., 2008; Klahre et al., 2011). For some angiosperm species, when a flower reaches sexual maturity, the recruitment of an active insect pollinator becomes paramount. This can be accomplished by broadcasting a chemical signal that will attract the pollinator to the flower, and increase the probability of pollen contact. However, the requirement for the correct pollinator based on flower structure, size, and/or location is fundamentally important (Andrews et al., 2007; Strauss and Whittall 2006).

This paradigm also holds true for human attraction to flowers. In the 1993 book *The Culture of Flowers*, Sir John Rankine Goody clearly illustrated the human fascination with flowers via anthropologic examples of human flower usage: in written language, social rituals, art works, and food stuffs (Goody, 1993). Breeding of ornamental angiosperm species dates back to the eighteenth century, and has produced countless angiosperm cultivars with varied traits (i.e. color, size, fragrance, longevity, and disease resistance). The continued generation of flower cultivars (e.g. roses, lilies, petunias, tulips, etc) speaks directly to the variation in human preference for flowers and plants. Three consumer segments have been identified with unique desired floral characteristics that motivate buying preferences through consumer-assisted selection methods (Levin et al., 2012). One of these segments is most interested in the floral trait

1
2
3
4 of fragrance. In order to understand and characterize the molecular and biochemical pathways
5
6 leading to floral volatile biosynthesis a model organism amenable to molecular biology studies is
7
8 needed.
9

10
11 Functional research into floral fragrance arguably began with the characterization of gene
12
13 products responsible for the synthesis of floral volatile compounds in the model organism
14
15 *Petunia x hybrida* cv. ‘Mitchell Diploid’ (MD) (Boatright et al., 2004; Negre et al., 2003;
16
17 Underwood et al., 2005; Verdonk et al., 2005). MD is a laboratory generated, homozygous
18
19 diploid species (Mitchell et al., 1980) that emits approximately 12 floral volatile
20
21 benzenoids/phenylpropanoids (FVBPs). Many of the protein encoding genes in the FVBP
22
23 metabolic network have been identified and characterized at some level. Several aspects of this
24
25 research have been translational in nature. For example: identification of the core β -oxidative
26
27 pathway for benzoic acid biosynthesis (Colquhoun et al., 2012; Klempien et al., 2012; Qualley et
28
29 al., 2012; Van Moerkercke et al., 2009), identification of novel enzymatic activities like a
30
31 decarboxylation-amine oxidation reaction (Kaminaga et al., 2006; Tieman et al., 2006), and
32
33 understanding of a complex and coordinated gene network (Colquhoun and Clark, 2011;
34
35 Colquhoun et al., 2010a; Van Moerkercke et al., 2012a; Van Moerkercke et al., 2011, 2012b;
36
37 Verdonk et al., 2005). Much of this research was aided by the relatively simple cellular
38
39 environment of the MD corolla limb tissue. The corolla limb tissue is mostly devoid of pigments,
40
41 only a few cell layers thick, and predominantly expresses the FVBP gene network. However, the
42
43 limited number of volatile compounds produced by MD flowers has confined the ability to
44
45 translate much of this research into other plant products like fruits and vegetables, which can
46
47 emit hundreds of different volatile compounds. Previous analysis of *Lilium* cv. ‘Casa Blanca’
48
49 established the presence of a diversity of volatile compounds synthesized from several
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 independent pathways such as terpenes, phenylpropanoids, and nitrogen-containing compounds
5
6 (Oyama-Okubo et al., 2011). The presence of compounds from diverse metabolic volatile
7
8 pathways positions lilies as a potential model system for the elucidation of gene products
9
10 responsible for the synthesis of volatiles which can build upon the information derived from
11
12 *Petunia x hybrida* cv. ‘Mitchell Diploid.’ Therefore, information derived from the lily system
13
14 could be translational to understanding the synthesis of volatiles from a wider range of important
15
16 agricultural crops such as blueberry (*Vaccinium Cyanococcus*) and strawberry (*Fragaria*).
17
18
19
20

21 Most contemporary *Lilium* flowers are stimulating to the human senses of vision and
22
23 olfaction, and are very popular with the home gardener and people that enjoy plants in general.
24
25 *Lilium* species and cultivars are among the top produced domestic fresh cut flowers and potted
26
27 plants (Vilsack and Clark, 2009). Presumably, due to societal popularity and the genus biology
28
29 itself, *Lilium* species have been a source of interest to humans for at least 3000 years according
30
31 to Minoan civilization paintings like The Prince of Lilies. More recently, the Madonna lily is
32
33 said to be the basis for the fleur-de-lys, which is first historically associated with the French
34
35 monarchy. Today, *Lilium* species and hybrids can be found most anywhere, and are represented
36
37 in many colors, shapes, and sizes.
38
39
40
41
42

43 Despite the overall popularity of *Lilium* plants, consumers have communicated specific
44
45 aspects of some lilies that are undesirable. Many *Lilium* flowers produce excessive amounts of
46
47 floral volatile organic compounds that may be overpowering for some consumers or uses. Few
48
49 scientific studies have addressed the different volatile chemicals that specific lily flowers
50
51 produce and emit, the variation of volatile compounds emitted throughout the *Lilium* genus, or
52
53 even individual volatile biosynthetic pathway elucidations. However, it seems that lily flower
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 fragrance may be a breeding trait of interest if it were quantitative and much more tenable from a
5
6 biochemical, molecular, and genetic level.
7
8

9 In the recent past, research using a pharmacological inhibitor, aminooxy acetic acid
10 (AOA), of a specific plant enzyme, phenylalanine ammonia-lyase (PAL), resulted in a significant
11 reduction of total floral fragrance from an oriental hybrid lily cultivar (Oyama-Okubo et al.,
12 2011). The results of this study were confusing since the pharmacological inhibitors used to
13 reduce floral fragrance emission inhibit the enzymatic ability of a key protein in the
14 phenylpropanoid pathway, but the lily cultivar used for these experiments produced and emitted
15 phenylpropanoids, terpenoids, and nitrogen-containing volatile compounds. Therefore, the
16 mechanism of action would be surprising and has not been elucidated currently. With that said,
17 this work did lay a foundation for what volatile organic compounds can be produced in an
18 oriental hybrid lily flower.
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33 The work presented here focused on the analytical characterization of volatile compounds
34 emitted from *Lilium* flowers. The overall goals of this work were to understand the diversity of
35 floral volatile compounds, quantitate the amounts of floral volatiles, and to begin to address any
36 opportunities for gene function identification.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

2. RESULTS AND DISCUSSION

Numerous oriental hybrid *Lilium* cultivars emit a strong floral fragrance, which can be an overall negative feature for consumers. *Lilium* breeders and industry leaders have communicated the desire to reduce the amount of floral fragrance via molecular breeding efforts. However, relatively little is known about *Lilium* floral fragrance from the perspectives of genetics, biochemistry, and molecular regulation. To begin to explore this issue, four hybrid *Lilium* cultivars (Tarrango, Belladonna, Robina, and Conca d'Or) were chosen with strong floral fragrance and one cultivar (Santander) with a weak floral fragrance (Fig. 1A-E). The qualitative determination of strong and weak fragrance was an anecdotal observation. All plants and flowers were grown, harvested, and shipped using common practices standard to the industry.

2.1 Floral volatile emission is differential between tepals and reproductive organs

The initial testable hypothesis was that the individual floral tissues from a single cultivar (Fig. 1E, F) produce and emit varying amounts of total volatile molecules. Similar to the anatomical separation and volatile collection conducted with petunia flowers (Underwood et al., 2005), the relatively large *Lilium* flowers were excised from the stalk and separated spatially into four tissue types: distal tepals (sepals), proximal tepals (petals), stamens, and carpels. Subsequent to biological staging and anatomical separation, volatile molecules were collected and analyzed from these tissues using a dynamic push/pull head space collection system (Underwood et al., 2005) and gas chromatography mass spectrometry (GC-MS) methods (Schmelz et al., 2001).

Significantly higher amounts of total volatile molecules were detected from tepal tissues compared to reproductive organ (stamen and carpel) tissues (Fig. S1). The reproductive organs of all cultivars emitted relatively low levels of total volatiles ranging from approximately 350 ng*gFW-1*hr-1 emitted by the carpel of 'Conca d'Or' to approximately 11,000 ng*gFW-1*hr-1 emitted by the carpel of 'Tarrango'. Total volatiles from 'Conca d'Or' tepal tissues were detected at approximately 145 times greater than total volatiles from all reproductive organs. Conversely, total volatiles from tepal tissues of 'Santander' were detected at approximately seven times greater levels than total volatiles from the reproductive organ tissues. Total volatile amounts from tepal tissues of all cultivars were relatively large except for 'Santander', which was the lowest floral volatile emitter among the five cultivars examined. Total volatiles from the proximal and distal tepals were detected at similar levels for every cultivar. Combined mean volatile emission from both proximal and distal tepals was greatest in 'Conca d'Or' which had an approximate mean emission of 60,000 ng*gFW-1*hr-1 compared to 'Santander' which had an approximate mean emission of 6,800 ng*gFW-1*hr-1. Our initial hypothesis was partially supported by a large differential in volatile emission between tepal and reproductive organ tissues for all cultivars. However, the proximal and distal tepals from each cultivar emitted similar levels of total volatiles. Additionally, the reproductive organs (stamen and carpel) from each cultivar emitted similar levels of total volatiles (Fig. S1). There are two important points after these observations: 1, the focus of our efforts should be geared towards understanding volatile production from tepal tissues (not reproductive organs) since they are the highest emitter of volatile molecules; and 2, compared to the other cultivars, 'Santander' is quantitatively confirmed as having a weaker floral fragrance, which implies 'Santander' could then be used as

an outgroup in further comparative experiments such as comparative transcriptomic experiments to identify candidate floral volatile biosynthesis/regulation transcripts in the future.

2.2 Chemotypic variations among floral tissues between cultivars

The second testable hypothesis was that each floral tissue emits a different volatile compound profile. A total of 66 individual volatile compounds were identified and quantitated from all *Lilium* tissues during these experiments (Table 1, S1). Among all cultivars, no single tissue emitted all 66 identified compounds. Between 57 and 63 individual volatile compounds were identified and quantitated from tepal tissues of all cultivars except ‘Santander’ (approximately 50 volatile compounds), while 35 to 40 individual volatile compounds were identified and quantitated from reproductive organs of all cultivars (Table 1). A principal components analysis clearly demonstrated that the reproductive organs and tepals have distinct floral volatile profiles (Fig. 2A). Likewise, the floral volatile profile from tepal tissue was cultivar specific (Fig. 2B).

Two-way Ward’s hierarchical cluster analysis was performed on the dataset of 66 individual volatile compounds (Fig. 3). The hierarchical cluster analysis provides a visual representation of relatedness for individual volatile compounds across samples (bottom dendrogram) in addition to the relatedness of all volatile contents between individual samples (right dendrogram). Standardization of metabolite content is by row mean ($n=4$) and standard deviation, with high values represented as red, average as green, and low as violet (spectral scale). It was hypothesized that reproductive organs and tepals would clade out into separate groups based the independent biological roles of each tissue in addition to the results from the PCA, however, the hierarchical clade on the right shows both sets of tepals from ‘Santander’

grouping with reproductive organs. It is possible that the hierarchical cluster analysis groups ‘Santander’ into the reproductive organ clade due to ‘Santander’s’ relatively low number of compounds emitted and low total volatile emission from the tepals. Although reproductive organs from the same cultivar did not group together, each set of tepals did, which shows the presence of a distinct volatile profile for each cultivar. In addition, proximal and distal tepals within a cultivar grouped together and independent of each other cultivar tepal set suggesting little variability between spatial tepal sets within a cultivar but distinct profiles between cultivars. Upon further inspection of the hierarchical cluster analysis, the volatile profiles from the tepals of ‘Belladonna’ and ‘Conca d’Or’ are within a group and separate from ‘Tarrango’ and ‘Robina’ with ‘Robina’ as the most dissimilar within the lower half of the dendrogram group. This observation may suggest that ‘Belladonna’ and ‘Conca d’Or’ have similar but still distinct volatile profiles based on total number and emission rate of emitted compounds. Taken together, results from both the PCA and hierarchical cluster analysis supported the initial and second testable hypotheses of cultivar specific chemical composition of floral fragrance and differential emission of floral volatiles anatomically.

2.3 Cultivar specific chemotypic variations

Chemical classes of vetted volatiles include monoterpene hydrocarbons, monoterpene alcohols and aldehydes, phenylpropanoids, benzenoids, fatty-acid-derived, nitrogen-containing, and amino-acid-derived compounds (Table 1). The two nitrogen containing volatile compounds, indole (**64**) and 5-inandol (**65**), were undetectable from all reproductive organs examined (Table

1
2
3
4 1). *Lilium* cv. ‘Conca d’Or’ accumulated the most indole (**64**) but has the lowest accumulation of
5
6 5-indanol (**65**) unlike ‘Robina’ which has the inverse accumulation profile.
7
8

9 Oriental hybrid ‘Santander’ has the lowest volatile emission, 9.6 % of ‘Conca d’Or’ (Fig.
10
11 4A) due to the near absence of monoterpenoid emission present in the other cultivars (Fig. 4B).
12
13 Oriental-Trumpet hybrids ‘Belladonna’, ‘Conca d’Or’, and ‘Robina’ have the greatest total
14
15 volatile emission (Fig. 4A.), which is comprised mostly of monoterpenoids (Fig. 4B).
16
17 Interestingly, ‘Tarrango’ lacks a substantial amount of monoterpene hydrocarbons (Fig. 4B) due
18
19 to low emission of (Z)- β -ocimene (**09**) (Table 1), but is still comparable in total volatile emission
20
21 due to a great amount of monoterpene alcohols and aldehydes (Fig. 4B). ‘Tarrango’ emits the
22
23 most linalool (**14**), a monoterpene alcohol, at approximately 30,000 ng*gFW-1*hr-1 (Table 1).
24
25

26
27 Emission of benzyl alcohol (**35**), 1,100 ng*gFW-1*hr-1, was only detected in ‘Santander’
28
29 which also had the highest emission of benzaldehyde (**24**), 500 ng*gFW-1*hr-1, but scarce
30
31 emission of methyl benzoate (**25**), 37 ng*gFW-1*hr-1 (Fig. 5, Table 1). In contrast,
32
33 ‘Belladonna’, ‘Conca d’Or’, ‘Robina’, and ‘Tarrango’ emit no detectable level, 1000, 1500, and
34
35 3700 ng*gFW-1*hr-1 of methyl benzoate (**25**) respectively (Fig. 5, Table 1). In the petunia floral
36
37 volatile benzenoid/ phenylpropanoid (FVBP) pathway model, benzaldehyde (**24**) is structured as
38
39 the downstream product of benzyl alcohol (**35**) after a yet uncharacterized set of oxidation
40
41 reactions. Metabolic channeling of benzyl-CoA (**69**) through a protein complex with varying
42
43 levels of activity for each protein in the complex could be responsible for the large variation in
44
45 compounds produced and emitted across cultivars. ‘Belladonna’ was the only cultivar to emit
46
47 detectable levels phenylacetaldehyde (**41**) and had the highest emission of 2-phenylethyl acetate
48
49 (**43**) within the phenylpropanoid-related compound class (Table 1). *Lilium* cv. ‘Belladonna’ has
50
51 the highest emission rate of phenylpropanoids at approximately 5000 ng*gFW-1*hr-1 (Fig. 4C)
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 largely composed of (Z)-isoeugenol (**64**) (Table 1). The large variation in benzenoid and
5
6 phenylpropanoid presence and emission across cultivars suggests potential differential presence
7
8 and regulation of genes which could lead to new insights on the regulation of the floral
9
10 benzenoid/ phenylpropanoid pathway.
11
12

13
14 All of the *Lilium* cultivars studied emit varying levels of C5 and C6 fatty acid derived
15
16 compounds (Fig. 4D); nine compounds were identified across cultivars (Table 1). Three C5 and
17
18 six C6 fatty acid derived green leaf volatiles were identified in tepal tissues across cultivars
19
20 (Table 1). ‘Santander’ has no detectable emission of (E)-2-hexenyl acetate (**59**) but has highest
21
22 emission of (Z)-3-hexenyl acetate (**57**) (Table 1) suggesting differential metabolic activities in
23
24 the green leaf volatile production pathway. ‘Santander’ emits the most C6 volatile compounds
25
26 (Fig. 4D); the two in greatest abundance were hexyl acetate (**58**) and (Z)-3-hexenyl acetate (**57**)
27
28 (Table 1). The latter compound, (Z)-3-hexenyl acetate (**57**), is derived from (Z)-3-hexenal, a so
29
30 called green leaf volatile that has been identified as a volatile constituent for flavor in tomatoes
31
32 (Buttery et al., 1989; Tieman et al., 2012) in addition to emission by many plants in response to
33
34 herbivory (Paré and Tumlinson, 1997). Also detected across cultivars, C4, C15, and C16 fatty
35
36 acid-derived volatile compounds constituted a minor addition to each cultivar’s overall volatile
37
38 profile but provide unique differential emission (Fig. 4D). ‘Tarrango’ has the highest emission of
39
40 C4 compounds (Fig. 4D) consisting of methyl butyrate (**49**) and methyl 2-methylbutyrate (**50**) at
41
42 19 and 26 ng*gFW⁻¹*hr⁻¹ respectively (Table 1). ‘Conca d’Or’s’ emission of C15 and C16
43
44 volatiles are at rates close to the emission of C6 compounds for ‘Belladonna’, ‘Conca d’Or’, and
45
46 ‘Tarrango’ while those cultivars emit very little C15 and C16 compounds (Fig. 4D).
47
48
49
50
51
52
53
54

55 Overall, information from the PCA, hierarchical cluster analysis, compound class
56
57 comparisons and individual volatile compound comparisons demonstrate the uniqueness and
58
59
60
61
62
63
64
65

diversity of each cultivar's volatile profile. Volatile profile uniqueness across cultivars will be advantageous towards identification of a multitude of candidate floral volatile biosynthesis/ regulation transcripts through future comparative transcriptomic experiments.

2.4 Comparative transcript accumulation of select biosynthetic candidates

The third testable hypothesis was that it was possible to identify transcripts that encode proteins involved in the biosynthesis of specific floral volatile compounds by comparing the differential transcript accumulation to the differential volatile compound emission between all cultivars. In an initial effort to identify floral volatile biosynthetic transcripts in *Lilium*, 5' and 3' tagged cDNA libraries were first generated from the tepal tissue of each cultivar. Tepal tissue was used exclusively in order to enrich in the abundance of transcripts from the floral tissue that emitted both the greatest diversity and content of volatile compounds. Primers were designed based on the homology of conserved sequences between various other plant species. The goal was to identify candidate genes in diverse pathways that may be crucial for carbon supply to generate groups of volatile molecules. A total of four transcripts were identified after careful selection, cloning, and sequencing of multiple amplicons. The four transcripts showed high homology to enzymes involved in the synthesis of benzenoid, terpene, terpenoid, and fatty-acid derived volatiles. Quantitative PCR primers were then constructed, vetted, and optimized (Fig. 5, Table S2).

As the last step of the β -oxidation pathway in petunia flowers, 3-ketoacyl-CoA thiolase (KAT) catalyzes the conversion of 3-oxo-3-phenylpropanoyl-CoA to benzyl-CoA (**69**) in the peroxisome organelle (Van Moerkercke et al., 2009). Benzyl-CoA (**69**) can be used as a substrate

1
2
3
4 for many metabolic processes including the synthesis of the volatile benzenoids phenylethyl
5
6 benzoate (39), benzyl benzoate (38), methyl benzoate (25), and benzyl alcohol (35). Cloning
7
8 efforts in *Lilium* cv. 'Robina' resulted in a putative KAT sequence (NCBI: KR998331) with the
9
10 longest open reading frame of 1476 nucleotides corresponding to 481 amino acid residues.
11
12
13
14 Quantitative transcript accumulation using KAT designed primers provided a single melt curve
15
16 for each reaction and revealed highest expression of KAT in *Lilium* cv. 'Belladonna' and
17
18 'Tarrango' with lowest expression in 'Robina' and 'Conca d'Or' (Fig. 5A). Volatile emission of
19
20 benzyl benzoate (38), a direct volatile metabolite of benzyl-CoA (69), was detected at the highest
21
22 level from 'Belladonna' compared to all other cultivars; however, 'Tarrango' emitted the second
23
24 lowest amount of benzyl benzoate (38) (Fig. 6). This may be a consequence of 'Tarrango'
25
26 emitting almost 4000 fold more methyl benzoate (25) than 'Belladonna', which would draw the
27
28 majority of the benzyl-CoA (69) (Colquhoun et al., 2010b; Corea et al., 2012; Langer et al.,
29
30 2014; Negre et al., 2003; Underwood et al., 2005).
31
32
33
34

35
36 Cloning efforts were also directed towards genes involved in the synthesis of the most
37
38 abundant volatile compound class emitted by the lilies studied here: monoterpenoids. Five
39
40 carbon isoprenoids used for the synthesis of terpenes and terpenoids are known to be synthesized
41
42 by two independent pathways: the 2-C-methyl-D-erythritol 4-phosphate (MEP) and mevalonate
43
44 pathways (Beytía and Porter, 1976; Rodríguez-Concepción and Boronat, 2002). In *Arabidopsis*
45
46 *thaliana*, the MEP pathway occurs in the plastid while the mevalonate pathway occurs in the
47
48 peroxisome (Tholl and Lee 2011). The MEP pathway provides isoprene units used for plastidal
49
50 synthesis of monoterpenes, diterpenes, carotenoids, and gibberellins. In the first step of the MEP
51
52 pathway, 1-deoxy-D-xylulose-5-phosphate synthase (DXS) (EC 2.2.1.7) synthesizes 1-deoxy-D-
53
54 xylulose 5-phosphate through a transketolase reaction between pyruvate and glyceraldehyde-3-
55
56
57
58
59
60
61
62
63
64
65

phosphate. There is evidence that DXS plays a role in the regulation of monoterpene synthesis as quantitative trait loci for monoterpene synthesis have been observed in close proximity to putative DXS genes and correlated to the synthesis of multiple monoterpene alcohols in grape (Battilana et al., 2009; Duchêne et al., 2009). DXS synthesizes precursors which can become isoprenoid substrates used to synthesize not only monoterpenes but also diterpenes, sesquiterpenes, chlorophylls, carotenoids, and other isoprenoid derived compounds. Cloning efforts were focused on genes in the MEP pathway since volatile analysis identified 12 monoterpene hydrocarbons and 8 monoterpene alcohols and aldehydes across cultivars. In *Lilium* cv. ‘Belladonna’, a putative DXS sequence (NCBI: KR998332) with a sequence length of 639 nucleotides was cloned. Quantitative transcript accumulation using DXS designed primers revealed three-fold higher expression of DXS in *Lilium* cv. ‘Conca d’Or’ and ‘Santander’ compared to expression in ‘Belladonna’ and ‘Robina’ (Fig. 5B). This is interesting because overall volatile emission of monoterpene hydrocarbons from *Lilium* cv. ‘Conca d’Or’ was 40-fold greater than ‘Santander’ but only two-fold greater than ‘Belladonna’ and ‘Robina’. The quantitative transcript accumulation of the putative *LhDXS* did not qualitatively trend with any of the monoterpene hydrocarbon emission levels of individual cultivars. However, a comparison of individual monoterpene hydrocarbon volatiles between cultivars reveals three compounds, (Z)- β -ocimene (**09**), eucalyptol (**13**), and linalool (**14**), which together constitutes the majority of monoterpene hydrocarbon composition in all cultivars, except for ‘Santander.’ *Lilium* cv. ‘Santander’ emitted higher amounts of non-cyclic monoterpenoids such as, ($\pm$)-citronellal (**15**), (R)-(+)- β -citronellol (**16**), neral (**17**), nerol (**18**), geranial (**19**) and geranyl acetate (**20**) than all other cultivars, suggesting ‘Santander’ may have absent or nonfunctional proteins involved in terpene cyclization. Previous studies have identified species with numerous TPS genes, but only

1
2
3
4 a small fraction of these genes are hypothesized to produce full length functional proteins (Chen
5
6 et al., 2011; Nieuwenhuizen et al., 2013). The author's hypothesize that the potential loss of
7
8 functional TPS genes in 'Santander' during selective breeding could have left monoterpene
9
10 hydrocarbon substrate pools open to oxidation to monoterpene alcohols and aldehydes for
11
12 floral volatile emission.
13
14

15
16 Terpenes synthases (TPS) are a mid-sized family of enzymes that can catalyze
17
18 polymerization and/or cyclization of the following isoprene subunits: isopentenyl diphosphate
19
20 (IPP), geranyl diphosphate (GPP), neryl diphosphate (NPP), farnesyl diphosphate (FPP) and
21
22 geranylgeranyl diphosphate (GGPP). The seven clades of TPS genes are involved in the diverse
23
24 synthesis of monoterpenes, diterpenes, and sesquiterpenes (Chen et al., 2011). Due to differing
25
26 reaction mechanisms and formation products, plant TPSs have been divided into two classes with
27
28 most known plant TPSs classified under class I (Chen et al., 2011). Class I TPSs form highly
29
30 reactive carbocation intermediates which can lead to the formation of multiple products from a
31
32 single substrate (Christianson, 2006; Steele et al., 1998). Sequence information from a previous
33
34 cloning effort in *Lilium* 'Siberia' (NCBI: KF734591) provided insight into cloning efforts. In
35
36 *Lilium* cv. 'Robina', a putative TPS sequence (NCBI: KR998333) with a longest open reading
37
38 frame of 1758 nucleotides corresponding to 586 amino acids, was attained. Quantitative
39
40 transcript accumulation using TPS designed primers revealed two-fold higher expression of TPS
41
42 in *Lilium* cv. 'Belladonna' compared to 'Conca d'Or' in addition to 25-fold higher expression
43
44 compared to 'Robina', 'Tarrango' and 'Santander' (Fig. 5C). A comparison of quantitative
45
46 *LhTPS* transcript accumulation across cultivars to individual terpene volatile emission
47
48 abundances reveals a qualitative trend between *LhTPS* expression and α -terpinene (**06**), a cyclic
49
50 monoterpene hydrocarbon, emission. Although α -terpinene (**06**) emission was between two and
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

four magnitudes below the emission of many other monoterpene hydrocarbons, emission was highest in ‘Belladonna’ at 18 ng*gFW⁻¹*hr⁻¹. These data provide a useful starting point toward understanding the degree of conservation for terpene synthesis and cyclization mechanisms between monocots and dicots.

The oxylipin metabolism pathway produces C₆ volatile compounds, often referred to as green leaf volatiles, through the metabolism of linolenic and linoleic acids. Within the oxylipin metabolism pathway, hydroperoxide lyase (HPL) (EC 4.2.1.92) catalyzes the cleavage of 13-hydroperoxide into (Z)-3-hexenal and 12-oxo-(Z)-9-dodecenoic acid (Vick and Zimmerman 1987). In *Lilium* cv. ‘Santander’ (NCBI: KR998330) and ‘Belladonna’ (NCBI: KR998329) a putative transcript with similar amino acid sequence to characterized hydroperoxide lyases were cloned. Quantitative transcript accumulation using HPL designed primers revealed 3-fold higher expression of HPL in *Lilium* cv. ‘Santander’ compared to ‘Belladonna’ and ‘Robina’ (Fig. 5D). In this study, the *Lilium* cultivars examined emitted six common C₆ green leaf volatiles: hexanal (**54**), (Z)-3-hexen-1-ol (**55**), 1-hexanol (**56**), (Z)-3-hexenyl acetate (**57**), hexyl acetate (**58**), (E)-2-hexenyl acetate (**59**). While the direct product of HPL, (Z)-3-hexenal, was not detected in any *Lilium* cultivar examined here, (Z)-3-hexenyl acetate (**57**), a known end product of oxylipin metabolism via modification of (Z)-3-hexenal was detected and shows a qualitatively similar emission trend with *LhHPL* expression across cultivars. Emission of (Z)-3-hexenyl acetate (**57**) was highest in ‘Santander’ at a rate of approximately 90 ng*gFW⁻¹*hr⁻¹.

Clearly, floral volatile biosynthesis in *Lilium* is much more complex than in more simple systems such as petunia (Colquhoun et al., 2010b). Cloning and assaying transcript accumulation from four floral volatile biosynthetic candidates provided few immediate or obvious trends with floral volatile emission itself. For example, even qualitative comparison of total benzenoid

1
2
3
4 emission downstream of benzyl-CoA (69) does not correlate with KAT transcript accumulation
5
6 across cultivars. This work did provide entry level transcript accumulation information, multiple
7
8 sequence information, and basic cloning tools and techniques that can be applied to *Lilium* going
9
10 forward. Future work involving stringent transcriptomics, soluble metabolomics, volatile
11
12 metabolomics, and correlation analysis (multiple regression/multivariate statistics) will be
13
14 required to identify the biosynthetic players and metabolic pathways leading to the floral
15
16 volatiles emitted by these *Lilium* cultivars.
17
18
19
20
21
22

23 **3 CONCLUSIONS**

24
25
26
27
28 A desire to understand *Lilium* floral fragrance from biochemical, genetic, and molecular
29
30 perspectives is economically and scientifically warranted. The ability of numerous hybrid *Lilium*
31
32 cultivars to emit many different volatile molecules and in large quantities can be a purchasing
33
34 deterrent for various consumers. Using contemporary volatile molecule collection, qualification,
35
36 and quantitation methods; cultivar and tissue specific floral volatile profiles of five hybrid *Lilium*
37
38 cultivars are now identified at high resolution. Thus, the biochemical and genetic
39
40 characterization of the metabolic pathways leading to these volatile molecules can now be
41
42 investigated with purpose. Four fairly well characterized enzymatic steps were simply selected
43
44 from the literature, and based on homology identified candidates associated with benzenoid,
45
46 terpene, terpenoid, and C₆ volatile biosynthesis. However, it is clear that large-scale comparative
47
48 transcriptomics between cultivars would be the most efficient and efficacious means to identify
49
50 biosynthetic and regulator transcripts going forward.
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 In addition, future work to identify specific volatile molecules that have direct effects on
5
6 human orthonasal olfaction and cognition will be of much interest for industry and science alike.
7
8 For example, recent work to identify specific volatile molecules that interact with human
9
10 retronasal olfaction and cognition to create certain aspects of flavor in select fruits has provided
11
12 the concept of sweet flavor volatiles (Schwieterman et al., 2014; Tieman et al., 2012). Certain
13
14 *Lilium* cultivars examined here emit many of these compounds (Fig. S3). Two fundamental
15
16 questions arise: 1, are these shared sweet flavor volatiles responsible for the sometimes referred
17
18 to as a sweet orthonasal smell, and 2, does the *Lilium* flower provide a much easier environment
19
20 to genetically and biochemically characterize the metabolism of these important volatile
21
22 molecules when compared to a perhaps more complex biological system of fruits?
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

4. EXPERIMENTAL

4.1 Plant material

Five interspecific *Lilium* hybrid cultivars were purchased from Oregon Flowers Inc (point of contact: Tyler Meskers). Oriental-trumpet hybrid *Lilium* cultivars included *Lilium* cv. ‘Belladonna,’ *Lilium* cv. ‘Conca d’Or,’ *Lilium* cv. ‘Robina,’ and oriental hybrid *Lilium* cultivars included *Lilium* cv. ‘Santander,’ *Lilium* cv. ‘Tarrango.’ Plants were packaged and shipped according to industry standards. Ten stalks of each cultivar were received containing five to seven flowers per stalk. Flower stalks were submerged in Floralife Chrystal Clear floral solution at a concentration of 16 mL per liter tap water. Flower stems were kept in a climate controlled room at a temperature of 21 °C, light level $10 \text{ } \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, photoperiod from 07:00-19:00 hours, and a relative humidity of $50\% \pm 10\%$. For all cultivars, ripe flower bulbs opened from bottom of the stalk to the top. Cut flowers retained ornamental value for approximately 12 to 14 days after opening.

4.2 Volatile analysis

All volatile emission collections occurred at 18:00 h. Flowers from each cultivar were tagged as flower opening occurred. Flowers were selected for volatile emission collection three days after tagging which corresponded with two days post anther dehiscence. Tepals, stamen, and carpels were removed from vegetative stems using a straight edged razor blade and quickly loaded into glass tubes (2.5 cm i.d., 61 cm long, and 300 mL volume). Tubes were connected to a dynamic

headspace collection system featuring a push-pull method of headspace air movement. At the outlet end of the tubes, volatiles were concentrated on glass columns containing approximately 50 mg HaySep Q 80-100 porous polymer adsorbent (Hayes Separations Inc., Bandera, TX). Collections occurred for one hour at room temperature. Volatiles were eluted from the polymer as described by Schmelz (Schmelz et al., 2001).

Quantification of volatiles from the elution matrix was performed on an Agilent 7890A Series gas chromatograph (GC) (carrier gas; He fixed at 13.56 psi; splitless injector, temperature 220°C, injection volume 2 µl) equipped with a DB-5 column ((5%-phenyl)-methylpolysiloxane, 30m length × 250µm i.d. × 1µm film thickness; Agilent Technologies, Santa Clara, CA, USA). Oven temperatures were programmed from 40°C (0.5 min hold) at 5°C min⁻¹ to 250°C (4 min hold). Signals were captured with a flame ionization detector (FID) at 280°C. Peaks from FID signal were integrated manually with Chemstation B.04.01 software (Agilent Technologies, Santa Clara, CA). Calculation of volatile emission mass (ng*gFW⁻¹*h⁻¹) was based on individual peak area relative to the peak area of an elution standard, nonyl acetate, within each sample and standardized for each samples corresponding biological mass. In addition, dilution curves for volatile standards were run on the GC-FID in duplicate to obtain a response factor for each compound that was used in the calculation of volatile emission mass. GC-Mass Spectrometry (MS) analysis of each elution was performed on an Agilent 6890N GC in tandem with an Agilent 5975 MS (Agilent Technologies, Santa Clara, CA, USA) and retention times are compared with authentic standards (Sigma Aldrich, St Louis, MO, USA) for volatile identification (Schmelz et al., 2001). Authentic volatile standards were obtained for 55 of the 66 compounds peaks identified; compounds where no authentic standard was obtained are marked with an asterisk following the compound name in Table 1. The response factor for the elution

standard was used to calculate volatile mass where authentic standards were unable to be obtained. Chemical Abstract Services (CAS) registry numbers were used to query SciFinder® substances database for associated chemical name in Table 1 and Table S1. The equation used to calculate mass for each individual volatile compound was as follows:

$$\frac{(\text{compound of interest peak area in sample}) * (\text{elution standard laboratory control spike peak area})}{(\text{peak area of elution standard in sample}) * (\text{biological sample mass}) * (\text{compound response factor})}$$

4.3 Tissue collection and RNA isolation

All tissue samples were collected concurrently during each volatile collection at 18:00 h and consisted of two biological replications per cultivar. Flowers from each cultivar were tagged as flower opening occurred followed by tissue collection three days after tagging which corresponded with two days post anther dehiscence. All six tepals from a single flower were excised with a single edged razor blade and pooled together into a single collection vessel which was immediately stored in liq. N₂ followed by long term storage at -80 °C. Samples were ground to a fine powder with a sterilized mortar and pestle that had been pre-chilled and maintained at an appropriate temperature with liq. N₂.

All RNA samples were isolated using a CTAB method from (Chang et al., 1993). All extracted RNA samples were subject to TURBO™ DNA-free™ treatment (Ambion Inc., Austin, TX) followed by on-column purification using the RNeasy® Mini protocol for RNA cleanup (Qiagen, Valencia, CA). Total RNA was then quantified via a NanoDrop™ 2000c spectrophotometer (Thermo Scientific, Wilmington, DE). Dilutions to 50 ng µL⁻¹ were created and kept at -20 °C.

4.4 Cloning and Nucleotide analysis

Lilium nucleic acid transcript sequence with sequence similarity to *P. hybrida* and *A. thaliana* was collected from publicly available databases: TAIR (<https://www.arabidopsis.org/>), NCBI (<http://www.ncbi.nlm.nih.gov/>), and ENA (<http://www.ebi.ac.uk/ena>). The resulting sequence fragments were constructed into contiguous sequences using Vector NTI® Advance 11 software (Invitrogen, Grand Island, NY). SMARTER™ RACE technology was used to clone full and partial length transcripts of *LhKAT*, *LhDXS*, *LhTPS*, and *LhHPL* from multiple lily cultivars via the SMARTER™ RACE cDNA amplification kit (Clontech Laboratories Inc., <http://www.clontech.com>) according to the manufacturer's protocol. DNA primers (Table S2) were used to amplify products through use of Advantage® 2 Polymerase Mix (Clontech Laboratories Inc., <http://www.clontech.com>). Gel extraction of amplicons was carried out using QIAquick™ Gel Extraction kit (<http://qiagen.com>) followed by on-column purification of amplicons using QIAquick™ Spin Columns (<http://qiagen.com>). Each independent amplicon was ligated into a pGEM®-T-easy vector (Promega, Madison, WI) which was then sequentially transformed into One Shot® Mach1™ T1^R chemically competent *E.coli* cells (Invitrogen Corp., <http://invitrogen.com>). Sanger sequencing (Big Dye V1–2; University of Florida Interdisciplinary Center for Biotechnology Research) to 2X coverage was carried out on multiple clones from multiple amplicons that had been screened and vetted for amplicons of appropriate size.

4.5 Transcript accumulation analysis

1
2
3
4
5
6
7 All transcript accumulation analyses were conducted multiple times with multiple biological
8
9 replicates and equivalent results were observed. Gene specific primers for qRT-PCR were
10
11 constructed in Primer Express® software v2.0 (Applied Biosystems, Foster City, CA) for
12
13 transcript accumulation of the following genes: *LhKAT*, *LhDXS*, *LhTPS* and *LhHPL*. Sequence of
14
15 the primers used are located in Table S2. $\Delta\Delta C_t$ Quantitative qRT-PCR was performed and
16
17 analyzed using StepOnePlus™ real-time PCR system (Applied Biosystems, Foster City, CA).
18
19 Power SYBR® Green RNA-to-Ct™ 1 and 2-Step kits (Applied Biosystems, Foster City, CA)
20
21 were used to amplify and detect the products according to protocol by the manufacturer.
22
23
24 Optimization of primers was conducted and demonstrated gene specificity during melt curve
25
26 analysis.
27
28
29
30
31
32

33 **4.6 Statistical analysis**

34
35
36
37

38 Means and standard errors were calculated from all volatile emission collection replicates using
39
40 JMP Pro (Version 10, SAS Institute Inc., Cary, NC, USA). Using JMP, the quantified dataset
41
42 was subjected to two independent principal component analyses each with overlaid data ellipsis
43
44 accounting for 90% of the data probability. Two-way Ward hierarchical cluster analysis of all
45
46 quantified volatile compounds was accomplished in JMP using the default settings.
47
48
49
50
51
52

53 **4.7 GenBank® accession numbers**

54
55
56
57
58
59
60
61
62
63
64
65

LhKAT ('Belladonna'), KR998331; *LhDXS* ('Belladonna'), KR998332; *LhTPS* ('Belladonna'),
KR998333; *LhHPL* ('Belladonna'), KR998329; *LhHPL* ('Santander'), KR998330.

ACKNOWLEDGEMENTS

The authors wish to thank Oregon Flowers Inc. and Gebr. Vletter & Den Haan. The manuscript and the work therein were supported by grants from the U.S. Department of Agriculture-Agricultural Research Service-Floriculture & Nursery Research Initiative (USDA-ARS-FNRI), the American Floral Endowment, and the Fred. C. Gloeckner Foundation.

FIGURE LEGENDS

Figure 1. Lily cultivars and floral anatomy. Type flowers of the five commercial hybrid *Lilium* cultivars analyzed used in this study. Flowers of three oriental-trumpet hybrids, (A) ‘Belladonna’, (B) ‘Conca d’Or’, (C) ‘Robina’, and two oriental hybrids, (D) ‘Santander’, (E) ‘Tarrango’, were studied in a tissue specific manner. Floral tissue specific volatile analysis was accomplished by dissecting the flowers prior to collection into (E) distal and proximal tepals, (F) stamens (anther + filament), and (F) carpels (stigma + style + ovary). Photographs and volatiles were captured two days post anther dehiscence.

Figure 2. Volatile emissions are tissue and cultivar specific. Principal component analysis of (A) mean ($n=4$) volatile compound emissions from floral tissues (carpel - $\circ$, stamen - x, proximal tepal - $\diamond$, distal tepal - +) of the five cultivars (‘Belladonna’ – blue, ‘Conca d’Or’ – red, ‘Robina’ – green, ‘Santander’ – purple, ‘Tarrango’ – orange). Component 1 accounts for 35.2% percent of observed variation among samples, while component 2 accounts for 19%. The separation of tepals from carpels and stamen along component 1 is independent of cultivar, thus it is likely accounting for tissue specific nature of volatile emission. The majority of component 2 variation is observed in the tepal samples. This is a result of cultivar specific tepal emission, which is highly evident given the minimal spread between proximal and distal tepal samples from the same cultivar. Density ellipses of floral tissue (carpel - blue, stamen - purple, proximal tepal – green, distal tepal – red) account for 90% of data probability. Principal component analysis of (B) four proximal tepal biological replicates from the same cultivars (‘Belladonna’ – blue, ‘Conca d’Or’ – red, ‘Robina’ – green, ‘Santander’ – purple, ‘Tarrango’ – orange). Component 1

and component 2 account for 21.8% and 20.5%, respectively; capturing biological variability within and among cultivars. Minimal overlap of cultivar-centric 90% density ellipses indicates a genetic basis of volatile emission variation and limited variation among biological replicates.

Principal component values were estimated using the default method in JMP Pro 10.

Figure 3. Two-way cluster analysis of volatile compounds from lily floral tissues. Two-way Ward hierarchical cluster analysis of mean ($n=4$) content of 66 volatile compounds (top) from floral tissues (carpel, stamen, proximal and distal tepals) of five cultivars ('Belladonna', 'Conca d'Or', 'Robina', 'Santander', and 'Tarrango') (left) was constructed using JMP Pro 10.

Standardization of metabolite content is by row mean and standard deviation, with high values represented as red, average as green, and low as violet (spectral scale). The hierarchy and distance of segments within the bottom dendrogram indicates relatedness of content across samples for single metabolites. Structure of the right dendrogram indicates relatedness of all volatile contents among individual samples.

Figure 4. Volatile profile of lily cultivars. Histograms for all panels represent volatiles emitted by proximal tepals in $\text{ng} \cdot \text{gFW}^{-1} \cdot \text{hr}^{-1}$ for all cultivars ('Belladonna' – B, 'Conca d'Or' – C, 'Robina' – R, 'Santander' – S, 'Tarrango' – T). Oriental-trumpet hybrids 'Belladonna', 'Conca d'Or', and 'Robina' have the greatest total emission (A), which is comprised mostly of monoterpene hydrocarbons and monoterpene alcohols and aldehydes (B). Interestingly, 'Tarrango' lacks a substantial amount of monoterpene hydrocarbons, but is still comparable in total volatile emission due to a greatest amount of monoterpene alcohols and aldehydes (B). 'Belladonna' has the greatest mission of phenylpropanoid and phenylpropanoid-related

compounds while ‘Tarrango’ has the greatest emission of benzenoids (C). (D) Comparison and variation of numerous fatty-acid related compounds across cultivars. Volatiles were captured using a dynamic headspace collection system and compounds were identified with GC-MS and quantitated using GC-FID. Chemical classes determined using biosynthetic pathway databases and literature search. Tukey’s HSD mean comparison was accomplished in JMP 10 across genotypes; different letters indicated significantly different means ($\alpha=0.05$).

Figure 5. Transcript accumulation analysis (one-step qRT-PCR) of putative *LhKAT*, *LhDXS*, *LhTPS* and *LhHPL* transcripts. Floral tepals were harvested by excision from the stalk and reproductive organs two days post anther dehiscence at 18:00 h in conjunction with volatile emission collections. Total RNA (50 ng) was used as the starting template per reaction in all cases. X-axis denotes corresponding lily cultivar in each case while each y-axis represents relative quantity of transcript as a fold change. Histograms are representative of multiple experiments with multiple biological replicates analyzed by the $\Delta\Delta C_t$ method (mean $\pm$ SE; $n=3$). For each histogram, expression between cultivars is relative to expression in *Lilium* cv. ‘Belladonna.’ Tukey’s HSD mean comparison was accomplished in JMP 10 across genotypes; different letters indicated significantly different means ($\alpha=0.05$).

Figure 6. Phenylpropanoid and benzenoid biosynthesis and cultivar emission. Histograms for all panels represent volatiles emitted by proximal tepals in $\text{ng} \cdot \text{gFW}^{-1} \cdot \text{hr}^{-1}$ for all cultivars (‘Belladonna’ – B, ‘Conca d’Or’ – C, ‘Robina’ – R, ‘Santander’ – S, ‘Tarrango’ – T). Graphic representation of chemical structure, biosynthetic pathway, and histogram of quantitated volatile compounds in which an arrow represents one enzymatic reaction. Phenylpropanoids

1
2
3
4 phenylacetaldehyde (41), phenylethyl alcohol (42), and 2-phenylethyl acetate (43) are
5
6 synthesized sequentially from phenylalanine (67). Counter, phenylalanine (67) can be converted
7
8 to (E)-cinnamic acid (68) and imported into the peroxisome which ultimately exports benzyl-
9
10 CoA (69). This activated benzoyl group is substrate to four different reactions which give rise
11
12 directly to phenylethyl benzoate (39), benzyl acetate (36), benzyl benzoate (38), and benzyl
13
14 alcohol (35). Benzyl alcohol (35) is the second precursor to benzyl benzoate (38) synthesis as
15
16 well as an indirect precursor of benzaldehyde (24), methyl salicylate (37), and methyl benzoate
17
18 (25) synthesis. Metabolic flow and bottlenecks are evident at different biosynthetic nodes. For
19
20 instance, all three phenylpropanoids are present in ‘Belladonna’ indicating enzymatic activity
21
22 and substrate buildup at the three nodes. A limiting activity of the third node and high activity of
23
24 the first node results in no phenylacetaldehyde (41) emission, substantial phenylethyl alcohol
25
26 (42) emission, and minimal 2-phenylethyl acetate (43) as observed in ‘Robina’. ‘Santander’ has
27
28 the greatest emission of benzyl alcohol (35) and benzaldehyde (24); a buildup due to non-
29
30 existent methyl salicylate (37), methyl benzoate (25), and benzyl benzoate (38) emission, which
31
32 is observed in all other cultivars.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

TABLES

Table 1. Means and standard errors of volatile replicates for each sample assayed ($n=4$) calculated in JMP. Includes volatile emission in $\text{ng} \cdot \text{gFW}^{-1} \cdot \text{hr}^{-1}$ and corresponding number of volatile compounds within means table. Compounds are separated into compound classes and ordered by GC-FID retention time within each class. Initial identification of compounds was based on GC-MS spectra compared to NIST 11 database search. Tentatively identified compounds were vetted by comparing peak retention times and peak mass spectra to authentic standards run under the same GC method. From left to right data has been arranged into spatial tissue groups: proximal tepals, distal tepals, stamen, and carpel. Chemical abstract service (CAS) numbers for each compound are listed to the left of compound names. Summation of total volatile emission for each cultivar within each tissue type listed in the bottom row of table. Asterisk following compound name denotes authentic standard was not obtained; compound identity based on MS spectra compared to NIST 11 database search only.

REFERENCES

- Andrews, E. S., Theis, N., Adler, L. S., 2007. Pollinator and herbivore attraction to cucurbita floral volatiles. *J Chem Ecol* 33, 1682-1691.
- Battilana J., Costantini L., Emanuelli F., Sevini F., Segala C., Moser S., Velasco R., Versini G., Stella Grando M., 2009. The 1-deoxy-D: -xylulose 5-phosphate synthase gene co-localizes with a major QTL affecting monoterpene content in grapevine. *Theor. Appl. Genet.* 118, 653–669.
- Beytia, E. D., Porter, J. W., 1976. Biochemistry of polyisoprenoid biosynthesis. *Annu. Rev. Biochem.* 45, 113-142.
- Boatright, J., Negre, F., Chen, X., Kish, C. M., Wood, B., Peel, G., Orlova, I., Gang, D., Rhodes, D., Dudareva, N., 2004. Understanding *in vivo* benzenoid metabolism in petunia petal tissue. *Plant Physiol.* 135, 1993-2011.
- Buttery, R.G., Teranishi, R., Flath, R.A., Ling, L.C., 1989. Fresh tomato volatiles-composition and sensory studies. In R Teranishi, RG Buttery, F Shahidi, eds, *Flavour Chemistry: Trends and Developments*. ACS Symposium Series 338. American Chemical Society, Washington, DC, 213–222.
- Chang, S., Puryear, J., Cairney, J., 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11, 113-116.
- Christianson, D. W., 2006. Structural biology and chemistry of the terpenoid cyclases. *Chem. Rev.* 106, 3412-3442.
- Chen, F., Tholl, D., Bohlmann, J., Pichersky, E., 2011. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* 66, 212-229.
- Colquhoun, T. A., Clark, D. G., 2011. Unraveling the regulation of floral fragrance biosynthesis. *Plant Signaling & Behavior* 6, 378-381.
- Colquhoun, T. A., Marciniak, D. M., Wedde, A. E., Kim, J. Y., Schwieterman, M. L., Levin, L. A., Van Moerkercke, A., Schuurink, R. C., Clark, D. G., 2012. A peroxisomally localized acyl-activating enzyme is required for volatile benzenoid formation in a *Petunia x hybrida* cv. 'Mitchell Diploid' flower. *J Exp Bot* 63, 4821-4833.
- Colquhoun, T. A., Schimmel, B. C., Kim, J. Y., Reinhardt, D., Cline, K., Clark, D. G., 2010a. A petunia chorismate mutase specialized for the production of floral volatiles. *Plant J.* 61, 145-155.
- Colquhoun, T. A., Verdonk, J. C., Schimmel, B. C., Tieman, D. M., Underwood, B. A., Clark, D. G., 2010b. Petunia floral volatile benzenoid/phenylpropanoid genes are regulated in a similar manner. *Phytochemistry* 71, 158-167.
- Corea, O. R. A., Ki, C., Cardenas, C. L., Kim, S. J., Brewer, S. E., Patten, A. M., Davin, L. B., Lewis, N. G., 2012. Aroenate dehydratase isoenzymes profoundly and differentially modulate carbon flux into lignins. *J. Biol. Chem.* 287, 11446-11459.

- 1
2
3
4 Duchêne, E., Butterlin, G., Claudel, P., Dumas, V., Jaegli, N., Merdinoglu, D., 2009. A
5 grapevine (*Vitis vinifera* L.) deoxy-D-xylulose synthase gene colocates with a major
6 quantitative trait loci for terpenol content. Theor. Appl. Genet. 118, 541–552.
7
8 Dudareva, N., Klempien, A., Muhlemann, J. K., Kaplan, I., 2013. Biosynthesis, function and
9 metabolic engineering of plant volatile organic compounds. New Phytologist 198, 16-32.
10
11 Gilbert, J. L., Guthart, M. J., Gezan, S. A., Pisaroglo de Carvalho, M., Schwieterman, M. L.,
12 Colquhoun, T. A., Bartoshuk, L. M., Sims, C. A., Clark, D. G., Olmstead, J. W., 2015.
13 Identifying breeding priorities for blueberry flavor using biochemical, sensory, and
14 genotype by environment analyses. PLoS ONE 10, e0138494.
15
16 Goody, J., 1993. The culture of flowers. Cambridge University Press, Cambridge ; New York.
17
18 Kaminaga, Y., Schnepf, J., Peel, G., Kish, C. M., Ben-Nissan, G., Weiss, D., Orlova, I., Lavie,
19 O., Rhodes, D., Wood, K., Porterfield, D. M., Cooper, A. J., Schloss, J. V., Pichersky, E.,
20 Vainstein, A., Dudareva, N., 2006. Plant phenylacetaldehyde synthase is a bifunctional
21 homotetrameric enzyme that catalyzes phenylalanine decarboxylation and oxidation. J.
22 Biol. Chem. 281, 23357-23366.
23
24 Kessler, D., Gase, K., Baldwin, I. T., 2008. Field experiments with transformed plants reveal
25 the sense of floral scents. Science 321, 1200-1202.
26
27 Klahre, U., Gurba, A., Hermann, K., Saxenhofer, M., Bossolini, E., Guerin, P. M., Kuhlmeier,
28 C., 2011. Pollinator choice in petunia depends on two major genetic loci for floral
29 scent production. Curr. Biol. 21, 730-739.
30
31 Klempien, A., Kaminaga, Y., Qualley, A., Nagegowda, D. A., Widhalm, J. R., Orlova, I.,
32 Shasany, A. K., Taguchi, G., Kish, C. M., Cooper, B. R., D'Auria, J. C., Rhodes, D.,
33 Pichersky, E., Dudareva, N., 2012. Contribution of CoA ligases to benzenoid
34 biosynthesis in petunia flowers. Plant Cell 24, 2015-2030.
35
36 Langer, K. M., Jones, C. R., Jaworski, E. A., Rushing, G. V., Kim, J. Y., Clark, D. G.,
37 Colquhoun, T. A., 2014. PhDAHPl is required for floral volatile
38 benzenoid/phenylpropanoid biosynthesis in *Petunia x hybrida* cv 'Mitchell Diploid'.
39 Phytochemistry 103, 22-31.
40
41 Levin, L. A., Langer, K. M., Clark, D. G., Colquhoun, T. A., Callaway, J. L., Moskowitz, H. R.,
42 2012. Using Mind Genomics (R) to identify essential elements of a flower product.
43 Hortscience 47, 1658-1665.
44
45 Mitchell, A. Z., Hanson, M. R., Skvirsky, R. C., Ausubel, F. M., 1980. Anther culture of petunia
46 - genotypes with high-frequency of callus, root, or plantlet formation. Z
47 Pflanzenphysiol 100, 131-146.
48
49 Negre, F., Kish, C. M., Boatright, J., Underwood, B., Shibuya, K., Wagner, C., Clark, D. G.,
50 Dudareva, N., 2003. Regulation of methylbenzoate emission after pollination in
51 snapdragon and petunia flowers. Plant Cell 15, 2992-3006.
52
53 Nieuwenhuizen, N. J., Green, S. A., Chen, X., Bailleul, E. J. D., Matich, A. J., Wang, M. Y.,
54 Atkinson, R. G., 2013. Functional genomics reveals that a compact terpene synthase
55
56
57
58
59
60
61
62
63
64
65

- gene family can account for terpene volatile production in apple. *Plant Physiol.* 161, 787-804.
- Oyama-Okubo, N., Nakayama, M., Ichimura, K., 2011. Control of floral scent emission by inhibitors of phenylalanine ammonia-lyase in cut flower of *Lilium* cv. 'Casa Blanca'. *J Jpn Soc Hortic Sci* 80, 190-199.
- Pare, P. W., Tumlinson, J. H., 1997. De novo biosynthesis of volatiles induced by insect herbivory in cotton plants. *Plant Physiol.* 114, 1161-1167.
- Qualley, A. V., Widhalm, J. R., Adebessin, F., Kish, C. M., Dudareva, N., 2012. Completion of the core beta-oxidative pathway of benzoic acid biosynthesis in plants. *Proc. Natl. Acad. Sci. U. S. A.* 109, 16383-16388.
- Rodriguez-Concepcion, M., Boronat, A., 2002. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol.* 130, 1079-1089.
- Schmelz, E. A., Alborn, H. T., Tumlinson, J. H., 2001. The influence of intact-plant and excised-leaf bioassay designs on volicitin- and jasmonic acid-induced sesquiterpene volatile release in *Zea mays*. *Planta* 214, 171-179.
- Schwieterman, M. L., Colquhoun, T. A., Jaworski, E. A., Bartoshuk, L. M., Gilbert, J. L., Tieman, D. M., Odabasi, A. Z., Moskowitz, H. R., Foltz, K. M., Klee, H. J., Sims, C. A., Whitaker, V. M., Clark, D. G., 2014. Strawberry flavor: diverse chemical compositions, a seasonal influence, and effects on sensory perception. *PLoS ONE* 9, e88446.
- Steele, C.L., Crock, J., Bohlmann, J. and Croteau, R., 1998. Sesquiterpene synthases from grand fir (*Abies grandis*). Comparison of constitutive and wound-induced activities, and cDNA isolation, characterization, and bacterial expression of delta-selinene synthase and gamma-humulene synthase. *J. Biol. Chem.* 273, 2078–2089.
- Strauss, S.Y., Whittall, J.W., 2006. Nonpollinator agents of selection on floral traits. In: Harder LD, Barrett SCH, eds. *Ecology and evolution of flowers*. New York, NY, USA: Oxford University Press, 120–138.
- Tholl, D., Lee, S., 2011. Terpene specialized metabolism in *Arabidopsis thaliana*. The *Arabidopsis* book / American Society of Plant Biologists 9, e0143.
- Tieman, D., Bliss, P., McIntyre, L. M., Blandon-Ubeda, A., Bies, D., Odabasi, A. Z., Rodriguez, G. R., van der Knaap, E., Taylor, M. G., Goulet, C., Mageroy, M. H., Snyder, D. J., Colquhoun, T., Moskowitz, H., Clark, D. G., Sims, C., Bartoshuk, L., Klee, H. J., 2012. The chemical interactions underlying tomato flavor preferences. *Curr. Biol.* 22, 1035-1039.
- Tieman, D. M., Taylor, M., Schauer, N., Fernie, A. R., Hanson, A. D., Klee, H. J., 2006. Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8287-8292.
- Underwood, B. A., Tieman, D. M., Shibuya, K., Dexter, R. J., Loucas, H. M., Simkin, A. J., Sims, C. A., Schmelz, E. A., Klee, H. J., Clark, D. G., 2005. Ethylene-regulated floral volatile synthesis in petunia corollas. *Plant Physiol.* 138, 255-266.

- 1
2
3
4 Van Moerkercke, A., Galvan-Ampudia, C. S., Verdonk, J. C., Haring, M. A., Schuurink, R. C.,
5 2012a. Regulators of floral fragrance production and their target genes in petunia are not
6 exclusively active in the epidermal cells of petals. *J Exp Bot* 63, 3157-3171.
7
8 Van Moerkercke, A., Haring, M. A., Schuurink, R. C., 2011. The transcription factor EMISSION
9 OF BENZENOIDS II activates the MYB ODORANT1 promoter at a MYB binding site
10 specific for fragrant petunias. *Plant J.* 67, 917-928.
11
12 Van Moerkercke, A., Haring, M. A., Schuurink, R. C., 2012b. A model for combinatorial
13 regulation of the petunia R2R3-MYB transcription factor ODORANT1. *Plant signaling*
14 & behavior 7, 518-520.
15
16 Van Moerkercke, A., Schauvinhold, I., Pichersky, E., Haring, M. A., Schuurink, R. C., 2009. A
17 plant thiolase involved in benzoic acid biosynthesis and volatile benzenoid production.
18 *Plant J.* 60, 292-302.
19
20 Verdonk, J. C., Haring, M. A., van Tunen, A. J., Schuurink, R. C., 2005. ODORANT1 regulates
21 fragrance biosynthesis in petunia flowers. *Plant Cell* 17, 1612-1624.
22
23 Vick, B. A., Zimmerman, D. C., 1987. Pathways of fatty acid hydroperoxide metabolism in
24 spinach leaf chloroplasts. *Plant Physiol.* 85, 1073-1078.
25
26 Vilsack, T., Clark, C. Z. F., 2009. US Census of Agriculture, 2007. United States Department of
27 Agriculture 1, 51.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1

Compound class	CAS	Compound name ^a		
			Belladonna	
Monoterpenoids (Monoterpene hydrocarbons C10)	99-83-2	α -PHELLANDRENE (01)	0.71 $\pm$	0.61
	7785-70-8	(+)- α -PINENE (02)	12.11 $\pm$	6.21
	79-92-5	CAMPHERE (03)	0.00 $\pm$	0.00
	555-10-2	β -PHELLANDRENE (04)*	9.60 $\pm$	9.60
	123-35-3 & 19902-08-0	MYRCENE & β -PINENE (05)	392.91 $\pm$	75.44
	99-86-5	α -TERPINENE (06)	18.14 $\pm$	5.07
	3779-61-1 & 5989-27-5	(<i>E</i>)- β -OCIMENE & LIMONENE (07)	200.64 $\pm$	37.63
	555-10-12	β -PHELLANDRENE 2 (08)*	1.95 $\pm$	1.95
	3338-55-4	(<i>Z</i>)- β -OCIMENE (09)	20125.76 $\pm$	3977.25
	99-85-4	γ -TERPINENE (10)	1.25 $\pm$	1.02
	460-01-5	COSMENE (11)*	38.28 $\pm$	8.28
	57396-75-5	3,4-DIMETHYL-2,4,6-OCTATRIENE (12)*	1.42 $\pm$	0.30
Monoterpenoids (Monoterpene alcohols & aldehydes C10)	470-82-6	EUCALYPTOL (13)	0.00 $\pm$	0.00
	78-70-6	LINALOOL (14)	21621.60 $\pm$	4302.16
	106-23-0	($\pm$)-CITRONELLAL (15)	28.94 $\pm$	8.01
	106-22-9	(<i>R</i>)-(+)- β -CITRONELLOL (16)	0.24 $\pm$	0.24
	106-26-3	NERAL (17)	7.30 $\pm$	1.38
	106-25-2	NEROL (18)*	97.25 $\pm$	26.27
	141-27-5	GERANIAL (19)	17.43 $\pm$	6.20
	105-87-3	GERANYL ACETATE (20)	15.70 $\pm$	4.51
Sesquiterpenes (C15)	502-61-4	α -FARNESENE (21)*	26.91 $\pm$	10.10
Terpenoid derivatives	110-93-0	6-METHYL-5-HEPTEN-2-ONE (22)	19.71 $\pm$	3.64
	689-67-8	GERANYL ACETONE (23)	8.63 $\pm$	5.60
Benzenoids (C6-C1)	100-52-7	BENZALDEHYDE (24)	176.66 $\pm$	53.99
	93-58-3	METHYL BENZOATE (25)	0.00 $\pm$	0.00
	121-33-5	VANILLIN (26)	3.95 $\pm$	0.73
	106-44-5	<i>p</i> -CRESOL (27)	9.17 $\pm$	7.10
	93-89-0	ETHYL BENZOATE (28)	119.75 $\pm$	68.54
	93-51-6	2-METHOXY-4-METHYLPHENOL (29)	276.83 $\pm$	76.84
	939-48-0	ISOPROPYL BENZOATE (30)*	9.52 $\pm$	5.14
	4463-33-6	2,3-DIMETHOXYTOLUENE (31)	396.93 $\pm$	30.92
	2315-68-6	PROPYL BENZOATE (32)	2.64 $\pm$	0.28
	25152-85-6	(<i>Z</i>)-3-HEXENYL BENZOATE (33)*	7.07 $\pm$	1.96
	6789-88-4	HEXYL BENZOATE (34)*	3.39 $\pm$	0.87
	100-51-6	BENZYL ALCOHOL (35)	0.00 $\pm$	0.00
	140-11-4	BENZYL ACETATE (36)	57.00 $\pm$	13.37
	119-36-8	METHYL SALICYLATE (37)	17.23 $\pm$	5.39
	120-51-4	BENZYL BENZOATE (38)	1035.36 $\pm$	321.40
	94-47-3	PHENYLETHYL BENZOATE (39)	49.31 $\pm$	29.09
	118-58-1	BENZYL SALICYLATE (40)	87.81 $\pm$	17.68
Phenylpropanoid-related (C6-C2)	122-78-1	PHENYLACETALDEHYDE (41)	66.17 $\pm$	48.75
	60-12-8	PHENYLETHYL ALCOHOL (42)	16.44 $\pm$	8.47
	103-45-7	2-PHENYLETHYL ACETATE (43)	793.01 $\pm$	206.40
Phenylpropanoids (C6-C3)	97-53-0	EUGENOL (44)	94.91 $\pm$	26.44
	1754-62-7	(<i>E</i>)-METHYL CINNAMATE (45)	2.26 $\pm$	1.18
	5932-68-3	(<i>E</i>)-ISOEUGENOL (46)	6.90 $\pm$	1.64
	97-54-1	(<i>Z</i>)-ISOEUGENOL (47)	4811.47 $\pm$	1049.98
	93-15-2	METHYL EUGENOL (48)	4.41 $\pm$	1.03
Fatty acid derived (C4)	623-42-7	METHYL BUTYRATE (49)	0.64 $\pm$	0.49
	53955-81-0	METHYL 2-METHYLBUTYRATE (50)*	0.71 $\pm$	0.71
Fatty acid derived (C5)	616-25-1	1-PENTEN-3-OL (51)	0.56 $\pm$	0.29
	71-41-0	1-PENTANOL (52)	0.07 $\pm$	0.07
	628-63-7	AMYL ACETATE (53)	0.82 $\pm$	0.21
Fatty acid derived (C6)	66-25-1	HEXANAL (54)	5.67 $\pm$	2.76
	928-96-1	(<i>Z</i>)-3-HEXEN-1-OL (55)	2.00 $\pm$	0.67
	111-27-3	1-HEXANOL (56)	5.95 $\pm$	1.08
	3681-71-8	(<i>Z</i>)-3-HEXENYL ACETATE (57)	19.03 $\pm$	4.80
	142-92-7	HEXYL ACETATE (58)	42.29 $\pm$	5.99
	2497-18-9	(<i>E</i>)-2-HEXENYL ACETATE (59)	4.25 $\pm$	2.48
Fatty acid derived (C15)	13360-61-7	1-PENTADECENE (60)	5.11 $\pm$	1.36
	629-62-9	PENTADECANE (61)	10.37 $\pm$	7.11
Fatty acid derived (C16)	53939-28-9	(<i>Z</i>)-11-HEXADECENAL (62)	13.75 $\pm$	7.99
	629-80-1	HEXADECANAL (63)*	4.40 $\pm$	2.45
Nitrogen-containing	120-72-9	INDOLE (64)	58.98 $\pm$	17.24
	1470-94-6	5-INDANOL (65)	83.34 $\pm$	22.78
Amino acid derived	590-86-3	ISOVALERALDEHYDE (66)	2.55 $\pm$	0.37
Total			50955.16 $\pm$	10547.51

Figure 1 black and white

Figure 1 color

Figure 2 black and white

Figure 2 color

Figure 3 black and white

Belladonna
Conca d'Or
Robina
Santander
Tarrango
Santander
Conca d'Or
Robina
Belladonna
Tarrango
Carpel
Proximal Tepals
Santander
Santander
Belladonna
Belladonna
Conca d'Or
Distal Tepals
Proximal Tepals
Distal Tepals
Proximal Tepals
Distal Tepals
Tarrango
Robina
Robina

Figure 3 color

Figure 4 black and white

Figure 4 color

Figure 5 black and white

Figure 6 black and white