Identification, Characterization, and Functional Analysis of Terpenoid Specialized Metabolism in Switchgrass (*Panicum virgatum*) and Carrot (*Daucus carota*)

Andrew Joseph Muchlinski

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Biological Sciences

Dorothea Tholl, Committee Chair

Brian Badgley
David Haak
Birgit Scharf
Bingyu Zhao

May10th 2019
Blacksburg, Virginia

Keywords: terpene, chemical defenses, roots, plant-microbe interactions

Identification, Characterization, and Functional Analysis of Terpenoid Specialized Metabolism in Switchgrass (*Panicum virgatum*) and Carrot (*Daucus carota*)

Andrew Joseph Muchlinski

ABSTRACT

Plants produce a large number of specialized or secondary compounds that aid in their reproduction and protection against biotic and abiotic stress. In this work I investigated the metabolism and function of terpenes, the largest class of specialized metabolites, in switchgrass and carrot. Switchgrass (*Panicum virgatum* L.), a perennial C4 grass of the Tallgrass Prairie, represents an important species in natural and anthropogenic grasslands of North America. Its natural resilience to abiotic and biotic stress has made switchgrass a preferred bioenergy crop. I have investigated the metabolism of terpenes in switchgrass leaves and roots in response to herbivory or defense hormone treatments and the application of drought. With a focus on volatile terpene metabolites, I functionally characterized over thirty genes (terpene synthases, TPSs), of which one third could be correlated with the production and release of volatile monoterpenes and sesquiterpenes that likely function in direct chemical defense or in the attraction of insect predators or parasitoids. Drought stress application caused switchgrass roots to accumulate a larger amount of oxygenated terpenes and presumably non-volatile terpenes, the function of which in direct or indirect drought stress protection requires further investigation. I also examined the metabolic dynamics and role of the monoterpene borneol, which accumulates at high concentrations in the roots of switchgrass and to a lower extent in the roots of the close relative *Setaria viridis,* in root microbe interactions. Although we demonstrated a successful RNAi based knock down of the borneol terpene synthase TPS04, we found no immediate evidence that borneol significantly modifies bacterial communities in the root. Further studies on *Setaria* and equivalent RNAi lines

in switchgrass will provide more detailed and needed insight to decipher the role of monoterpene accumulation in grasses interactions with mutualists, pathogens, and pests.

In an applied project, I investigated terpene specialized metabolism in carrot (*Daucus carota L.*) to identify genetic determinants of carrot aroma and flavor. To determine central enzymes which contribute to the terpene component of carrot volatile blends, we first analyzed tissue specific expression patterns of carrot terpene synthase genes (*TPS*) in the genomic model carrot (cv. DH1) and in roots of four aromatically unique colored carrot genotypes (orange-4943B, red-R6637, yellow-Y9244A and purple-P7262). We selected nineteen key biosynthetic enzymes involved in terpene formation and compared in vitro products from recombinant proteins with native volatile profiles obtained from DH1 and colored carrot genotypes. We biochemically characterized several highly expressed TPSs with direct correlations to major compounds of carrot flavor and aroma including germacrene-D (*Dc*TPS11), (*Dc*TPS30) and α-terpinolene (*Dc*TPS03). Random forest analysis of colored carrot volatiles revealed that nine terpene compounds are sufficient for distinguishing the flavor and aroma of raw colored carrots. Interestingly, accumulation of specific terpene compounds rather than chemical diversity is responsible for differences in sensory quality traits in colored genotypes. As accumulations of specific terpene compounds can contribute to the undesired flavor in carrot, our report provides a detailed roadmap for future breeding efforts to enhance carrot flavor and aroma.

Identification, Characterization, and Functional Analysis of Terpenoid Specialized Metabolism in Switchgrass (*Panicum virgatum*) and Carrot (*Daucus carota*)

Andrew Joseph Muchlinski

GENERAL AUDIENCE ABSTRACT

Plants produce a large number of chemicals that are important for growth, defense, flavor, and aroma. While chemical production has been studied in some major food crops (corn, tomato, rice), knowledge of the formation and function of chemicals in switchgrass and carrot is still limited. Switchgrass (*Panicum virgatum* L.), a grass of the Tallgrass Prairie, represents an important species grasslands of North America. Its natural resilience to stress has made switchgrass a preferred bioenergy crop. I found that switchgrass produces many compounds in the chemical class of terpenoids in roots and leaves that likely serve as a defense against damage from pests. In addition, I found that drought stress leads to the production of terpenoid compounds that may have roles in protection when water is limited. My research also demonstrates that roots of switchgrass and the related grass *Setaria* maintain substantial levels of the essential oil compound borneol. This terpenoid compound can act as a nutrient source for specific bacteria and/or an antimicrobial agent. Therefore, I proposed that switchgrass and *Setaria* roots produce borneol to establish a distinct root microbiome by recruitment of beneficial bacteria and deterrence of harmful microorganisms. To test this hypothesis, we genetically engineered plants to reduce borneol formation and accumulation in roots. Using these plants, we evaluated changes in the root microbiome in response to altered borneol levels. We found that interfering with borneol production in *Setaria* roots has limited influence on the microbiome inside roots. Although a similar approach was used for switchgrass, we were unable to significantly reduce borneol

formation in roots. Results from this study provide a better understanding of belowground plant-microbe interactions, and potential for enhancing resistance traits into other crop species.

I also investigated the flavor and aroma compounds produced in carrots, which are considered a key supplemental vegetable due to high nutritional value and pleasant taste. Surprisingly, little has been known about the genetic factors that control flavor and aroma traits in colored carrot varieties. Therefore, I performed a robust characterization of the biosynthesis of terpenoids, which are the predominant aroma and flavor compounds in carrot. I identified several enzymes in carrot that can produce a diverse blend of terpenoids which are associated with sweet, spicy and bitter tastes. In addition, I discovered that carrot stems and leaves also maintain a rich chemistry of terpenoids similar to that in roots. Results from this work provide a baseline for engineering enhanced flavor in carrot and provide a deeper insight into essential oil formation in root crops.

DEDICATION

I dedicate this dissertation in memory of my grandfather

Charles Jassen

July 6, 1918 – December 29, 2013

ACKNOWLEDGEMENTS

Throughout my academic journey, I have received an incredible amount of support and encouragement from family, friends, colleagues and mentors. I must first express my gratitude to Dr. Dorothea Tholl for her strong support over the course of my Ph.D., and for helping me to become a better scientist. My excitement and passion for science has only grown over the last five years, and I owe a great deal of that to Dr. Tholl for her encouragement and support. She has been an incredible advocate for my current and continued success, and provided me with many opportunities to learn and grow.

I am incredibly thankful to my Ph.D. committee members, Drs. Brian Badgley, David Haak, Birgit Scharf and Bingyu Zhao, who have both collectivity and individually provided continued support and guidance throughout my time at VT. I would also like to thank current and former members of their labs for all of their help and training.

I would like to especially thank Dr. James Tokuhisa for his encouragement and support. His mentorship has been an incredible asset and has allowed me to reach my full potential at VT. I would also like to thank current and former members of the Tholl laboratory, especially Dr. Jason Lancaster, who provided me with early mentorship and friendship which really helped me when I first joined the lab.

I would like to thank the Biological Sciences Department and the Translational Plant Sciences program which have provided me with a great community as well as financial support to attend international conferences.

I would like to thank Drs. Eric Hogan and Rich Walker for giving me the opportunity to teach my own class and helping me be successful at it. I would also like to thank Wooram Lee for his guidance and helping me develop and improve my teaching abilities.

I must also give a special acknowledgment to my family for their incredible support throughout my entire academic journey. Without their support this would not have been possible.

TABLE OF CONTENTS

ABSTRACT (ACADEMIC)……..…….……II-III

ABSTRACT (PUBLIC)……..……IV-V

DEDICATION……..……..VI

ACKNOWLEDGEMENTS…….……..VII-VIII

TABLE OF CONTENTS……...IX-XI

LIST OF FIGURES AND TABLES…….XII-XIII

CHAPTER **PAGE**

1. INTRODUCTION AND OVERVIEW OF RESEARCH

PLANT VOLATILE ORGANIC COMPOUNDS: BACKGROUND, FUNCTION AND APPLICATION……..…..1-3
TERPENOIDS METABOLISM AND STRESS RESPONSE…….……..3-11
PLANT MICROBE INTERACTIONS IN THE RHIZOSPHERE…….……..11-14
SWITCHGRASS AS A MODEL STUDY SYSTEM……...15-17
CARROT AS A MODEL STUDY SYSTEM…….....17-19
OVERVIEW OF RESEARCH…….19-21
REFERENCES…….21-27

2. BIOSYNTHESIS AND EMISSION OF STRESS-INDUCED VOLATILE TERPENES IN ROOTS AND LEAVES OF SWITCHGRASS

ABSTRACT…….29-30
INTRODUCTION……30-32
RESULTS
Emission of volatile terpenes from leaves in response to insect feeding ….....……...32-35
Emission of volatile terpenes from roots and leaves upon belowground treatment with methyl jasmonate or salicylic acid……..35-37
Genome-wide Identification of putative terpene synthases in switchgrass..…..…….37-40
Biochemical characterization of mono- and sesquiterpene synthases from switchgrass……..…….40-46
Expression analysis of PvTPSs following treatment with FAW, MeJA and SA…….....47-50
DISCUSSION……...51-56
METHODS……...…56-60
ACKNOWLEDGMENTS…….……60
REFERENCES…….…60-65
SUPPLEMENTAL MATERIALS……...65-67

3. BIOSYNTHESIS AND FUNCTION OF VOLATILE TERPENOIDS IN ROOT-MICROBE INTERACTIONS

INTRODUCTION……68-70
RESULTS
Borneol and camphor accumulation is conserved across upland and lowland switchgrass ecotypes……71-72
PvTPS04 is expressed in leaf tissues……72-74
*Pv*TPS04 *produces (-)-borneol in leaf tissues*……74-75
Salicylic acid modulates borneol and camphor levels in switchgrass roots……75-76
Switchgrass rhizosphere microorganisms alter terpenoid levels in roots.....76-79
The switchgrass root endosphere is enriched with Proteobacteria……80-83
Isolation of root endophytic bacteria on monoterpene containing medium……83-86
TPS04 RNAi plants produce lower borneol levels compared to wild-type.....86-89
DISCUSSION……90-93
METHODS……93-99
REFERENCES……99-104
SUPPLEMENTARY MATERIALS……105-108

4. CHARACTERIZATION OF TERPENE SYNTHASES INVOLVED IN PRODUCTION OF MAJOR AROMA AND FLAVOR COMPOUNDS IN CARROT CULTIVARS

ABSTRACT……110-111
INTRODUCTION……111-115
RESULTS
Analysis of terpene volatiles in leaves, petioles and roots……115-116
Identification of TPS gene models in the carrot genome……116-120
Gene candidate selection……121
Characterization of TPS-a clade genes……121-123
Characterization of TPS-b clade genes…… 123-125
DcTPS19 and DcTPS23 are members of the TPS-g subfamily……125-127
DcTPS25 belongs to the TPS-c clade……128-129
DcTPS28 in an ent-kaurene synthase in the TPS-e/f subfamily……129
Colored carrot varieties have distinct volatile profiles……129-131
DISCUSSION……132-134
MATERIALS AND METHODS……134-141
REFERENCES……138-140
SUPPLEMENTAL MATERIALS……144-161

5. FINAL DISCUSSION AND FUTURE PERSPECTIVE……162-179

APPENDIX A: EFFECTS OF DROUGHT STRESS ON VOLATILE TERPENOID PRODUCTION IN SWITCHGRASS ROOTS……180-194

APPENDIX B: BIOSYNTHESIS OF BORNEOL AND EXAMINATION OF ITS EFFECT ON THE ROOT MICROBIAL COMMUNITY OF *SETARIA VIRIDIS* …….195-222

APPENDIX C: QIIME STANDARD OPERATING PROCEDURE FOR 16S DATA ANALYSIS……..……...….223-225

LIST OF MAIN CHAPTER FIGURES AND TABLES

FIGURE/TABLE **PAGE**

Chapter 1

Figure 1.1 Terpene biosynthesis through the MVA and MEP pathways……...5
Figure 1.2 Biosynthesis of terpene compounds by TPSs……..7-8
Figure 1.3 Plant-microbe interactions in the rhizosphere……10
Figure 1.4 Phylogeny of putative full-length TPSs from seven sequenced plant genomes…….14
Figure 1.5 Root chemicals in belowground interactions……16
Figure 1.6 Major carrot root terpenes……19

Chapter 2

Figure 2.1 GC-MS analysis of volatile emission from treated switchgrass leaves…….34
Figure 2.2 GC-MS analysis of root sesquiterpene emission following MeJA treatment……37
Figure 2.3 Phylogenetic analysis of TPS proteins……..42
Figure 2.4 Relative locations of *TPSs* on the pseudochromosomes of switchgrass……43
Figure 2.5 Hierarchical cluster analysis of *TPS* expression……44
Figure 2.6 SPME-GC-MS analysis of volatile products produced by recombinant TPSs with (*E*,*E*)-FPP as a substrate……....……...45
Figure 2.7 SPME-GC-MS analysis of volatile products produced by truncated recombinant TPSs with GPP as a substrate…….....…….46
Figure 2.8 Gene expression analysis of selected *TPS* genes in switchgrass……..50

Table 2.1 Volatile emission from switchgrass leaves of switchgrass treated with SA, FAW and MeJA…….……..35
Table 2.2 Identified mono- and sesqui*TPS* gene models in the switchgrass genome…….....39

Chapter 3

Figure 3.1 GC-MS analysis of volatile emission switchgrass roots……72
Figure 3.2 Expression analysis of *PvTPS04* across switchgrass tissues……..73-74
Figure 3.3 Chiral GC/MS analysis of (-) borneol produced in switchgrass roots……..75
Figure 3.4 Effect of salicylic acid (SA) treatment on borneol and camphor levels……78
Figure 3.5 Effect of microbial inoculation treatment on borneol and camphor levels…..……..79
Figure 3.6 Relative abundance of phyla across switchgrass root zones……82
Figure 3.7 Principle coordinate analysis of soil grown switchgrass plants……82
Figure 3.8 Growth curve analysis of endosphere isolated bacteria……..85-86

Figure 3.9 Analysis of *PvTPS04* expression and borneol production in wild-type and transgenic switchgrass……88

Table 3.1 Analysis of indicator species enriched in the endosphere compartment of switchgrass roots…….....……..83

Chapter 4

Figure 4.1 GC-MS analysis of hexane extracts from 11-week old carrot (cv. DH1) leaves, petioles and root……...116
Figure 4.2 Maximum-likelihood phylogenetic tree of predicted TPSs……...…….120
Figure 4.3 Analysis of gene expression by RT-qPCR for selected gene candidates…...…….123
Figure 4.4 SPME-GC-MS analysis of the products formed in vitro by partially purified TPS enzymes…….....…127
Figure 4.5 Functional characterization of recombinant DcTPS25 and DcTPS28……..…….....128
Figure 4.6 Multidimensional scaling plot root terpene volatiles……..……..…….131
Figure 4.7 Comparison of relative terpene levels .....…….131

Table 4.1 Characteristics of the 43 terpene synthase genes analyzed in this study……118

CHAPTER I. INTRODUCTION AND OVERVIEW OF RESEARCH

Plant Volatile Organic Compounds: Background, Function and Application

Plants release large quantities of carbon through the emission of volatile organic molecules (Penuelas et al., 2014). Biogenic volatile organic compounds (VOCs) represent a diverse class of metabolites consisting of terpenoids, phenylpropanoids, and fatty acid/amino acid derivatives (Dudareva et al., 2006). These compounds may be synthesized *de novo* under induced conditions, or stored in specialized structures to be released in response to specific stimuli (Gershenzon, 1994). Increased biosynthesis and release of VOCs has been generally associated with defense responses against herbivores and pathogens, as well as protection against abiotic stressors such as high temperature, high light, and drought (Loreto et al., 1998; Vaughan et al., 2015). Additional functions associated with the release of VOCs are the attraction of pollinators and parasitoids of herbivores, and the recruitment of beneficial microorganisms (Cheng et al., 2007; Junker et al., 2011; Huang et al., 2012; Chen et al., 2018). As an example for the use of VOCs by microbes, emissions of methanol from leaves during plant cell wall remodeling provide a carbon source for *Methylobacterium*, which is found at high densities in the phyllosphere (Sy et al., 2005). As methylotrophic bacteria can promote plant growth and induce systemic resistance (Madhaiyan et al., 2004), VOCs likely serve as a type of chemical currency mediating plant-microbe interactions.

While many studies have investigated VOC production in aboveground tissues, much less is known about the biosynthesis and function of VOCs in belowground tissues (Junker and Tholl, 2013). Most investigations have occurred across major model crops i.e. maize and rice, revealing a large chemical and biosynthetic diversity. Increased gene expression and resulting emission of blends of 15-carbon sesquiterpene VOCs from maize roots, for example, have been shown to occur in response to herbivore damage by the Western corn root worm, a major agricultural pest in North

America (Degenhardt et al., 2009). Although likely involved in direct defense, specific VOCs from this induced blend, i.e. (*E*)-beta-caryophyllene, have also been shown to promote tritrophic interactions via the recruitment of entomopathogenic nematodes as natural enemies of maize root herbivores (Kollner et al., 2008). Furthermore, emissions of phenylpropanoids from roots of milkweed and grape have been shown to be induced as a direct defense against insect feeding (Lawo et al., 2011). More recently, investigations of the complex blend of metabolites in dandelion latex identified a semi-volatile sesquiterpene lactone responsible for protecting dandelion roots against larvae from cockchafer - its natural herbivore (Huber et al., 2016). Collectively, these studies highlight conserved functions of VOCs in belowground interactions; however, determinations of exact biological functions of VOCs remains challenging due to measuring VOCs in soil, limitations in transformation capabilities, genomic resources and overall lack of clear phenotype from genetic manipulations of secondary metabolic pathways (Nascimento and Fett-Neto, 2010).

Because of their defensive functions, VOCs, especially those that accumulate as essential oils in tissues of aromatic plants in the Lamiaceae and Apiaceae families, have been used by humans for a long time as insect repellents and antimicrobial agents (Norton and Hamm, 2004; Julsing et al., 2006; Tassou et al., 2012; Pichersky and Raguso, 2018). In addition, VOCs represent major flavor and aroma compounds for a large number of plant-derived foods (Daferera et al., 2002). Many herbs (e.g. dill, lemongrass) and spices (e.g. saffron), contain blends of volatile terpenoids which are the major contributors to their flavor and aroma components (Stewart, 2013). In addition, much of the flavor attribute to alcoholic drinks (e.g. wine) is derived from the presence of terpenes (Stewart, 2013). Storage root and rhizome producing crops such as carrot and ginger accumulate high levels of VOCs, which directly contribute to the aroma and flavor of these crops

(Alasalvar et al., 2001; Shirin and Prakash, 2010). As an example investigated in this study, the aroma of carrot roots is largely determined by blends of volatile terpenes (Fukuda et al., 2013). Changes in environmental conditions can also affect the accumulation of VOCs, but also lead to an increased harsh taste modifying taste perception and overall palatability (Alasalvar et al., 2001; Rosenfeld et al., 2004). Despite the importance of VOCs for the aroma of root crops, knowledge on the biosynthesis of volatiles has been surprisingly limited and mostly restricted to transcriptomic analyses (Jiang et al., 2017). However, recent sequencing of the carrot genome now allows for deciphering the genetic determinants of terpenes as major carrot aroma and flavor components (Iorizzo et al., 2016; Keilwagen et al., 2017). Since terpenes represent the largest group of plant VOCs, I next present a condensed review of terpenoid metabolism.

Terpenoid Metabolism and Stress Response

Compounds in the class of terpenoids or isoprenoids have diverse roles in both primary and secondary metabolism. Terpenoids typically represent oxygenated compounds, while the term "terpenes", in a narrower sense, refers to non-oxygenated hydrocarbons or olefins. However, both terms are often used interchangeably and will both appear in this dissertation. Primary terpene metabolites are required for normal plant growth and development, and include defined roles in steroid formation, phytohormone precursors and molecular side chains via prenylation reactions e.g. chlorophyll (Fig. 1; Mcgarvey and Croteau, 1995). Secondary or specialized terpenoid metabolites may be cyclic or linear and their structural diversity is seemly infinite due secondary modifications of terpene scaffolds and stereochemical arrangements (Dudareva et al., 2004). The biosynthesis of all the central precursors of terpenoids occurs along two independent pathways resulting in the formation of the 5-carbon terpenoid precursors, isopentenyl pyrophosphate (IPP)

and its isomer dimethylallyl pyrophosphate (DMAPP; Fig. 1). The methylerythritol phosphate (MEP) pathway generates IPP and DMAPP exclusively in plastids as precursors in monoterpenoid (C_{10}) and diterpenoid (C_{20}) biosynthesis (Lichtenthaler et al., 1997). This pathway is present in all photosynthetic eukaryotes and in most bacteria and is absent in animals, fungi and archaea (Tholl and Lee, 2011). The MEP pathway progresses through seven enzymatic steps that were first identified in *E. coli* and later confirmed in *A. thaliana* (Fig. 1; Lichtenthaler et al., 1997). Intermediates in the MEP pathway have also received increased attention due to involvement in pathway regulation and plant stress responses. For example, 2-*C*-methyl-D-erythritol 2,4, cyclodiphosphate (MEcPP), an intermediate of the MEP pathway, has been implicated in retrograde signaling in response to abiotic stress, as well as in adaptive growth (Xiao et al., 2012; Jiang et al., 2018).

The mevalonic acid (MVA) pathway occurs independently of the MEP pathway and is present in most eukaryotes including plants (Tholl and Lee, 2011). Unlike the MEP pathway, the six enzymatic steps of the MVA pathway occur in the endoplasmic reticulum, peroxisome and/or cytosol to produce IPP as the precursor of C_{15} sesquiterpenoid and C_{30} triterpenoid compounds (Tholl and Lee, 2011).

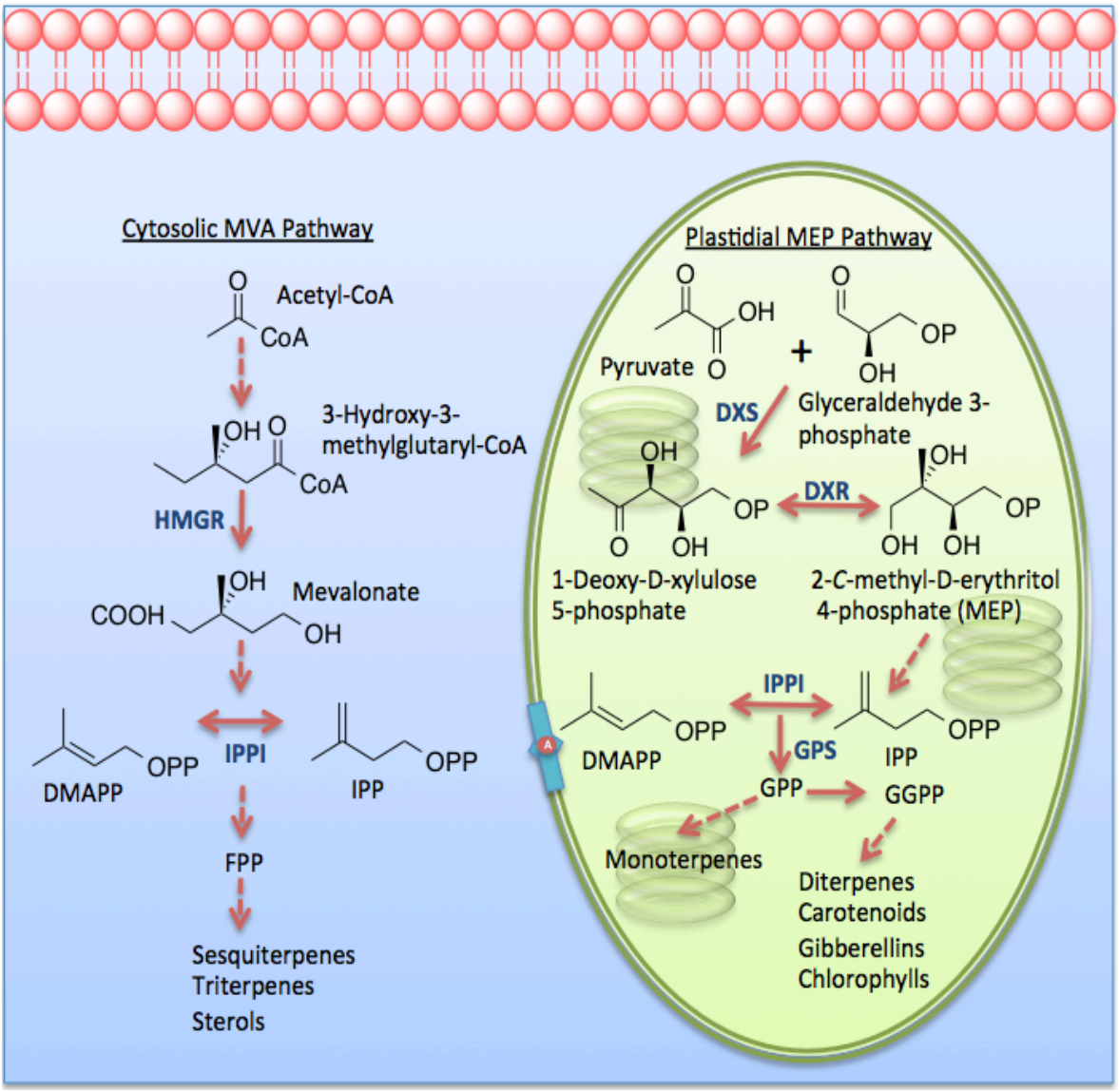
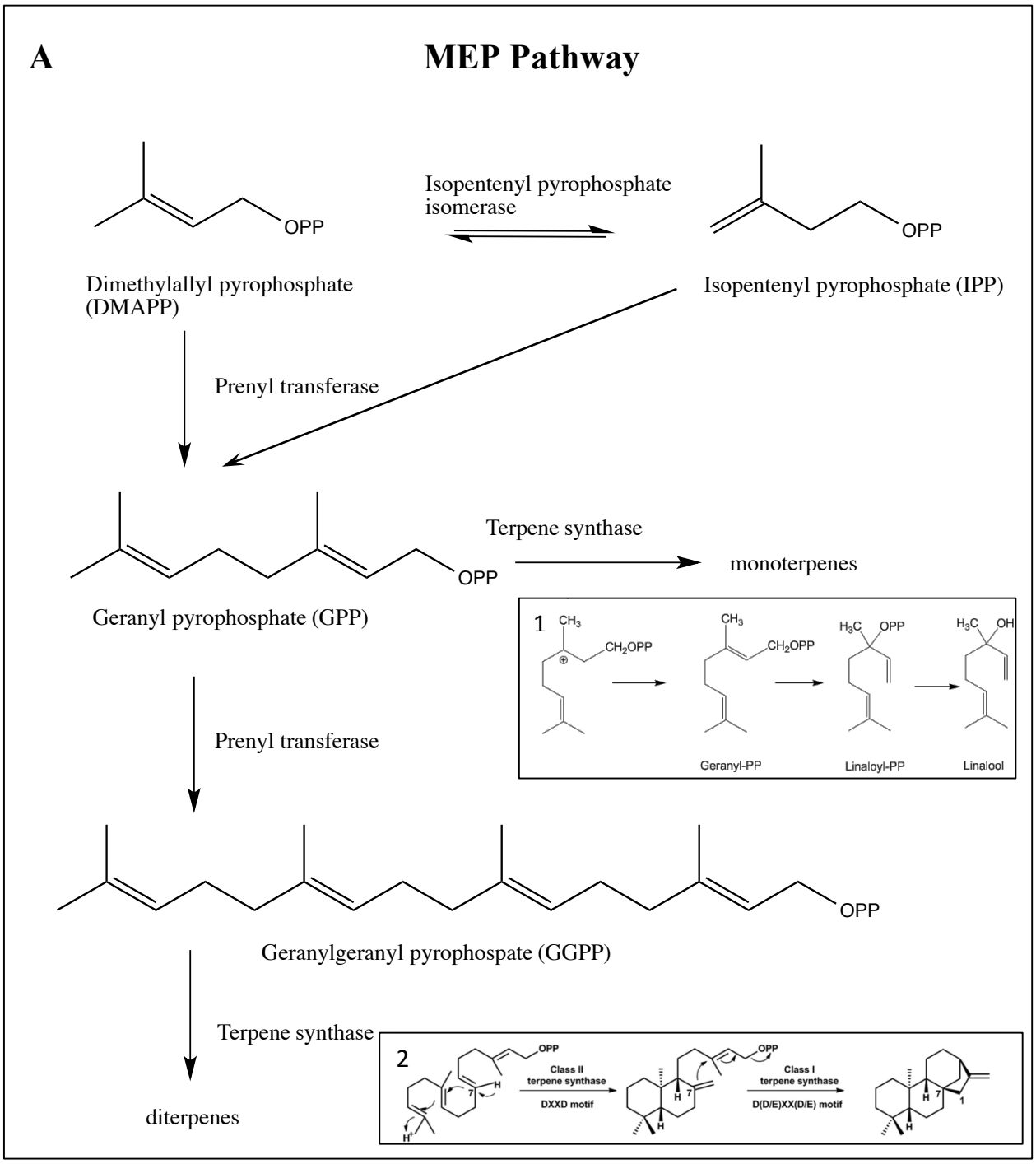


Figure 1. Terpenoid biosynthesis through the mevalonic acid (MVA) and methylerythritol phosphate (MEP) pathways. "A" indicates that the isoprenoid units isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) may be shared between pathways via transport between the compartments. Solid red arrows indicate single reaction steps with respective enzymes listed in blue. Double arrows indicate reversible reactions or isomeric conversions. Dashed red arrows indicate multiple enzymatic steps required for downstream products. HMGR: 3-hydroxy-3-methylglutaryl-CoA reductase, DXS: deoxyxylulose phosphate synthase, DXR: deoxy-D-xylulose 5-phosphate reductoisomerase, IPPI: isopentenyl pyrophosphate isomerase, GPS: geranyl pyrophosphate synthase.

The second major step in terpenoid biosynthesis is the condensation of the 5-carbon units IPP and DMAPP to form the central prenyl pyrophosphate precursors of terpenoids (Fig. 2A and B). These reactions are catalyzed by enzymes called prenyl transferases or isoprenyl diphosphate synthases. Specifically, a head to tail (1'-4') condensation between IPP and DMAPP in the form of an ionization-condensation-elimination reaction leads to the formation of C_{10} geranyl diphosphate (GPP), the precursor substrate for monoterpenoid biosynthesis (Croteau and Shaskus, 1985; Lichtenthaler et al., 1997; Tholl and Lee, 2011). Further condensation reactions can occur with the addition of IPP resulting in C_{15} farnesyl diphosphate (*E,E*-FPP) and C_{20} geranylgeranyl diphosphate (all *trans*-GGPP) as the precursors in sesquiterpenoid and diterpenoid formation, respectively (Tholl and Lee, 2011). The prenyl diphosphate intermediates (GPP, FPP, GGPP) are then further converted to monoterpenes, sesquiterpenes and diterpenes, respectively, by the activity of terpene synthase enzymes (TPS, Figs. 1 and 2A and B). All TPSs form carbocations from their prenyl diphosphate substrates, which undergo numerous reactions in the catalytic center of the TPS enzyme that can include cyclizations, hydride shifts, rearrangements, and deprotonations (Christianson, 2017). These reactions can lead to the formation of several linear or cyclic products by a single enzyme. Monoterpenoids are synthesized by dephosphorylation and ionization of GPP resulting in a geranyl carbocation, which is the precursor of cyclic and acyclic C_{10} terpenoids. The formation of sesquiterpenoids begins with the ionization of FPP generating a farnesyl cation that can undergo isomerization to a nerolidyl cation. The latter cation is the primary cation for cyclic terpenoid formation whereas the farnesyl cation gives rise to acyclic C_{15} terpenoids. C_{20} terpenoids result from carbocation formation from GGPP based on two proposed mechanisms. In the first described reaction, ionization of the pyrophosphate of GGPP occurs

resulting in a geranylgeranyl cation that can undergo distinctive reactions to form various diterpenoid products (class I). Substrate protonation catalyzed by class II TPSs results in the cyclization of GGPP to *ent*-copalyl pyrophosphate (CPP) that can then be further converted to diterpenoid products by class I TPS activity.



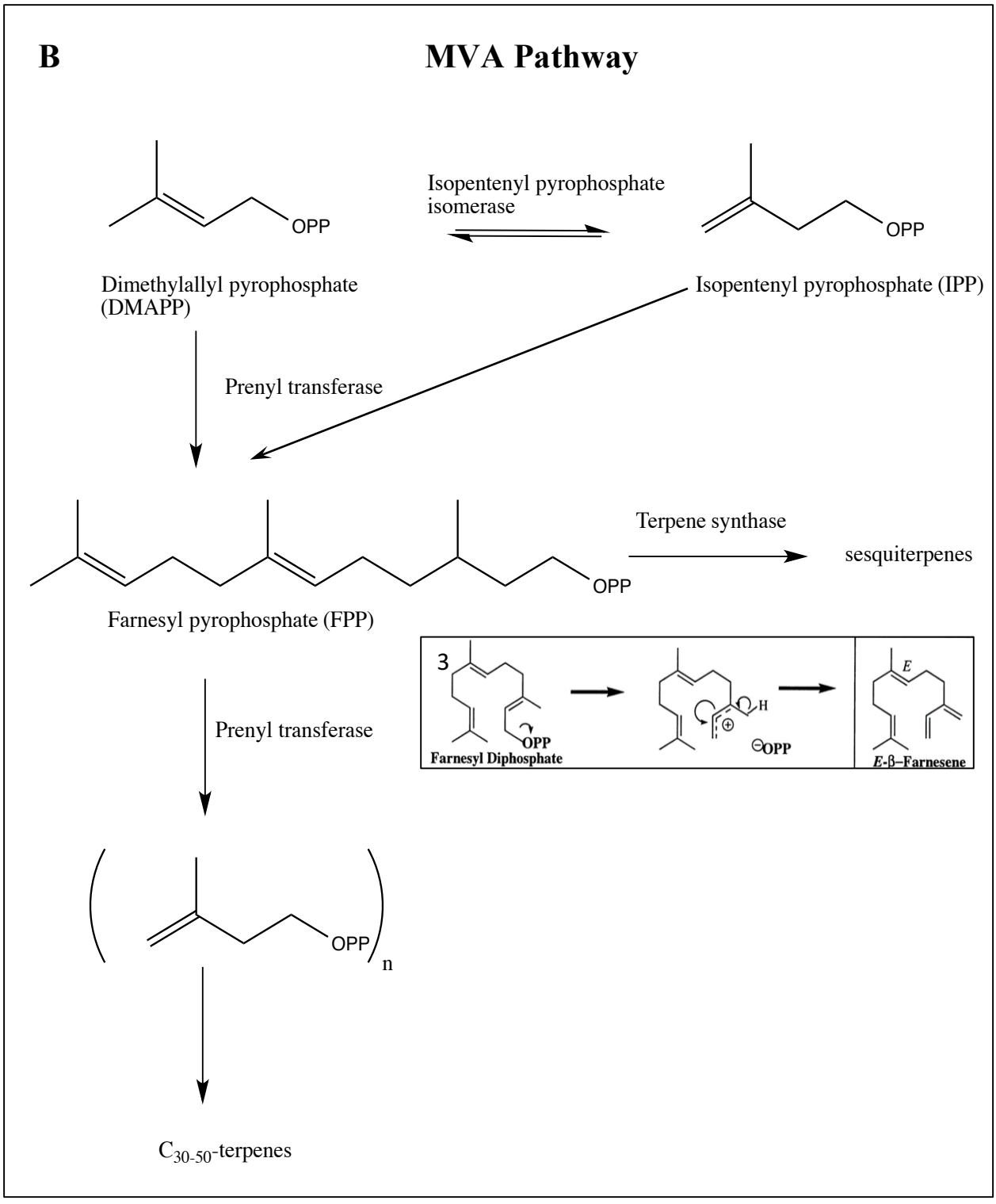


Figure 2. Generalized scheme for the biosynthesis of terpenoid compounds by the activities of prenyl transferase and terpene synthase enzymes. A) Formation of monoterpenes and diterpenes derived from the plastidial MEP pathway B) Sesquiterpene and $C_{30\text{-}50}$ terpene biosynthesis through the cytosolic MVA pathway. Insert 1) conversion of GPP to linalool via class I monoTPS activity 2) conversion of GGPP to CPP via a class II diTPS and conversion to CPP to ent-kaurene via class I diTPS activity 3) conversion of FPP to farnesene via class I sesquiTPS activity.

The products made by TPS enzymes can be further modified by secondary reactions such as hydroxylations, methylations and dehydrogenations, which contribute to the structural diversity of terpenoid compounds (Christianson, 2017). Moreover, gene duplication often leads to the evolution of large species-specific families of TPS genes and is a common mechanism to generate biosynthetic complexity and structural diversity in terpenoid metabolism (Chen et al., 2014). Bohlmann et al. (2000) categorized *TPS* genes into seven sub-families based on sequence homology of putative full length *TPSs* derived from sequenced plant genomes. Specifically, the TPS-c family, which represents the most conserved clade among land plants, contains primarily diterpene synthases (Fig. 3). In contrast, the gymnosperm specific clade (TPS-d) is represented by only monoterpene and sesquiterpene synthases with similar results found in the angiosperm specific clade (TPS-g). Similar patterns are observed for the remaining gene families in addition to an organism specific family (TPS-h, *Selaginella moellendorffii)*. Investigations within gene families provide evidence for tissue-specific expression and protein localization based on promoter-reporter and transit peptide analysis, respectively. Further, the presence of aspartate rich motifs D(D/E)XX(D/E) (Class I) and DXXD (Class II), necessary for metal cofactor chelation and required for enzyme activity, is found among all functional TPSs (Chen et al., 2011). Although high sequence similarly is found among many TPS proteins, sustained mutation over time and the promiscuous nature of TPSs (i.e. many products from the same substrate) may be sufficient to generate the high chemical diversity of terpenoid products (Pichersky and Gang, 2000).

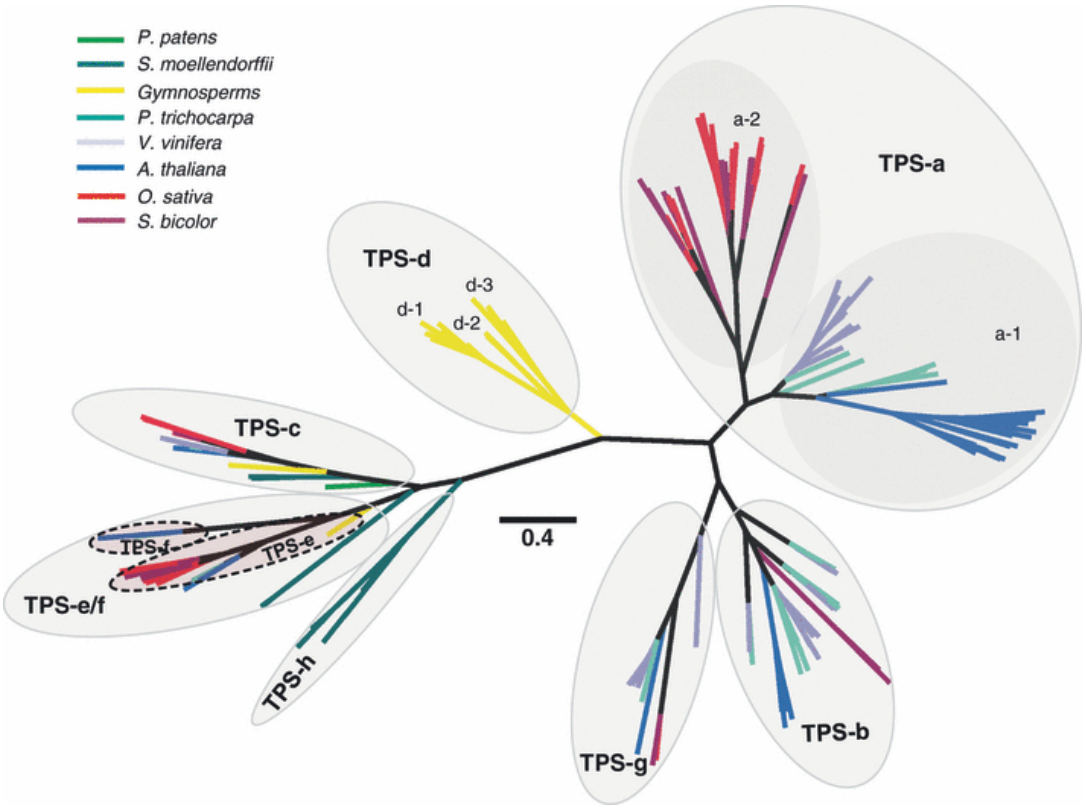


Figure 3. Phylogeny of putative full-length TPSs from seven sequenced plant genomes and representative characterized TPSs from gymnosperms. Subfamily TPS-c (most conserved among land plants), subfamily TPS-e/f (conserved among vascular plants), subfamily TPS-h (*Selaginella moellendorffii* specific), subfamily TPS-d (gymnosperm specific), and three angiosperm-specific subfamilies TPS-b, TPS-g and TPS-a. The TPS-a subfamily is further divided into two groups, a-1 being dicot-specifc and a-2 being monocot-specific. The TPS-d subfamily is further divided into three groups, d-1, d-2 and d-3, which show distinction in function of TPSs in each group. The TPS-e/f subfamily is merged from the previously separate TPS-e and TPS-f subfamilies, which are also shown on the phylogenetic tree. Figure used with permission from Chen et al., 2011.

Terpene secondary metabolism can occur constitutively in above- and belowground tissues and compounds are either released directly into the environment or stored in specialized structures such as oil glands or ducts (Dudareva et al., 2004). In addition, terpene biosynthesis is often induced in response to biotic stress due to the induced expression of terpene synthases. Terpenoids are found in a wide range of tissue types (flowers, leaves, roots) and often accumulate in defensive

secretory structures for quick release in response to attack or infection. An example of such cellular structures are glandular trichomes that enhance the effectiveness of chemical defenses by stopping predator attack preemptively (Theis and Lerdau, 2003). Volatile terpenoids can also be constitutively released from flowers or are emitted under induced conditions leaf tissue to attract pollinators or natural enemies of herbivores (Kessler and Baldwin, 2001). Additionally, the biosynthesis of terpenoid compounds in belowground non-storage organs can occur *de novo* due to the potential for cytotoxicity (Theis and Lerdau, 2003).

The biosynthesis of terpenoids can often be induced by abiotic stress, presumably to protect plants from reactive oxygen species (Vickers et al., 2009). In response to high light and temperature, emission of isoprene, one of the smallest volatile terpenes, can increase from 5% to 50% of the carbon fixed by photosynthesis (Loreto and Fineschi, 2015). Drought stress should limit aboveground volatile terpene production by inducing stomatal closure thereby reducing carbon supply; however, terpenoid biosynthesis often increases in response to drought albeit below severe stress conditions (Loreto and Schnitzler, 2010). In maize roots, biosynthesis and accumulations of terpene phytoalexins (low molecular weight antimicrobials) occurs in response to water deficit, and phytoalexins mutants are less tolerant to drought (Vaughan et al., 2015). Recent work in the bioenergy crop switchgrass provided additional evidence that terpenoids accumulate in response to abiotic stressors (i.e. osmotic damage via copper sulfate and UV radiation) suggesting a conserved stress response across model grasses (Pelot et al., 2018). Together, these studies open further questions as to what extent terpenoids mediate biotic and abiotic stress responses either directly or indirectly through chemically mediated interactions.

Plant-Microbe Interactions in the Rhizosphere

Plants interact with an incredible diversity of microorganisms in a complex ecosystem located around the plant root called the rhizosphere (Zhang et al., 2015). The importance of rhizosphere plant-microbial associations on plant productivity and resilience has been well described, however knowledge of the specific mechanisms for such associations remains limited (Junker and Tholl, 2013; Massalha et al., 2017). Classical examples of small molecules mediating belowground plant-microbe interactions include the release of sesquiterpene lactones (strigolactones) by plant roots for the recruitment of arbuscular mycorrhiza fungi (Fig. 4; AMF). In this mutualistic exchange, the host plant receives increased access to nutrients and water while providing the AMF with a stable supply of carbohydrates (Akiyama and Hayashi, 2006).

Plant-microbe associations have also been directly linked to increased plant fitness through enhancement of nutrient assimilation and resilience against biotic and abiotic stressors (Yang et al., 2009; Ker et al., 2014). Root-derived metabolites mediate a chemical dialogue resulting in the formation of complex relationships with plant growth promoting rhizobacteria (PGPR). PGPR can mediate resilience to unfavorable environmental conditions by stimulating plant growth and decreasing susceptibility to disease (Huang et al., 2014). PGPR can also contribute to plant defense by eliciting a response of induced systemic resistance (Yang et al., 2009). During drought stress conditions, plants accumulate ethylene, which leads to the inhibition of plant growth. PGPR can allow plants to overcome this inhibition by degrading the ethylene precursor 1-aminocyclopropane-1-carboxylate enhancing stress tolerance (Yang et al., 2009). The basis of these interactions is largely mediated by the exudation or release of primary and secondary metabolites, thereby supporting the development of a taxonomically rich microbial community (Badri et al., 2013). Bacterial cell densities in the rhizosphere show a 100-fold increase over

densities found in the surrounding soil, illustrating strong enrichment of microbes to the root zone (Whipps, 2001). Moreover, in rice, microbial communities outside the roots have been shown to be distinct from the community inside the root (Edwards et al., 2015). Although this may be due to the physical barrier of the epidermal tissue, a selective chemic barrier may also exist. In Arabidopsis, for example, salicylic acid is responsible for the recruitment of specific bacterial families (Lebeis et al., 2015). To what extent additional root metabolites influence microbial community assembly and maintenance is not well known.

Complex plant–microbe interactions have been described for bacteria inhabiting the root interior (endosphere) of the essential oil producing grass vetiver (*Vetiveria zizanioides*). In vetiver roots, sesquiterpenes are used by by endosphere microbes as unique carbon sources and metabolized producing a richer chemical diversity than in sterile grown roots (Del Giudice et al., 2008). A similar degradation of volatile terpenes has been observed for rhizobacteria including *Pseudomonas putida* which can utilize the monoterpene camphor aa a sole carbon source in vitro (Kleinheinz et al., 1999; Prasad et al., 2013). There is growing evidence that similar microbial metabolic activities occur with other root-derived metabolites such as phenylpropanoids and hydroxamic acids i.e. benzoxazinoids (Badri et al., 2013). In the case of benzoxazinoids (e.g. DIMBOA), which are exuded by maize roots in defense against the European corn borer larvae, were also found to recruit the beneficial microbe *Pseudomonas putida* to the rhizosphere during young and vulnerable plant growth stages (Neal et al., 2012). It is therefore plausible that root specialized metabolites have multifunctional semiochemical (signaling) and defensive roles in mediating belowground interactions. Further investigations are needed to demonstrate fundamental roles of plant-derived secondary metabolites including VOCs in the root

environment. This includes reverse genetic attempts, although challenging, to provide a better avenue for characterizing biological functions of small molecules released from plant roots.

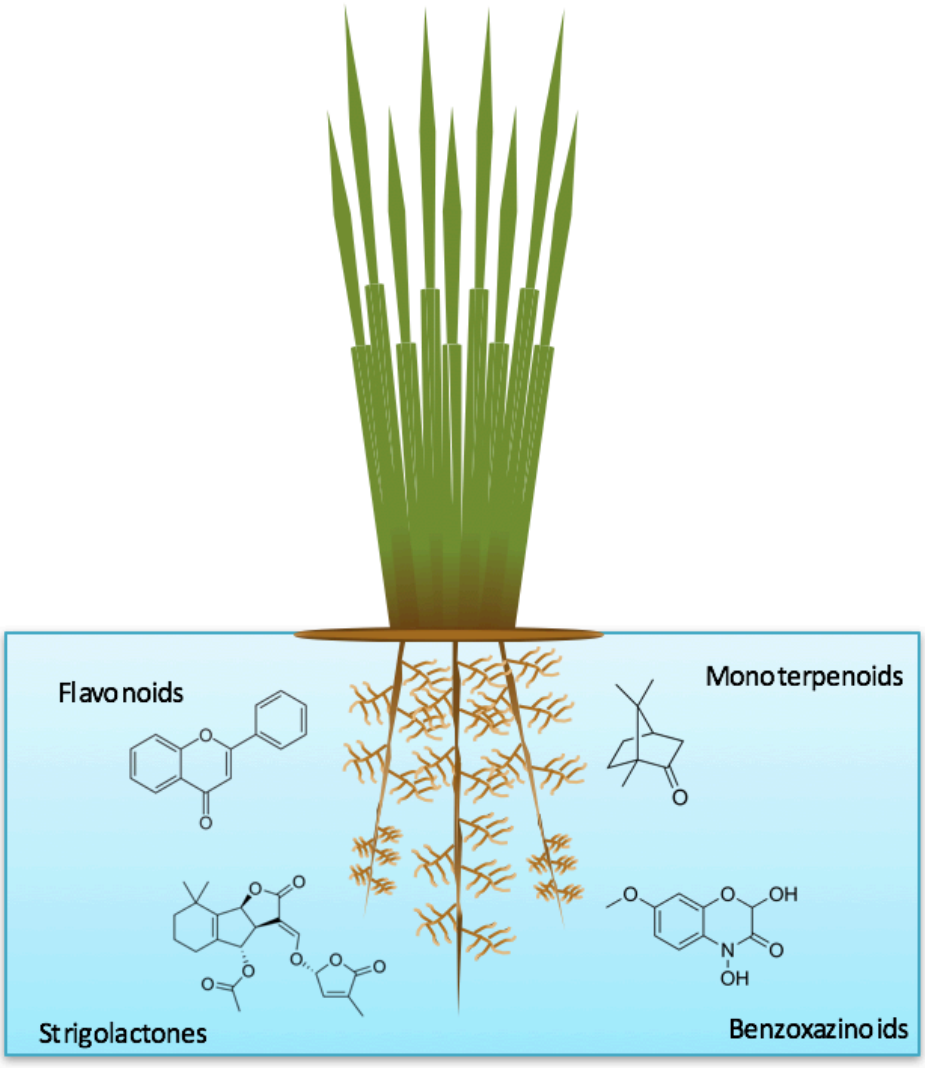


Figure 4. Structural examples of selected chemicals involved in belowground plant-microbe interactions. Structures shown: general flavonoid backbone; (+)-orobranchyl acetate (strigolactone); camphor (monoterpenoid); DIMBOA (benzoxazinoid).

Switchgrass as a Model Study System

Switchgrass (*Panicum virgatum*, Poaceae) is a C_4 warm season perennial prairie grass native to North America. Of the prairie species, switchgrass is considered one of the more important taxa for sustaining prairie biodiversity (Sanderson et al., 2006). Historically used as a forage grass, increased breeding efforts were implemented beginning over 20 years ago to develop switchgrass as model bioenergy crop (Casler et al., 2011). The high lignocellulosic content of switchgrass, its limited growth requirements, and expansive habitat range have resulted in extensive research and resource allocation to switchgrass as a bioenergy model (Casler et al., 2004; McLaughlin and Kszos, 2005). In addition, switchgrass has considerable resistance to drought, pests and diseases, although little is still known about the mechanism of this resistance, partially due to the genetic complexity of this species (Bouton, 2008). However, field observations of widespread rust fungal infection of switchgrass leaves and resulting ethanol reduction have increased the urgency to elucidate resilience mechanisms (Sykes et al., 2016).

Switchgrass genotypes are divided into two distinct ecotypes (upland, lowland) based on ecological origin, ploidy levels, and life history characteristics (Sanderson and Wolf, 1995). Lowland switchgrass ecotypes generally produce more biomass, are predominantly tetraploid, and often grow in wet mild climates (Sanderson and Wolf, 1995). In contrast, upland ecotypes are of mixed ploidy (tetra-, hexa- and octoploid), produce less biomass, and are found in colder, drier areas (Sanderson and Wolf, 1995). Although cross-pollination is common in grasses, the complex genetic background and varying ploidy levels appear to act as a prezygotic barrier between switchgrass cultivars (Wullschleger et al., 2010). However, increased genomic, transcriptomic and transformation resources have allowed more in depth analyses of switchgrass stress resilience.

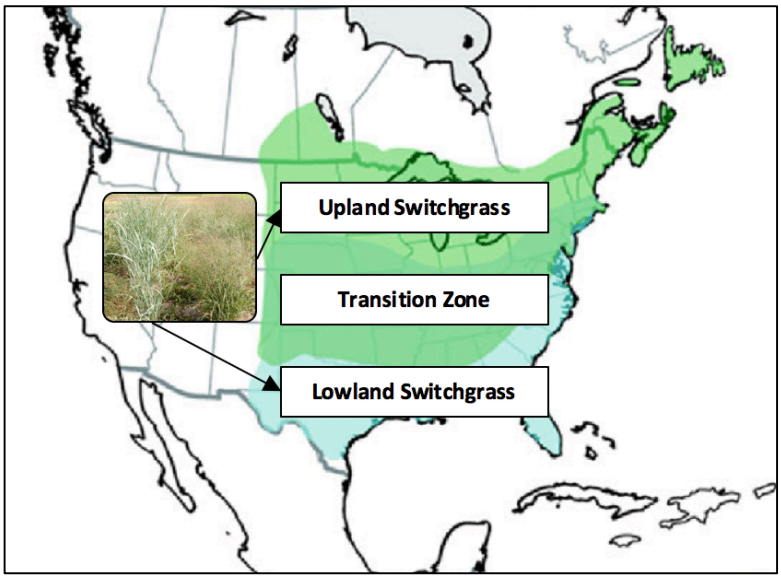


Figure 5. Distribution of upland and lowland switchgrass ecotypes across the United States. Adapted from Casler et al., 2011.

Recent transcriptional analysis of switchgrass (cv. AP13) leaves during herbivory from green bug (*Schizaphis graminum*) revealed a complex defensive response resulting in increased reactive oxygen species and upregulation of terpene synthase genes (Donze-Reiner et al., 2017). In addition, comparative transcriptome profiling of upland (cv. VS16) and lowland (cv. AP13) varieties revealed an expanded metabolic complexity of biotic and abiotic stress related genes in switchgrass (Ayyappan et al., 2017). Four switchgrass cultivars were also found to accumulate high levels of glycosidic triterpenes (steroidal saponins) in leaves and stems, similar to observations in other model plant systems e.g. *Medicago truncatula* (Huhman et al., 2005; Lee et al., 2009). Such compounds have high toxicity to ruminants due to photosensitization, and are known defensive compounds used by plants against pathogens and pests (Thimmappa et al., 2014).

Functional genomics investigations, although limited in switchgrass, have revealed direct correlations between diterpene production and drought and UV stress suggesting terpenoids could be involved in abiotic stress resistance in switchgrass (Pelot et al., 2018). However, no studies to date have investigated the formation and function of volatile terpenes (C_{10} and C_{15}) in switchgrass in association with biotic or abiotic stressors.

In addition, little is known about the role of secondary metabolites such as terpenes in mediating beneficial associations of switchgrass with colonizing microbes. As detailed above, microbial communities provide strong support of plant growth and stress tolerance. Switchgrass plants inoculated with PGPR, for example, showed a 40% yield increase compared to non-inoculated samples (Ker et al., 2012). Moreover, switchgrass seedlings inoculated with the endophtye *Burkholderia phytofirmans* showed a significant increase in belowground biomass with substantially enhanced root length and diameter (Yang et al., 2009). These results demonstrate the significance of microbial interactions for the growth performance of switchgrass without addressing the nature of host-specific factors controlling these interactions. Therefore, a better understanding of chemical effects in switchgrass-microbe interactions and their underlying genetics may allow for breeding or engineering more resistant genotypes.

Carrot as a Model Study System

In addition to studies addressing the biological functions of volatile terpenes, much knowledge has been gained on the contributions of these metabolites to flavor and aroma (Stewart, 2013; Pichersky and Raguso, 2018). Carrot (*Daucus carota L.*) has been widely cultivated since 900 AD due to the presence of high levels of beta carotene, the precursor to vitamin A (Simon et al., 1980). The pleasant flavor and aroma of carrots are largely determined by blends of volatile

terpenes as well as phenolic compounds, and sugar content (Simon et al., 1980; Alasalvar et al., 2001; Kreutzmann et al., 2008). Volatile terpenes are synthesized and stored as part of an essential oil consisting predominantly of monoterpenes and sesquiterpenes accumulated in highly interconnected phloem oil ducts (Senalik and Simon, 1986; Kreutzmann et al., 2008). Such sensory traits can be highly genotype specific. Orange cultivars, for example, have a spicy and woody aroma, whereas yellow and purple cultivars have a sweeter odor (Fukuda et al., 2013). Red cultivars can have a sharp odor and bitter flavor which is attributed to low levels of sugars and high levels of sesquiterpenes (Kreutzmann et al., 2008). As mentioned earlier, abiotic factors such as elevated temperature can also significantly increase harsh and bitter taste in carrot which is directly correlated with increases in terpene levels (Rosenfeld et al., 2002). These genotype and environmentally dependent differences in aroma and flavor attributes have raised increased attention to understand the genetic and biochemical determinants of carrot taste (Guerin and Ryan, 1984) and support breeding (conventional or accelerated) of carrots with more desirable traits. First efforts in investigating terpene metabolism in carrot have been made (Yahyaa et al., 2013; Yahyaa et al., 2015; Keilwagen et al., 2017), and were expanded as part of this study for a more rigorous characterization of volatile terpene biosynthesis in correlation with representative terpenes in different carrot genotypes.

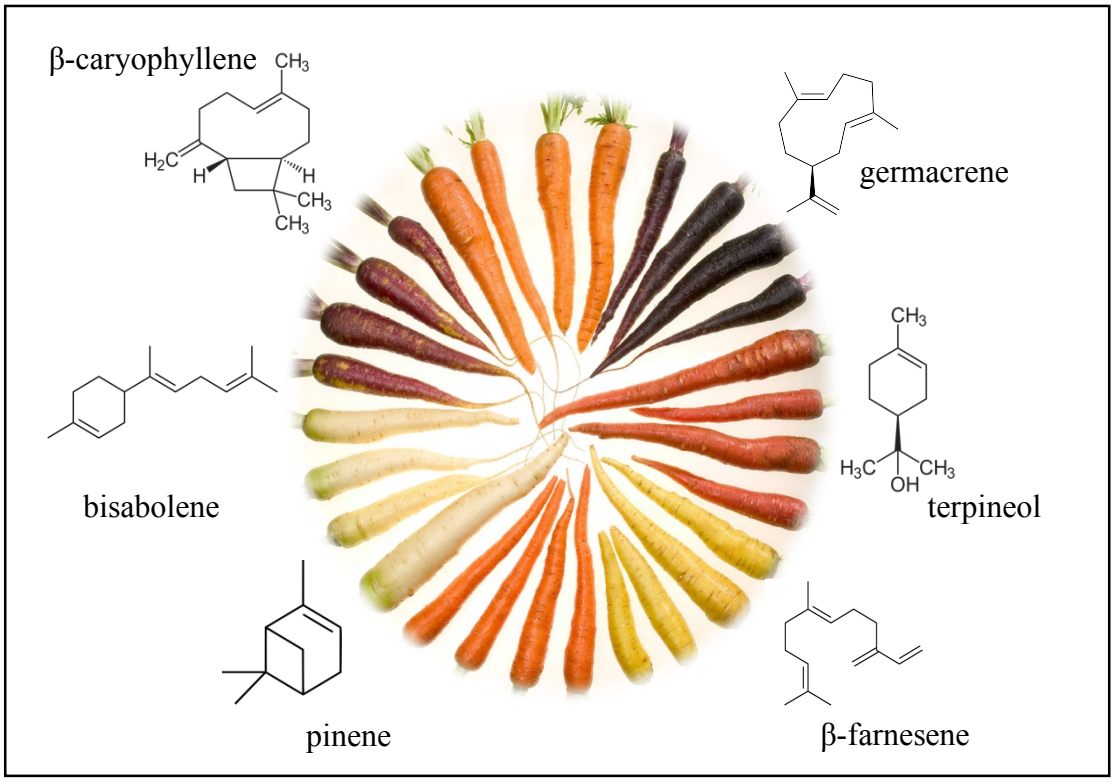


Figure 6. Major volatile compounds found in carrot roots. Figure modified from www.carrotmuseum.co.uk.

Overview of Research

Plant specialized metabolites have received much attention for their role as signaling molecules, defensive compounds and their involvement in environmental stress response. Especially, VOCs have been implicated in mediating chemical interactions of plants. Such functions have been studied mostly in photosynthetic tissues, while knowledge on the role of volatile compounds in roots is still limited. In addition, how the metabolism of VOCs contributes to the flavor and aroma of root crops and their change in response to environmental stress is not well understood. We have investigated these questions at the example of terpene metabolism in two different plant systems: In switchgrass we have examined terpene biosynthesis and dynamics in response to biotic stress and also performed a limited study on responses of terpene metabolism

to drought as an abiotic stress. Furthermore, we have investigated the potential role of monoterpenes (camphor and borneol) that accumulate in roots of switchgrass and the related grass *Setaria* as host specific factors in root microbial colonization. In an independent study of carrot we report the biochemical characterization of several terpene synthases and show that differences in carrot flavor and aroma are related to levels of specific terpene compounds.

The objectives of this study were to:

i. Identify and characterize root-specific genes involved in terpenoid formation in switchgrass abiotic and biotic stress responses.
ii. Examine the production and effect of volatile terpenoids in switchgrass and *Setaria* root-microbe interactions.
iii. Identify terpene synthases as the genetic determinants of carrot aroma and flavor.

Chapter II presents the identification and biochemical characterization of terpene synthase genes in switchgrass. Several TPSs were found that are likely to be responsible or contribute to stress- or defense hormone-induced emissions of volatile terpenes in leaves and roots.

Chapter III presents the gene functional analysis of borneol production in switchgrass roots. We show that borneol is accumulated in roots; however, its biosynthesis occurs in leaves and can be stimulated in leaves by treatments of roots with communities of rhizobacteria. Furthermore, experimental approaches to determine the effects of borneol on root microbial community composition are presented.

Chapter IV represents an independent project focused on terpene biosynthesis in carrot. We present the identification and biochemical characterization of terpene synthases contributing to carrot aroma and flavor.

The dissertation includes three appendices to Chapter II and III. Appendix A contains experimental results on volatile terpene and non-volatile metabolic profiles as well as terpene synthase transcriptional profiles in switchgrass roots upon drought stress treatment. Appendix B presents results on borneol biosynthesis in *Setaria viridis* and microbial community composition in roots of *Setaria* wild type and borneol biosynthetic mutants. Appendix C presents the detailed data analysis pipeline for analysis of microbial communities using QIIME.

References

1. Akiyama, K., and Hayashi, H. (2006). Strigolactones: Chemical signals for fungal symbionts and parasitic weeds in plant roots. *Ann. Bot.* 97**,** 925-931. doi: 10.1093/aobmc1063.
2. Alasalvar, C., Grigor, J.M., Zhang, D.L., Quantick, P.C., and Shahidi, F. (2001). Comparison of volatiles, phenolics, sugars, antioxidant vitamins, and sensory quality of different colored carrot varieties. *J. Agric. Food Chem.* 49**,** 1410-1416. doi: 10.1021/jf000595h.
3. Ayyappan, V., Saha, M.C., Thimmapuram, J., Sripathi, V.R., Bhide, K.P., Fiedler, E., Hayford, R.K., and Kalavacharla, V. (2017). Comparative transcriptome profiling of upland (VS16) and lowland (AP13) ecotypes of switchgrass. *Plant Cell Rep.* 36**,** 129-150. doi: 10.1007/s00299-016-2065-0.
4. Badri, D.V., Chaparro, J.M., Zhang, R.F., Shen, Q.R., and Vivanco, J.M. (2013). Application of natural blends of phytochemicals derived from the root exudates of Arabidopsis to the soil reveal that phenolic-related compounds predominantly modulate the soil microbiome *J. Biol. Chem.* 288**,** 30503-30503. doi: 10.1074/jbc.A112.433300.
5. Bohlmann, J., Gershenzon, J., and Aubourg, S. (2000). Biochemical, molecular genetic and evolutionary aspects of defense-related terpenoid metabolism in conifers. *Recent Adv Phytochem* 34**,** 109-150. doi: 10.1016/S0079-9920(00)80006-4.
6. Bouton, J. (2008). "Improvement of Switchgrass as a Bioenergy Crop," in *Genetic Improvement of Bioenergy Crops,* ed. W. Vermerris. (New York, NY: Springer), 309-345.
7. Casler, M.D., Tobias, C.M., Kaeppler, S.M., Buell, C.R., Wang, Z.Y., Cao, P.J., Schmutz, J., and Ronald, P. (2011). The switchgrass genome: tools and strategies. *Plant Genome* 4**,** 273-282. doi: 10.3835/plantgenome2011.10.0026.
8. Casler, M.D., Vogel, K.P., Taliaferro, C.M., and Wynia, R.L. (2004). Latitudinal adaptation of switchgrass populations. *Crop Sci.* 44**,** 293-303. doi: 10.2135/cropsci2004.2930.
9. Chen, F., Tholl, D., Bohlmann, J., and 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. doi: 10.1111/j.1365-313X.2011.04520.x.
10. Chen, H., Li, G.L., Kollner, T.G., Jia, Q.D., Gershenzon, J., and Chen, F. (2014). Positive Darwinian selection is a driving force for the diversification of terpenoid biosynthesis in the genus Oryza. *BMC Plant Biol.* 14. doi: 10.1186/s12870-014-0239-x.
11. Chen, X.J., Chen, H., Yuan, J.S., Kollner, T.G., Chen, Y.Y., Guo, Y.F., Zhuang, X.F., Chen, X.L., Zhang, Y.J., Fu, J.Y., Nebenfuhr, A., Guo, Z.J., and Chen, F. (2018). The rice terpene synthase gene OsTPS19 functions as an (S)-limonene synthase in planta, and its overexpression leads to enhanced resistance to the blast fungus Magnaporthe oryzae. *Plant Biotechnol. J.* 16**,** 1778-1787. doi: 10.1111/pbi.12914.
12. Cheng, A.X., Xiang, C.Y., Li, J.X., Yang, C.Q., Hu, W.L., Wang, L.J., Lou, Y.G., and Chen, X.Y. (2007). The rice (*E*)-beta-caryophyllene synthase (OsTPS3) accounts for the major inducible volatile sesquiterpenes. *Phytochemistry* 68**,** 1632-1641. doi: 10.1016/j.phytochem.2007.04.008.

13. Christianson, D.W. (2017). Structural and Chemical Biology of Terpenoid Cyclases. *Chem Rev* 117**,** 11570-11648. doi: 10.1021/acs.chemrev.7b00287.
14. Croteau, R., and Shaskus, J. (1985). Biosynthesis of monoterpenes - demonstration of a geranyl pyrophosphate - (-)-bornyl pyrophosphate cyclase in soluble enzyme preparations from Tansy (Tanacetum vulgare). *Arch Biochem Biophys* 236**,** 535-543. doi: 10.1016/0003-9861(85)90656-3.
15. Daferera, D.J., Tarantilis, P.A., and Polissiou, M.G. (2002). Characterization of essential oils from Lamiaceae species by Fourier transform Raman spectroscopy. *J. Agric. Food Chem.* 50**,** 5503-5507. doi: 10.1021/jf0203489.
16. Degenhardt, J., Hiltpold, I., Kollner, T.G., Frey, M., Gierl, A., Gershenzon, J., Hibbard, B.E., Ellersieck, M.R., and Turlings, T.C.J. (2009). Restoring a maize root signal that attracts insect-killing nematodes to control a major pest. *Proc. Natl. Acad. Sci. U.S.A.* 106**,** 13213-13218. doi: 10.1073/pnas.0906365106.
17. Del Giudice, L., Massardo, D.R., Pontieri, P., Bertea, C.M., Mombello, D., Carata, E., Tredici, S.M., Tala, A., Mucciarelli, M., Groudeva, V.I., De Stefano, M., Vigliotta, G., Maffei, M.E., and Alifano, P. (2008). The microbial community of Vetiver root and its involvement into essential oil biogenesis. *Environ. Microbiol.* 10**,** 2824-2841. doi: 10.1111/j.1462-2920.2008.01703.x.
18. Dudareva, N., Negre, F., Nagegowda, D.A., and Orlova, I. (2006). Plant volatiles: Recent advances and future perspectives. *Crit. Rev. Plant Sci.* 25**,** 417-440. doi: 10.1080/07352680600899973.
19. Dudareva, N., Pichersky, E., and Gershenzon, J. (2004). Biochemistry of plant volatiles. *Plant Physiol.* 135**,** 1893-1902. doi: 10.1104/pp.104.049981.
20. Edwards, J., Johnson, C., Santos-Medellin, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., Eisen, J.A., and Sundaresan, V. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* 112**,** E911-920. doi: 10.1073/pnas.1414592112.
21. Fukuda, T., Okazaki, K., and Shinano, T. (2013). Aroma characteristic and volatile profiling of carrot varieties and quantitative role of terpenoid compounds for carrot sensory attributes. *J. Food Sci.* 78**,** S1800-S1806. doi: 10.1111/1750-3841.12292.
22. Gershenzon, J. (1994). Metabolic costs of terpenoid accumulation in higher plants. *J. Chem. Ecol.* 20**,** 1281-1328. doi: 10.1007/Bf02059810.
23. Guerin, P.M., and Ryan, M.F. (1984). Relationship between root volatiles of some carrot cultivars and their resistance to the carrot fly *Psila rosae*. *Entomol. Exp. Appl.* 36**,** 217-224. doi: 10.1111/j.1570-7458.1984.tb03431.x.
24. Huang, M., Sanchez-Moreiras, A.M., Abel, C., Sohrabi, R., Lee, S., Gershenzon, J., and Tholl, D. (2012). The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (*E*)-β-caryophyllene, is a defense against a bacterial pathogen. *New Phytol.* 193**,** 997-1008. doi: 10.1111/j.1469-8137.2011.04001.x.
25. Huang, X.F., Chaparro, J.M., Reardon, K.F., Zhang, R.F., Shen, Q.R., and Vivanco, J.M. (2014). Rhizosphere interactions: root exudates, microbes, and microbial communities. *Botany* 92. doi: 10.1139/cjb-2013-0225.
26. Huber, M., Epping, J., Gronover, C.S., Fricke, J., Aziz, Z., Brillatz, T., Swyers, M., Kollner, T.G., Vogel, H., Hammerbacher, A., Triebwasser-Freese, D., Robert, C.A.M., Verhoeven, K., Preite, V., Gershenzon, J., and Erb, M. (2016). A latex metabolite benefits

plant fitness under root herbivore attack. *PLoS Biol.* 14. doi: 10.1371/journal.pbio.1002332.
27. Huhman, D.V., Berhow, M.A., and Sumner, L.W. (2005). Quantification of saponins in aerial and subterranean tissues of *Medicago truncatula*. *J. Agric. Food Chem.* 53**,** 1914-1920. doi: DOI 10.1021/jf0482663.
28. Iorizzo, M., Ellison, S., Senalik, D., Zeng, P., Satapoomin, P., Huang, J.Y., Bowman, M., Iovene, M., Sanseverino, W., Cavagnaro, P., Yildiz, M., Macko-Podgorni, A., Moranska, E., Grzebelus, E., Grzebelus, D., Ashrafi, H., Zheng, Z.J., Cheng, S.F., Spooner, D., Van Deynze, A., and Simon, P. (2016). A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nat. Genet.* 48**,**
29. . doi: 10.1038/ng.3565.
30. Jiang, J.S., Rodriguez-Furlan, C., Wang, J.Z., de Souza, A., Ke, H.Y., Pasternak, T., Lasok, H., Ditengou, F.A., Palme, K., and Dehesh, K. (2018). Interplay of the two ancient metabolites auxin and MEcPP regulates adaptive growth. *Nat. Commun.* 9. doi: 10.1038/s41467-018-04708-5.
31. Jiang, Y.S., Liao, Q.H., Zou, Y., Liu, Y.Q., and Lan, J.B. (2017). Transcriptome analysis reveals the genetic basis underlying the biosynthesis of volatile oil, gingerols, and diarylheptanoids in ginger (*Zingiber officinale Rosc.*). *Botanical Studies* 58. doi: 10.1186/s40529-017-0195-5.
32. Julsing, M.K., Koulman, A., Woerdenbag, H.J., Quax, W.J., and Kayser, O. (2006). Combinatorial biosynthesis of medicinal plant secondary metabolites. *Biomol. Eng.* 23**,** 265-279. doi: 10.1016/j.bioeng.2006.08.001.
33. Junker, R.R., Loewel, C., Gross, R., Dotter, S., Keller, A., and Bluthgen, N. (2011). Composition of epiphytic bacterial communities differs on petals and leaves. *Plant Biol.* 13**,** 918-924. doi: 10.1111/j.1438-8677.2011.00454.x.
34. Junker, R.R., and Tholl, D. (2013). Volatile Organic Compound Mediated Interactions at the Plant-Microbe Interface. *J. Chem. Ecol.* 39**,** 810-825. doi: 10.1007/s10886-013-0325-9.
35. Keilwagen, J., Lehnert, H., Berner, T., Budahn, H., Nothnagel, T., Ulrich, D., and Dunemann, F. (2017). The terpene synthase gene family of carrot (*Daucus carota* L.). *Front Plant Sci* 8. doi: 10.3389/fpls.2017.01930.
36. Ker, K., Seguin, P., Driscoll, B.T., Fyles, J.W., and Smith, D.L. (2012). Switchgrass establishment and seeding year production can be improved by inoculation with rhizosphere endophytes. *Biomass Bioenerg* 47**,** 295-301. doi: 10.1016/j.biombioe.2012.09.031.
37. Ker, K., Seguin, P., Driscoll, B.T., Fyles, J.W., and Smith, D.L. (2014). Evidence for enhanced N availability during switchgrass establishment and seeding year production following inoculation with rhizosphere endophytes. *Arch Agron Soil Sci* 60**,** 1553-1563. doi: 10.1080/03650340.2014.898840.
38. Kessler, A., and Baldwin, I.T. (2001). Defensive function of herbivore-induced plant volatile emissions in nature. *Science* 291**,** 2141-2144. doi: 10.1126/science.291.5511.2141.
39. Kleinheinz, G.T., Bagley, S.T., St John, W.P., Rughani, J.R., and McGinnis, G.D. (1999). Characterization of alpha-pinene-degrading microorganisms and application to a bench-scale biofiltration system for VOC degradation. *Arch. Environ. Contam. Toxicol.* 37**,** 151-157. doi: 10.1007/s002449900500.

40. Kollner, T.G., Held, M., Lenk, C., Hiltpold, I., Turlings, T.C.J., Gershenzon, J., and Degenhardt, J. (2008). A maize (*E*)-beta-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most American maize varieties. *Plant Cell* 20**,** 482-494. doi: 10.1105/tpc.107.051672.
41. Kreutzmann, S., Thybo, A.K., Edelenbos, M., and Christensen, L.P. (2008). The role of volatile compounds on aroma and flavour perception in coloured raw carrot genotypes. *Int. J. Food Sci. Technol.* 43**,** 1619-1627. doi: 10.1111/j.1365-2621.2007.01662.x.
42. Lawo, N.C., Weingart, G.J.F., Schuhmacher, R., and Forneck, A. (2011). The volatile metabolome of grapevine roots: First insights into the metabolic response upon phylloxera attack. *Plant Physiol. Biochem.* 49**,** 1059-1063. doi: 10.1016/j.plaphy.2011.06.008.
43. Lebeis, S.L., Paredes, S.H., Lundberg, D.S., Breakfield, N., Gehring, J., McDonald, M., Malfatti, S., del Rio, T.G., Jones, C.D., Tringe, S.G., and Dangl, J.L. (2015). Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* 349**,** 860-864. doi: 10.1126/science.aaa8764.
44. Lee, S.T., Mitchell, R.B., Wang, Z., Heiss, C., Gardner, D.R., and Azadi, P. (2009). Isolation, characterization, and quantification of steroidal saponins in switchgrass (*Panicum virgatum* L.). *J. Agric. Food Chem.* 57**,** 2599-2604. doi: 10.1021/jf803907y.
45. Lichtenthaler, H.K., Schwender, J., Disch, A., and Rohmer, M. (1997). Biosynthesis of isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent pathway. *FEBS Lett.* 400**,** 271-274. doi: 10.1016/S0014-5793(96)01404-4.
46. Loreto, F., and Fineschi, S. (2015). Reconciling functions and evolution of isoprene emission in higher plants. *New Phytol.* 206**,** 578-582. doi: 10.1111/nph.13242.
47. Loreto, F., Forster, A., Durr, M., Csiky, O., and Seufert, G. (1998). On the monoterpene emission under heat stress and on the increased thermotolerance of leaves of *Quercus ilex* L. fumigated with selected monoterpenes. *Plant Cell Environ* 21**,** 101-107. doi: 10.1046/j.1365-3040.1998.00268.x.
48. Loreto, F., and Schnitzler, J.P. (2010). Abiotic stresses and induced BVOCs. *Trends Plant Sci.* 15**,** 154-166. doi: 10.1016/j.tplants.2009.12.006.
49. Madhaiyan, M., Poonguzhali, S., Senthilkumar, M., Seshadri, S., Chung, H.Y., Yang, J.C., Sundaram, S., and Sa, T.M. (2004). Growth promotion and induction of systemic resistance in rice cultivar Co-47 (*Oryza sativa L.*) by *Methylobacterium spp. Botan Bulletin Acad* 45**,** 315-324.
50. Martinat, S., Mintalova, T., Dvorak, P., Navratil, J., Klusacek, P., and Kunc, J. (2013). Does rural space benefit from location of anaerobic digestion plants? Perspective of communal administration. *Geographia Cassoviensis* 7**,** 41-49.
51. Massalha, H., Korenblum, E., Tholl, D., and Aharoni, A. (2017). Small molecules below-ground: the role of specialized metabolites in the rhizosphere. *Plant J.* 90**,** 788-807. doi: 10.1111/tpj.13543.
52. Mcgarvey, D.J., and Croteau, R. (1995). Terpenoid Metabolism. *Plant Cell* 7**,** 1015-1026.
53. McLaughlin, S.B., and Kszos, L.A. (2005). Development of switchgrass (*Panicum virgatum*) as a bioenergy feedstock in the United States. *Biomass Bioenerg* 28**,** 515-535. doi: 10.1016/j.biombioe.2004.05.006.
54. Nascimento, N.C., and Fett-Neto, A.G. (2010). *Plant Secondary Metabolism and Challenges in Modifying Its Operation: An Overview.*

55. Neal, A.L., Ahmad, S., Gordon-Weeks, R., and Ton, J. (2012). Benzoxazinoids in root exudates of maize attract Pseudomonas putida to the rhizosphere. *PLoS One* 7. doi: 10.1371/journal.pone.0035498.
56. Norton, R.E., and Hamm, J. (2004). Plant resins: Chemistry, evolution, ecology, and ethnobotany. *J Am Inst Conserv* 43**,** 285-288. doi: Doi 10.2307/4129644.
57. Pelot, K.A., Chen, R., Hagelthorn, D.M., Young, C.A., Addison, J.B., Muchlinski, A., Tholl, D., and Zerbe, P. (2018). Functional diversity of diterpene synthases in the biofuel crop switchgrass. *Plant Physiol.* 178**,** 54-71. doi: 10.1104/pp.18.00590.
58. Penuelas, J., Asensio, D., Tholl, D., Wenke, K., Rosenkranz, M., Piechulla, B., and Schnitzler, J.P. (2014). Biogenic volatile emissions from the soil. *Plant Cell Environ.* 37**,** 1866-1891. doi: 10.1111/pce.12340.
59. Pichersky, E., and Gang, D.R. (2000). Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci.* 5**,** 439-445.
60. Pichersky, E., and Raguso, R.A. (2018). Why do plants produce so many terpenoid compounds? *New Phytol.* 220**,** 692-702. doi: 10.1111/nph.14178.
61. Prasad, B., Mah, D.J., Lewis, A.R., and Plettner, E. (2013). Water oxidation by a cytochrome P450. *PLoS One* 8. doi: 10.1371/journal.pone.0061897.
62. Rosenfeld, H.J., Aaby, K., and Lea, P. (2002). Influence of temperature and plant density on sensory quality and volatile terpenoids of carrot (*Daucus carota* L.) root. *J. Sci. Food Agric.* 82**,** 1384-1390. doi: 10.1002/jsfa.1200.
63. Rosenfeld, H.J., Vogt, G., Aaby, K., and Olsen, E. (2004). Interaction of terpenes with sweet taste in carrots (*Daucus carota* L.). *Advan Veg Breed* 377-386. doi: 10.17660/ActaHortic.2004.637.47.
64. Sanderson, M.A., Adler, P.R., Boateng, A.A., Casler, M.D., and Sarath, G. (2006). Switchgrass as a biofuels feedstock in the USA. *Can. J. Plant Sci.* 86**,** 1315-1325. doi: 10.4141/P06-136.
65. Sanderson, M.A., and Wolf, D.D. (1995). Switchgrass biomass composition during morphological development in diverse environments. *Crop Sci.* 35**,** 1432-1438. doi: 10.2135/cropsci1995.0011183X003500050029x.
66. Senalik, D., and Simon, P.W. (1986). Relationship between oil ducts and volatile terpenoid content in carrot roots. *Am. J. Bot.* 73**,** 60-63. doi: 10.2307/2444277.
67. Shirin, A.P.R., and Prakash, J. (2010). Chemical composition and antioxidant properties of ginger root (*Zingiber officinale*). *J Med Plant Res* 4**,** 2674-2679. doi: 10.5897/Jmpr09.464.
68. Simon, P.W., Peterson, C.E., and Lindsay, R.C. (1980). Correlations between sensory and objective parameters of carrot flavor. *J. Agric. Food Chem.* 28**,** 559-562. doi: 10.1021/jf60229a041.
69. Stewart, A. (2013). *The Drunken Botanist: The Plants That Create the World's Great Drinks*.
70. Sy, A., Timmers, A.C.J., Knief, C., and Vorholt, J.A. (2005). Methylotrophic metabolism is advantageous for Methylobacterium extorquens during colonization of *Medicago truncatula* under competitive conditions. *Appl. Environ. Microbiol.* 71**,** 7245-7252. doi: 10.1128/Aem.71.11.7245-7252.2005.
71. Sykes, V.R., Allen, F.L., Mielenz, J.R., Stewart, C.N., Windham, M.T., Hamilton, C.Y., Rodriguez, M., and Yee, K.L. (2016). Reduction of ethanol yield from switchgrass infected with rust caused by *Puccinia emaculata*. *Bioenergy Research* 9**,** 239-247. doi: 10.1007/s12155-015-9680-4.

72. Tassou, C.C., Chorianopoulos, N.G., Skandamis, P.N., and Nychas, G.J.E. (2012). Herbs, spices and their active components as natural antimicrobials in foods. *Woodhead Publ Food S* 228**,** 17-50.
73. Theis, N., and Lerdau, M. (2003). The evolution of function in plant secondary metabolites. *Int. J. Plant Sci.* 164**,** S93-S102. doi: 10.1086/374190.
74. Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P., and Osbourn, A. (2014). Triterpene Biosynthesis in Plants. *Annual Review of Plant Biology, Vol 65* 65**,** 225-257. doi: 10.1146/annurev-arplant-050312-120229.
75. Tholl, D., and Lee, S. (2011). Terpene Specialized Metabolism in *Arabidopsis thaliana*. *Arabidopsis Book* 9**,** e0143. doi: 10.1199/tab.0143.
76. Vaughan, M.M., Christensen, S., Schmelz, E.A., Huffaker, A., Mcauslane, H.J., Alborn, H.T., Romero, M., Allen, L.H., and Teal, P.E.A. (2015). Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. *Plant Cell Environ* 38**,** 2195-2207. doi: 10.1111/pce.12482.
77. Vickers, C.E., Gershenzon, J., Lerdau, M.T., and Loreto, F. (2009). A unified mechanism of action for volatile isoprenoids in plant abiotic stress. *Nat. Chem. Biol.* 5**,** 283-291. doi: 10.1038/nchembio.158.
78. Whipps, J.M. (2001). Microbial interactions and biocontrol in the rhizosphere. *J. Exp. Bot.* 52**,** 487-511. doi: 10.1093/jexbot/52.suppl_1.487.
79. Wullschleger, S.D., Davis, E.B., Borsuk, M.E., Gunderson, C.A., and Lynd, L.R. (2010). Biomass production in switchgrass across the United States. *Agron. J.* 102**,** 1158-1168. doi: 10.2134/agronj2010.0087.
80. Xiao, Y.M., Savchenko, T., Baidoo, E.E.K., Chehab, W.E., Hayden, D.M., Tolstikov, V., Corwin, J.A., Kliebenstein, D.J., Keasling, J.D., and Dehesh, K. (2012). Retrograde signaling by the plastidial metabolite MEcPP regulates expression of nuclear stress-response genes. *Cell* 149**,** 1525-1535. doi: 10.1016/j.cell.2012.04.038.
81. Yahyaa, M., Bar, E., Dubey, N.K., Meir, A., Davidovich-Rikanati, R., Hirschberg, J., Aly, R., Tholl, D., Simon, P.W., Tadmor, Y., Lewinsohn, E., and Ibdah, M. (2013). Formation of norisoprenoid flavor compounds in carrot (*Daucus carota* L.) roots: characterization of a cyclic-specific carotenoid cleavage dioxygenase 1 Gene. *J. Agric. Food Chem.* 61**,** 12244-12252. doi: 10.1021/jf404085k.
82. Yahyaa, M., Tholl, D., Cormier, G., Jensen, R., Simon, P.W., and Ibdah, M. (2015). Identification and characterization of terpene synthases potentially involved in the formation of volatile terpenes in carrot (*Daucus carota* L.) Roots. *J. Agric. Food Chem.* 63**,** 4870-4878. doi: 10.1021/acs.jafc.5b00546.
83. Yang, J., Kloepper, J.W., and Ryu, C.M. (2009). Rhizosphere bacteria help plants tolerate abiotic stress. *Trends Plant Sci.* 14**,** 1-4. doi: 10.1016/j.tplants.2008.10.004.
84. Zhang, Y.X., Ruyter-Spira, C., and Bouwmeester, H.J. (2015). Engineering the plant rhizosphere. *Curr. Opin. Biotechnol.* 32**,** 136-142. doi: 10.1016/j.copbio.2014.12.006.

CHAPTER II. BIOSYNTHESIS AND EMISSION OF STRESS-INDUCED VOLATILE TERPENES IN ROOTS AND LEAVES OF SWITCHGRASS

In review: Frontiers in Plant Science

Contributions:

Andrew Muchlinski: Transcriptome assembly, RNA-seq analysis, gene identification/annotation, recombinant protein expression, nickel affinity chromatography, terpene synthase in vitro assays (SPME and solvent extraction), analysis and annotation of enzyme products by GC-MS, plant management and experimental design, hexane extractions of root tissues, root phytohormone treatments, relative localization of TPSs on chromosomes, phylogenetic tree reconstruction, amino acid alignments, hierarchical cluster analysis of TPS expression, manuscript writing and figure preparation.

Other contributions:

Xinlu Chen, Research Associate at the University of Tennessee, Knoxville: performed RNA extractions and qRT-PCR, analysis of phytohormone and herbivore treatments in leaves, assisted with figure development; Tobias Köllner, PI at the Max Plank Institute for Chemical Ecology, Germany: assisted with preliminary phylogenetic analysis; Kyle Pelot, PhD student at the University of California, Davis: assisted with gene identification; Philipp Zerbe, Assistant Professor at the University of California, Davis: assisted with gene identification; Meredith Ruggiero, former undergraduate student in Tholl Lab: assisted with TPS enzyme assays and plant maintenance; LeMar Callaway III, former undergraduate student in Tholl Lab: assisted with TPS enzyme assays and plant maintenance; Suzanne Laliberte, former undergraduate student in Tholl Lab: assisted with TPS enzyme assays and plant maintenance; Feng Chen, Professor at the University of Tennessee, Knoxville: assisted in manuscript review; Dorothea Tholl, Professor in

the Dept. of Biology at Virginia Tech: assisted with experimental design, manuscript preparation and review.

ABSTRACT

Switchgrass (*Panicum virgatum* L.), a perennial C4 grass, represents an important species in natural and anthropogenic grasslands of North America. Its resilience to abiotic and biotic stress has made switchgrass a preferred bioenergy crop. However, little is known about the mechanisms of resistance of switchgrass against pathogens and herbivores. Volatile compounds such as terpenes have important activities in plant direct and indirect defense. Here we show that switchgrass leaves emit blends of monoterpenes and sesquiterpenes upon feeding by the generalist insect herbivore *Spodoptera frugiperda* (fall armyworm) and in a systemic response to the treatment of roots with defense hormones. Belowground application of methyl jasmonate also induced the release of volatile terpenes from roots, whereas this response was not observed upon aboveground feeding by *S. frugiperda*. To correlate the emission of terpenes with the expression and activity of their corresponding biosynthetic genes, we identified a gene family of 44 monoterpene and sesquiterpene synthases (mono- and sesqui-TPSs) of the type-a, b, g and e subfamilies, of which 32 TPSs were found to be functionally active in vitro. The TPS genes are distributed over most of the nine chromosomes of the K and N sub-genomes with a higher abundance and clustering occurring on three chromosomes. Synteny analysis revealed syntenic networks for approximately 30-40% of the switchgrass TPS genes in the genomes of *Panicum hallii*, *Setaria viridis* and *Sorghum bicolor* suggesting shared TPS ancestry in the common progenitor of these grass lineages. Sixteen switchgrass TPS genes were substantially induced upon insect and hormone treatment and the enzymatic products of nine of these genes correlated with compounds of the induced volatile blends. Our results demonstrate complex above and

belowground responses of induced volatile terpene metabolism in switchgrass and provide a framework for more detailed investigations of the function of terpenes in stress resistance in this monocot crop.

KEYWORDS: switchgrass, terpene synthase, volatile, herbivory, defense

INTRODUCTION

Switchgrass (*Panicum virgatum L.*, Poaceae) is a native warm-season C4 perennial grass common to natural and anthropogenic grasslands in North America. Characteristic of the Tallgrass Prairie, switchgrass is considered an important species for sustaining natural prairie biodiversity (Sanderson et al., 2006). Used mostly for forage since the 1950s, more intensive breeding of switchgrass began over 20 years ago to develop the species as an herbaceous model species for biofuel feedstock development (Casler et al., 2011). Major advantages for cultivating switchgrass are its resilience to extreme weather conditions, capability of growing on marginal soils, and a high cellulosic content (Vogel, 2004). Switchgrass also exhibits considerable resistance to pests and diseases (Parrish and Fike, 2005). With an increase in cultivation, growing interest has focused on elucidating the resistance mechanisms of switchgrass as well as engineering more resistant varieties. However, surprisingly little is still known about the modes of pathogen and pest defense in this species.

Plants deploy a biosynthetic and structurally diverse mosaic of specialized or secondary metabolites for chemical defense (Dudareva et al., 2004). Terpenes constitute the majority of such metabolites with important defensive activities. For instance, non-volatile triterpenes are potent growth inhibitors of fungal pathogens (Osbourn, 1996). By contrast, low molecular weight 10-carbon monoterpenes and 15-carbon sesquiterpenes are emitted by plants as volatile compounds

and serve important roles in direct defenses against pathogens and herbivores or function indirectly by the attraction of parasitoids or intra- and interplant priming (Turlings et al., 1990; Dicke, 1994; Kost and Heil, 2006; Köllner et al., 2008; Vaughan et al., 2013; Erb et al., 2015).

The formation of terpenes in plants is catalyzed by enzymes of the terpene synthase superfamily (TPSs). TPS enzymes convert 10-carbon and 15-carbon *cis*- or *trans*-isoprenyl diphosphates such as GDP (geranyl diphosphate), NDP (neryl diphosphate), (*E*,*E*)-FDP (farnesyl diphosphate), or (*Z*,*Z*)-FDP into monoterpenes or sesquiterpenes, respectively (Tholl and Lee, 2011). TPS genes often undergo species specific divergence and duplications resulting in terpene metabolic plasticity and adaptations (Pichersky and Gang, 2000). The structural diversity and biosynthetic evolution of terpene secondary metabolites has been studied extensively in crops including grasses such as maize, rice, and sorghum (Chen et al., 2011; Boutanaev et al., 2015; Block et al., 2019). Terpene related defenses have been well described in these monocot crops and reveal diverse chemical mechanisms for resistance against above- and belowground stressors. For example, the sesquiterpene (*E*)-β-caryophyllene, one of the major VOCs released by maize leaves and roots, is involved in indirect defense by attracting parasitoids of herbivores and entomopathogenic nematodes (Turlings et al., 1990; Rasmann et al., 2005; Köllner et al., 2008). Monoterpenes have also been implicated in defensive roles; for example linalool confers resistance against rice bacterial blight caused by *Xanthomonas oryzae* (Taniguchi et al., 2014). More recently, a rice (*S*)-limonene synthase (*OsTPS19*) was shown to be involved in direct defense against the blast fungus *Magnaporthe oryzae* (Chen et al., 2018).

In contrast to these findings in highly domesticated grasses, the biosynthesis and dynamics of terpenes in switchgrass have not been fully investigated, in part because of its complex genetic background. Lowland ecotypes are allotetraploid (2n = 4x = 36) while upland cultivars are

frequently octoploid (2n = 8x = 72). Recent transcriptional analysis of defense responses to green bug herbivory (*Schizaphis graminum*, Aphididae) in switchgrass leaves revealed a global transcriptional remodeling resulting in increased ROS production and upregulation of genes with predicted terpene synthase function (Donze-Reiner et al., 2017). Moreover, the presence of a few triterpene saponins (C_{30}) (Lee et al., 2009), and the synthesis of diterpenes (C_{20}) related to abiotic stress have been described (Pelot et al., 2018). However, no prior studies have investigated the formation and function of volatile terpenes in this grass. Therefore, we sought to identify and characterize TPS genes from the switchgrass genome and correlate stress induced terpene synthases with compound production in roots and leaves. Particular focus was placed on TPSs that were readily inducible when challenged with a generalist herbivore and the defense related phytohormones methyl jasmonate and salicylic acid with the future goal to investigate these genes for their broad defensive functions against pathogens and herbivores. Results from this study provide further insight into the genetic organization of terpene metabolism in switchgrass and illustrate the metabolic potential of terpene related defenses in perennial polyploid grasses.

RESULTS

Emission of volatile terpenes from leaves in response to insect feeding

To assess whether switchgrass leaves emit volatile compounds upon aboveground herbivory, emissions from switchgrass plants (cv. Alamo) damaged by larvae of *Spodoptera frugiperda* (fall armyworm; FAW) were collected by open headspace sampling and analyzed by gas chromatography-mass spectrometry (GC-MS). We found that FAW treatment induced the emission of nine terpene compounds, which were not detected in plants that only received physical wounding or remained untreated (Fig. 1 and Table 1). Total induced volatile emission from FAW treatment was approximately 2 μg $hr^{-1}g$ FW^{-1} (Table 1). Among the released compounds, the

sesquiterpenes (*E*)-β-caryophyllene and (*E*)-β-farnesene were strongly induced by herbivore damage accounting for ~21% and ~33% of the total VOC emission (Fig. 1 and Table 1). Additional major compounds induced by FAW included the monoterpene (*E*)-β-ocimene, the homoterpene (*E*)-DMNT and the sesquiterpenes β-elemene, α-bergamotene, α-humulene and β-copaene (Fig. 1 and Table 1).

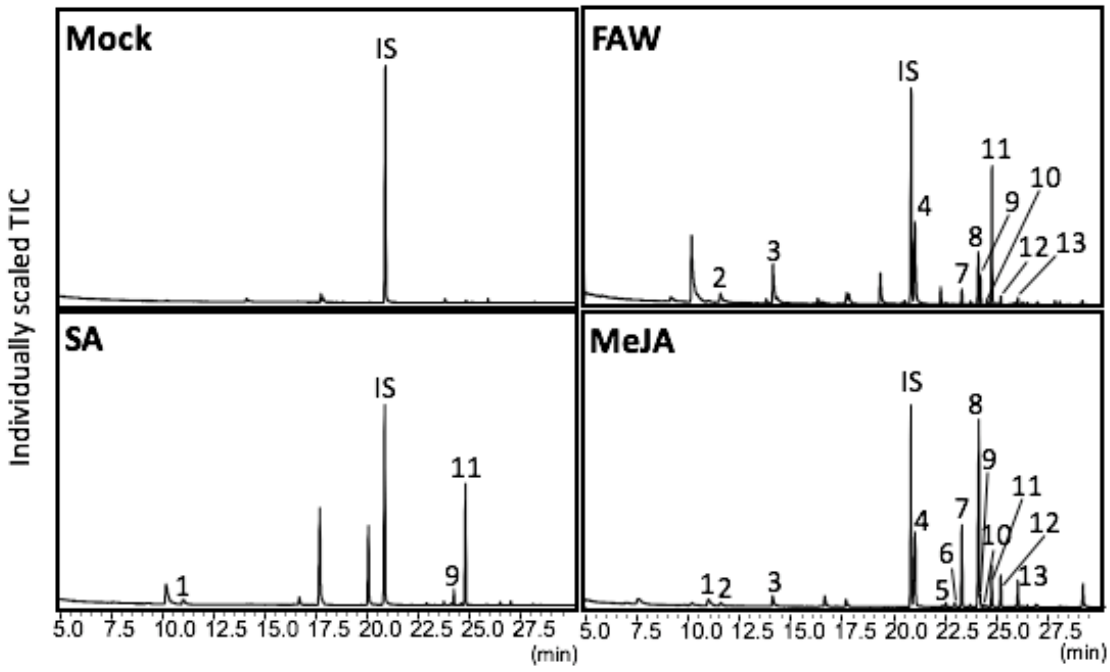


Figure 1. GC-MS analysis of volatile emission from leaves of switchgrass (cv. Alamo) treated with salicylic acid (SA), fall armyworms (FAW) or methyl jasmonate (MeJA). Individual chromatograms are representative of samples analyzed in biological triplicate. Compound identification was based on similarity to library matches (NIST, WILEY) and comparisons to authentic standards. 1: limonene; 2: (*E*)-β-ocimene, 3: (*E*)-DMNT; 4: indole; 5: α-ylangene; 6: elemene isomer; 7: β-elemene; 8: (*E*)-β-caryophyllene; 9: α-bergamotene; 10: unidentified sesquiterpene; 11: (*E*)-β-farnesene; 12: α-humulene and 13: β-copaene.

Table 1. Emission of voltile compounds from switchgrass leaves (cv. Alamo) treated with fall armyworms (FAW), methyl jasmonate (MeJA), salycilic acid (SA) or mock control. Values were obtained in biological triplicate and are reported as mean±SD (ng hr^{-1} g^{-1}. nd = not detected, tr = trace. Compound identification was based on similarity to library matches (NIST, WILEY) and comparisons to authentic standards.

Compound	Control	MeJA	FAW	SA
limonene	nd	89.3 ± 34.7	nd	63.8 ± 8.3
(*E*)-β-ocimene	nd	30.5 ± 11.2	124.7 ± 5.7	nd
(*E*)-DMNT	nd	52.6 ± 27.8	328.2 ± 50.2	nd
α-ylangene	nd	27.2 ± 10.4	nd	nd
elemene isomer	nd	31.3 ± 12.9	nd	nd
β-elemene	nd	633.3 ± 106.2	218.7 ± 46.2	nd
(*E*)-β-caryophyllene	nd	1454.4 ± 248.7	506.5 ± 68.4	6.9 ±1.3
α-bergamotene	nd	61.6 ± 23.5	187.4 ± 48.0	88.5 ± 7.1
unidentified sesquiterpene	nd	15.6 ± 6.2	8.2 ± 3.8	nd
(*E*)-β-farnesene	tr	310.4 ± 153.6	770.4 ± 54.7	772.4 ± 58.1
α-humulene	nd	222.6 ± 24.9	95.1 ± 25.1	nd
β-copaene	nd	188.7 ± 24.4	76.6 ± 19.5	nd
Total emission	**-**	**3117.5 ± 684.5**	**2315.8 ± 321.6**	**929.8 ± 73.5**

Emission of volatile terpenes from roots and leaves upon belowground treatment with methyl jasmonate or salicylic acid

We further determined whether emissions of volatile compounds from switchgrass roots could be induced by root treatment with phytohormones mimicking herbivory or pathogen infection. Different concentrations of methyl jasmonate (MeJA) were tested (0.1 mM, 1 mM, 5 mM) by watering plants directly with each solution. Because of the volatility of MeJA, we expected that the compound diffused further into the substrate at a lower concentration. Volatiles were collected from detached roots using solid phase microextraction (SPME) and analyzed by GC-MS. Concentrations of 1 mM and 5 mM MeJA caused a similar relative release of sesquiterpene compounds from the root tissue (shown for 5 mM treatment; Fig. 2), while no volatiles were induced upon treatment with 0.1 mM MeJA. Of the seven identified compounds, (*E*)-β-

caryophyllene was the most abundant (~43% of total), while cycloisosativene, β-elemene, α-humulene, α-selinene, germacrene D and δ-cadinene were present at low levels. We also applied salicylic acid (SA) at a concentration of 5 mM; however, no release of sesquiterpenes was observed from root tissue. We further found two monoterpenoids, camphor and borneol, to be released from root tissue of untreated plants. Emissions of these compounds were reduced by MeJA and SA treatments although not at statistically significant levels (Supplementary Fig. S1).

We also tested whether a drench with MeJA and SA at 5 mM could induce volatile emissions in aboveground tissues. Treatment with MeJA strongly induced volatile emission from leaves compared to mock controls, with twelve compounds identified and a total emission rate of approximately 3 μg hr^{-1}g FW^{-1} (Fig. 1 and Table 1). Major induced compounds were (*E*)-*β*-caryophyllene and β-elemene accounting for ~46% and ~20% of total volatile emissions (Fig. 1 and Table 1). Other minor compounds included limonene, (*E*)-β-ocimene, (*E*)-DMNT, α-ylangene, α-bergamotene, α-humulene, (*E*)-β-farnesene and β-copaene (Fig. 1 and Table 1). Two additional putative sesquiterpenes were also emitted, however these compounds could not be further identified based on available standards. Treatment with SA induced the emission of four terpene compounds with (*E*)-β-farnesene accounting for 83% of total emissions. Trace amounts of limonene, (*E*)-β-caryophyllene and α-bergamotene were detected, which were not observed in the untreated controls (Fig. 1 and Table 1).

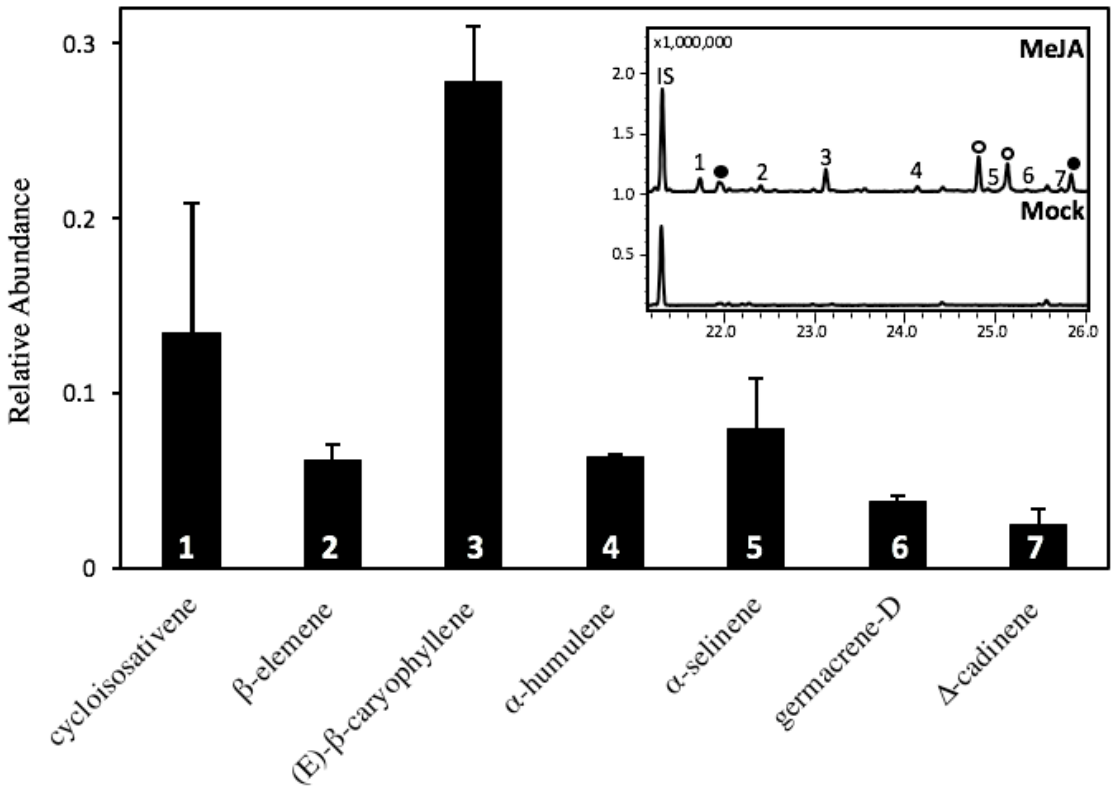


Figure 2. SPME-GC-MS analysis of root sesquiterpene emission following MeJA treatment. Volatile compounds were analyzed in triplicate from detached pooled root material of 6-week old plants. No sesquiterpenes were identified in the mock treatment (insert). IS: internal standard 1-bromodecane; shaded circles indicate putative terpene compounds and open circles indicate non-terpene hydrocarbons (insert). Samples were normalized to the internal standard and grams fresh. Compound identification was based on similarity to library matches (NIST, WILEY) and authentic standards.

Genome-wide Identification of putative terpene synthases in switchgrass

Based on the inducible emission of diverse volatile terpenes from switchgrass roots and leaves, we sought to identify the TPS genes responsible for their formation. Following a genome-wide search of the switchgrass draft genome v.1, we originally identified 144 putative TPS gene models. Of these putative gene models, 108 were confirmed in the draft genome v.4, with 74 putative full-length mono-, sesqui-, and di-TPS genes identified. Manual sequence curation through multiple

sequence alignments and comparison to genomic and transcriptomic data resulted in the identification of 44 putative full-length mono-TPS and sesqui-TPS genes (Table 2, Supplementary Table 1, Supplementary Fig.S2). Identified di-TPS genes (30 in total) were previously reported and therefore not included in this study (Pelot et al., 2018).

Alignment and phylogenetic analysis of amino acid sequences from the mono-TPSs and sesqui-TPSs together with select TPSs from maize, rice, sorghum, tomato, and snapdragon showed that 35 members belong to the TPS type-a clade (Fig. 3). In addition, five proteins aligned to the TPS-g sub-family and three clustered with the proteins in the TPS-b subfamily. Only *Pv*TPS15 (TPS-e) was predicted to be involved in volatile formation outside of the TPS-a, b and g subfamilies. Like in other plant TPS proteins, switchgrass TPSs of the TPS-a, b, e and g subfamilies carry the conserved aspartate-rich 'DDXXD' motif and the less conserved 'NSE/DTE' motif in the C-terminal α-domain (Tholl et al., 2011).

When we examined the relative chromosomal position of the identified TPS genes, we found that 22 genes are distributed across the 9 chromosomes in the switchgrass sub-genome K with highest abundance of genes occurring on chromosomes 1K, 6K, and 9K (Fig. 4; Table 2). In sub-genome N, we identified the relative location of 20 genes with highest abundance on chromosomes 1N, 6N, and 9N (Fig. 4; Table 2). Several genes are positioned in loose gene clusters throughout the genome (Fig. 4; Table 2). The relative positions of *PvTPS02* and *PvTPS07* could not be determined based on incomplete genomic data (Table 2).

Investigation of syntenic orthologous genes between the two switchgrass sub-genomes identified networks between eight genes on sub-genome K and 10 genes on sub-genome N (including two putative diterpene-TPSs: 3K6400900; 3NG171200) (Fig. 4, Supplementary Fig. S3; Supplementary Table 1). Comparisons between the genomes of switchgrass and sorghum

showed that 13 switchgrass TPS loci have syntenic orthologs on six of the 10 sorghum chromosomes (Fig. 4, Supplementary Fig. S3; Supplementary Table 1). Several of these switchgrass TPS genes also occur in syntenic gene networks with genomes of the more closely related grasses *Setaria viridis* and *Panicum hallii* suggesting conserved genomic regions in TPS gene evolution in these species.

Table 2. Identified mono- and sesqui*TPS* gene models in the switchgrass (cv. AP13) genome. Genomic coordinates were determined based on draft genome data available in Phytozome (phytozome.jgi.doe.gov). Putative biosynthetic gene clusters are annotated based on loci proximity.

Locus ID	Designation	Protein Length	Genomic Coordinates	Sub-family
Pavir.1KG024200	PvTPS28	520	Chr01K:2898821..2904002	tps-g
Pavir.1KG213700	PvTPS83	607	Chr01K:30203541..30207022	tps-a
Pavir.1KG026700	PvTPS13	557	Chr01K:3056228..3059326	tps-g
Pavir.1KG250000	PvTPS03	612	Chr01K:41080529..41082963	tps-a
Pavir.1KG359700	PvTPS33	574	Chr01K:60696889..60702775	tps-a
Pavir.1NG048300	PvTPS106	551	Chr01N:2491525..2496085	tps-g
Pavir.1NG047800	PvTPS27	552	Chr01N:2543130..2546026	tps-g
Pavir.1NG173900	PvTPS04	607	Chr01N:30384431..30393731	tps-a
Pavir.1NG245800	PvTPS26	583	Chr01N:56485952..56491693	tps-a
Pavir.2KG086400	PvTPS01	548	Chr02K:11130096..11132928	tps-a
Pavir.2KG150300	PvTPS14	552	Chr02K:19642302..19645362	tps-a
Pavir.2KG163100	PvTPS11	554	Chr02K:22460036..22464132	tps-a
Pavir.2NG122900	PvTPS85	515	Chr02N:18521205..18525163	tps-a
Pavir.2NG123000	PvTPS19	555	Chr02N:18533973..18537033	tps-a
Pavir.3KG098200	PvTPS36	592	Chr03K:8855955..8859293	tps-a
Pavir.3KG098400	PvTPS81	621	Chr03K:8870023..8873401	tps-a
Pavir.3KG102800	PvTPS08	616	Chr03K:9179806..9183931	tps-a
Pavir.5KG204300	PvTPS79	567	Chr05K:26556249..26559391	tps-a
Pavir.5KG258500	PvTPS09	491	Chr05K:38083739..38087664	tps-a
Pavir.6KG122500	PvTPS69	549	Chr06K:16365911..16369794	tps-a
Pavir.6KG141400	PvTPS10	259	Chr06K:23665655..23669023	tps-a
Pavir.6KG047500	PvTPS109	582	Chr06K:6526582..6531758	tps-a
Pavir.6KG066200	PvTPS17	549	Chr06K:9019105..9024173	tps-a
Pavir.6NG075100	PvTPS18	545	Chr06N:14008882..14012062	tps-a
Pavir.6NG075300	PvTPS16	626	Chr06N:14020511..14024883	tps-a
Pavir.6NG075700	PvTPS20	549	Chr06N:14043871..14051763	tps-a
Pavir.6NG192300	PvTPS94	545	Chr06N:36030001..36038500	tps-a
Pavir.6NG135600	PvTPS06	544	Chr06N:58557205..58561380	tps-a
Pavir.7NG405000	PvTPS15	618	Chr07N:63119161..63123381	tps-e
Pavir.8KG329300	PvTPS62	597	Chr08K:67429955..67433773	tps-a

Pavir.8NG101200	PvTPS101	549	Chr08N:15863965..15870973	tps-a
Pavir.9KG112900	PvTPS12	579	Chr09K:10384938..10389576	tps-b
Pavir.9KG454600	PvTPS56	570	Chr09K:66968728..66972114	tps-a
Pavir.9KG443500	PvTPS54	576	Chr09K:67815040..67822909	tps-a
Pavir.9KG442600	PvTPS55	552	Chr09K:67885565..67896951	tps-a
Pavir.9KG486800	PvTPS73	546	Chr09K:72969142..72975315	tps-b
Pavir.9NG517400	PvTPS50	551	Chr09N:86982070..86991194	tps-a
Pavir.9NG529400	PvTPS71	549	Chr09N:88336547..88343612	tps-a
Pavir.9NG574000	PvTPS05	560	Chr09N:92237810..92240737	tps-a
Pavir.9NG595500	PvTPS53	606	Chr09N:94252282..94256051	tps-a
Pavir.9NG595900	PvTPS104	467	Chr09N:94271551..94277158	tps-a
Pavir.9NG693900	PvTPS52	601	Chr09N:99983657..99989362	tps-b
Pavir.J27731	PvTPS07	502	contig312548:164..1302	tps-a
Pavir.J691000	PvTPS02	515	scaffold_732:11619..13941	tps-g

Biochemical characterization of mono- and sesquiterpene synthases from switchgrass

To determine the in vitro function of the 44 identified TPS genes, open reading frames were synthesized and cloned into the bacterial pET28b expression vector. The recombinant proteins were expressed in *E. coli* and protein lysates tested for TPS activity with GDP and (*E,E*)-FDP as substrates. We expected many TPSs in the subfamily-a (Fig. 3) to function as sesqui-TPSs. Indeed, 20 recombinant TPS proteins in this family produced one or more sesquiterpene olefins, among them (*E*)-β-caryophyllene, (*E*)-*β*-farnesene, and other common plant sesquiterpenes (Figs. 3 and 5). All of these proteins except *Pv*TPS83 did not carry a plastidial transit peptide indicating that they are likely to function in the cytosolic compartment. *Pv*TPS02 was the only TPS protein found in the g-subfamily to exhibit sesquiterpene synthase activity in vitro. However, since a plastidial targeting sequence typical of subtype-g TPSs has been predicted for this protein, its function as a sesqui-TPS in vivo might be limited.

Twelve TPS proteins distributed over the TPS-a, -b, -g, and –e subfamilies functioned as monoterpene synthases in vitro (Figs. 3 and 6). *Pv*TPS04 produced a mixture of monoterpenes from GDP with α-terpinolene and borneol as major products (Fig. 6). *Pv*TPS36 and *Pv*TPS56 converted GDP into multiple volatile products with predominantly limonene and α-terpineol as

the major products respectively (Fig. 6). The remaining enzymes produced either linalool (*Pv*TPS12, *Pv*TPS13, *Pv*TPS15, *Pv*TPS27, *Pv*TPS52 and *Pv*TPS71) or geraniol (*Pv*TPS53 and *Pv*TPS101) (Fig. 6). *Pv*TPS13 and *Pv*TPS15 also converted (*E,E*)-FDP into nerolidol (Supplementary Fig. S4); however, this activity might be limited in vivo because of the predicted plastidial localization of these proteins. On the contrary, no plastidial transit peptides were predicted for *Pv*TPS12, *Pv*TPS56, *Pv*TPS71, and *Pv*TPS101, which questions their function as monoterpene synthases in vivo.

Only trace amounts of compounds were detected for recombinant proteins encoded by *PvTPS07*, *PvTPS62*, *PvTPS81*, and *PvTPS106*. In addition, no substantial enzymatic activity was found for eight proteins (*Pv*TPS18, *Pv*TPS26, *Pv*TPS28, *Pv*TPS33, *Pv*TPS54, *Pv*TPS73, *Pv*TPS85, *Pv*TPS104), which is in accordance with the presence of several deletions and/or insertions in the open reading frames of the corresponding genes (Supplementary Fig. S2). Sequence truncations were furthermore found at the N-terminus and C-terminus of the functionally active *Pv*TPS09 and *Pv*TPS02 proteins, respectively (Supplementary Fig. S2).

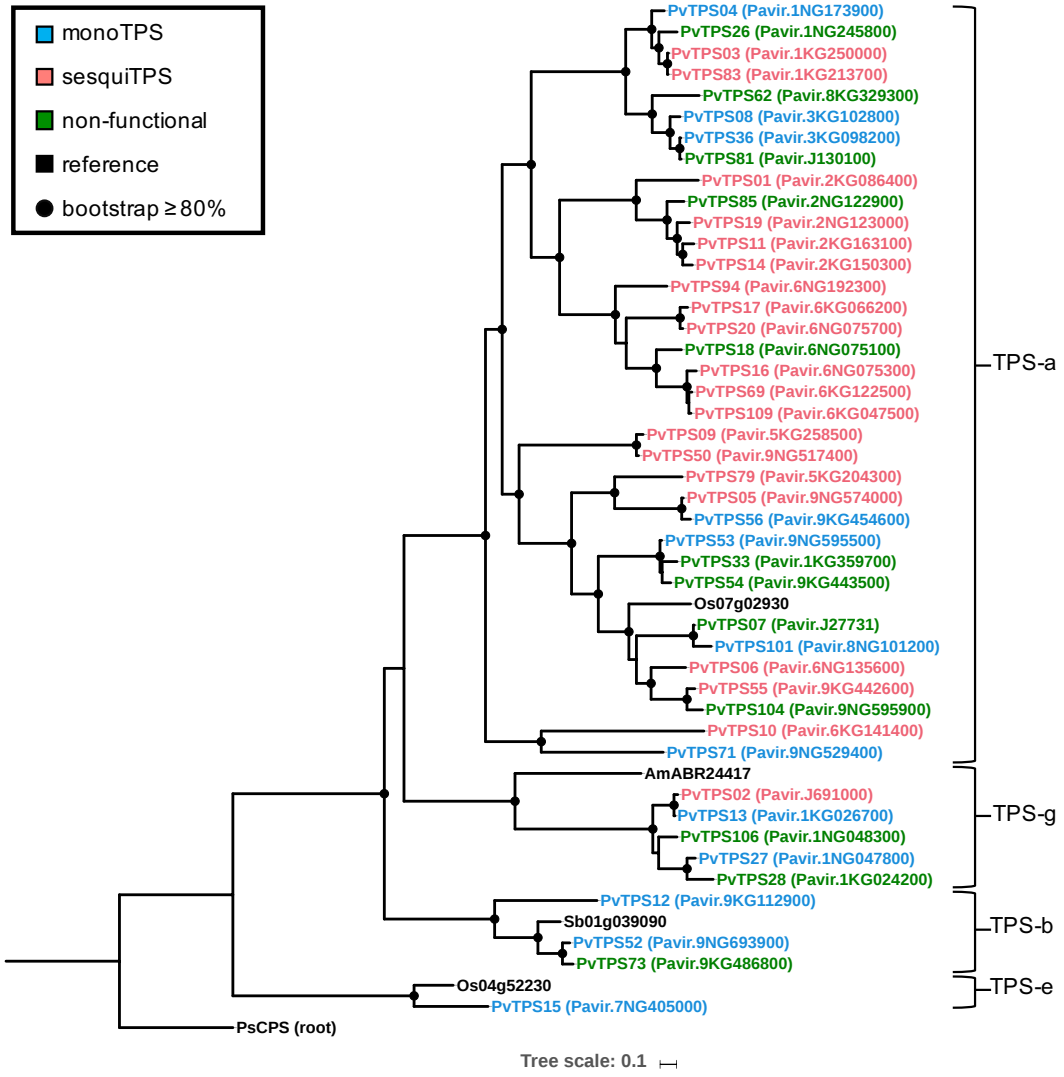


Figure 3. Maximum-likelihood phylogenetic tree based on the MAFFT alignment of 44 characterized mono- and sesquiterpene synthases in switchgrass and reference TPSs from rice, sorghum and snapdragon. Shaded circles represent branches with bootstrap support ≥ 80% (bootstrap replications = 500).

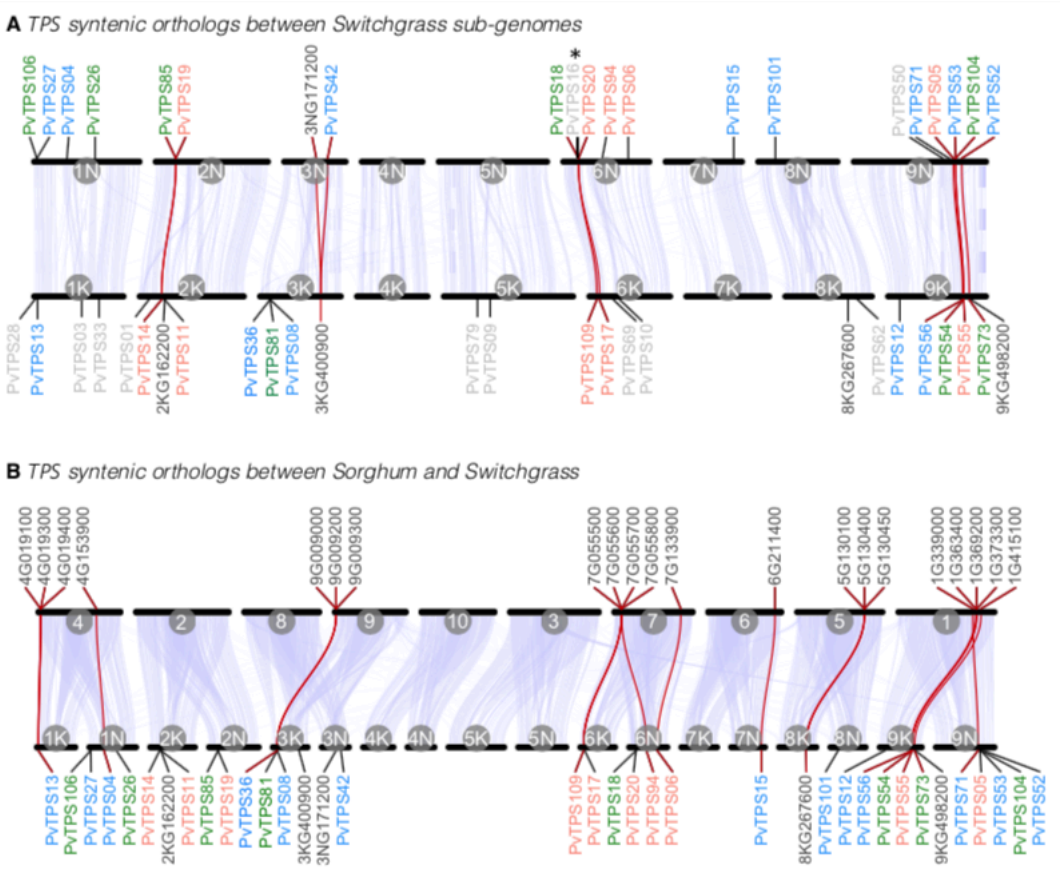


Figure 4. Synteny and orthology of TPS gene family members between *P. virgatum* sub-genomes (A) and between *P. virgatum* and *S. bicolor* (B). In each plot, the genomic positions of chromosomes are plotted along the x-axis. The scale is independent for each genome and chromosomes are ordered to maximize synteny with *P. virgatum*. Chromosome IDs are printed on the inside of each line segment. Syntenic blocks between each pair of (sub)genomes are presented as light blue polygons. All characterized *P. virgatum* mono-TPS and sesqui-TPS genes (labeled PvTPS, color code according to Fig. 3) and annotated (functional or non-functional) TPS genes (labeled with the alpha-numeric Phytozome gene IDs) that occur in orthologous gene networks of any of the analyzed genomes (this figure and Supplementary Fig. S3) are shown. Characterized *P. virgatum* mono-TPS and sesqui-TPS genes that do not occur in orthologous gene networks are printed in light grey in (A). The positions of genes are indicated by a straight line. An orange line indicates that there is an ortholog in the alternative genome. *, PvTPS16 has only been partially annotated in the *P. virgatum* v.4 genome.

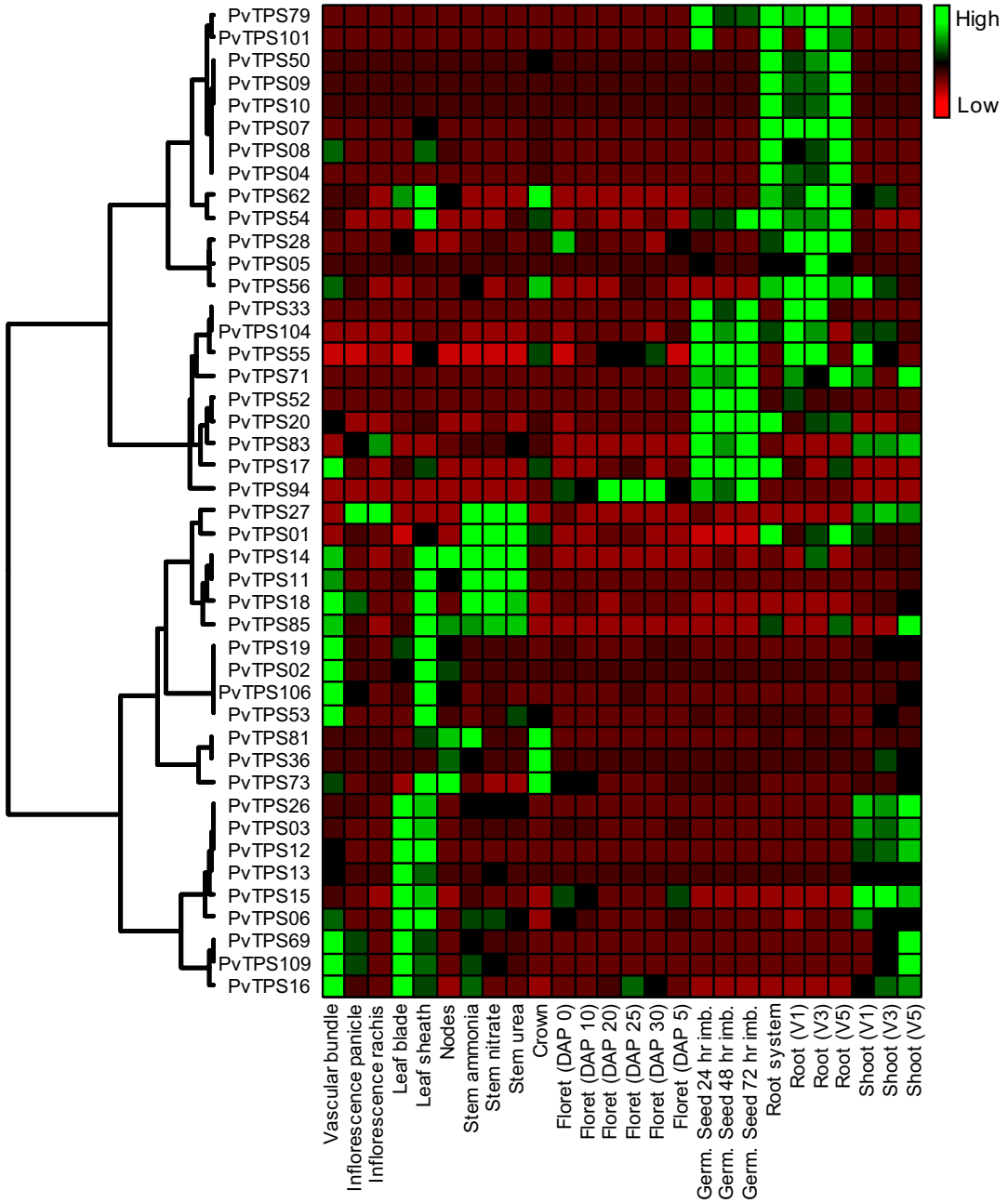


Figure 5. Hierarchical cluster analysis of *TPS* expression profiles across switchgrass ("AP13") tissues. Heat map compares relative transcript abundance for mono and sesqui*TPS* genes in FPKM (Fragments Per Kilobase of transcript per Million mapped reads) across 26 gene expression data sets (https://phytozome.jgi.doe.gov/).

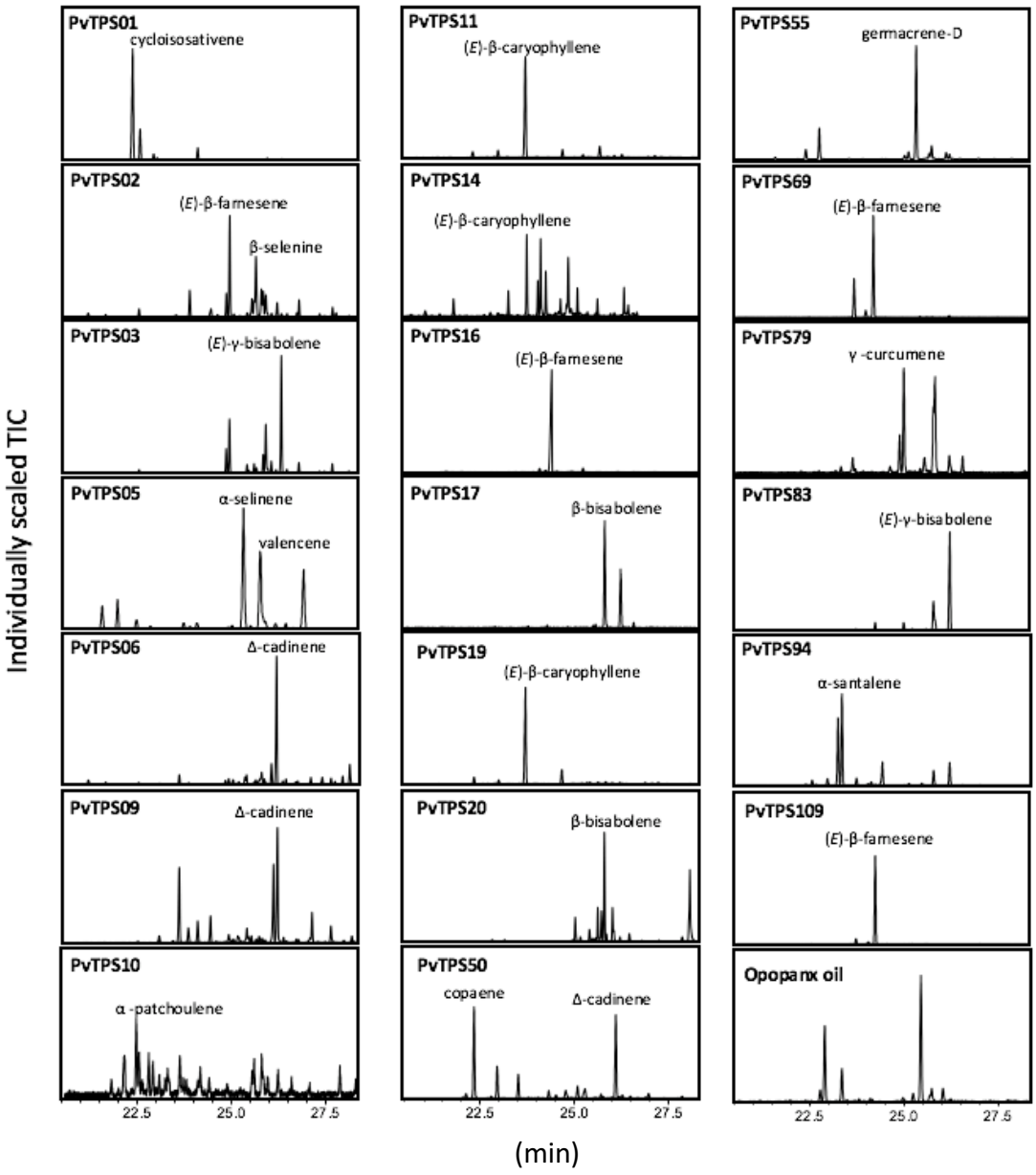


Figure 6. SPME-GC-MS analysis of volatile products produced by recombinant TPSs with (*E*,*E*)-FPP as a substrate. Compound identification was based on similarity to library matches (NIST, WILEY), authentic standards and comparison to Opopanax oil (*Commiphora guidotti*).

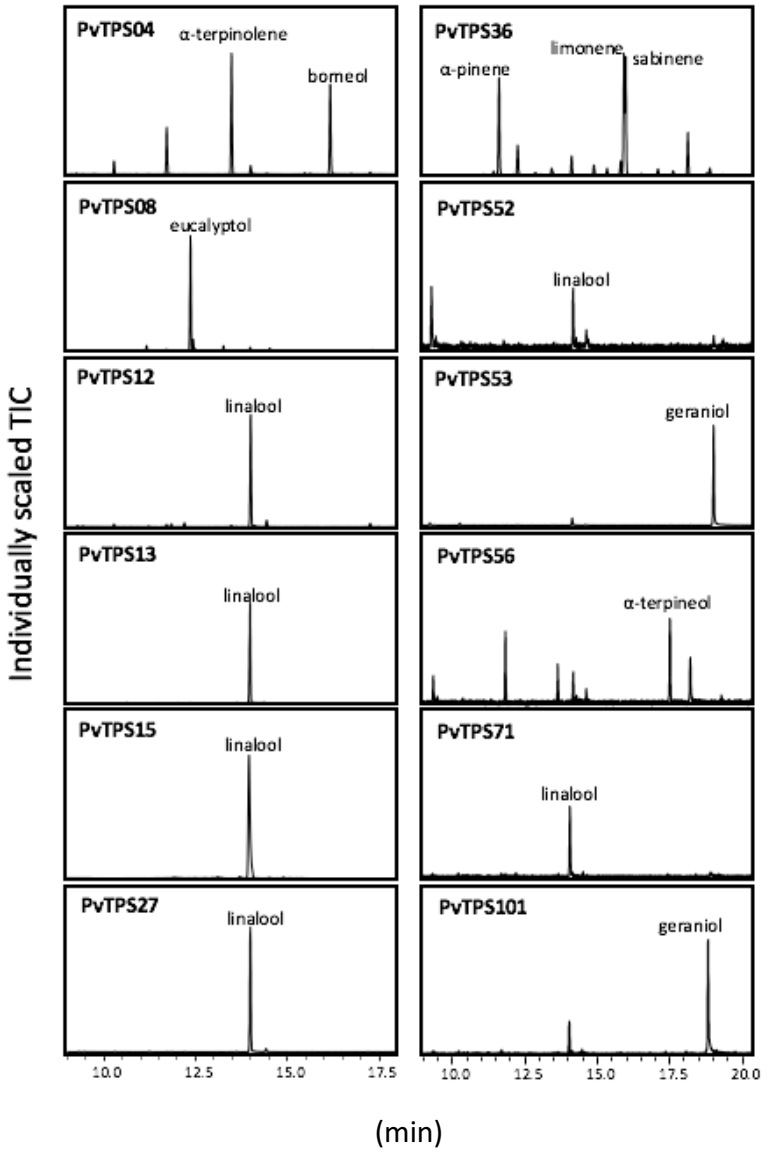


Figure 7. SPME-GC-MS analysis of volatile products produced by truncated recombinant TPSs with GPP as a substrate. Compound identification was based on similarity to library matches (NIST, WILEY), authentic standards and comparison to Opopanax oil.

Expression analysis of PvTPSs following treatment with FAW, MeJA and SA

Global expression patterns for all 44 TPS genes were analyzed by hierarchical cluster analysis based on publically available data (https://phytozome.jgi.doe.gov/). We found specific patterns of transcript abundance in vascular tissue, leaf blade and sheath tissues as well as roots and germinating seeds (Fig. 7). Transcripts included those of the 12 genes that lack in vitro functional activity. There was little overlap in expression between above- and belowground tissues indicating gene specific adaptations in these tissues. Despite the observed transcriptional patterns, we were unable, with the exception of borneol, to detect volatile terpenes in leaves and roots of the Alamo cultivar under constitutive conditions.

To determine whether correlations between transcript abundance and volatile terpene products could be established in response to treatment with FAW, MeJA and SA, we selected multiple TPS genes for expression analysis by quantitative RT-qPCR (Fig. 8A, Supplementary Fig. S5). In leaves, substantial induction at the transcript level (>10 fold) following herbivory was observed for 12 TPS genes (*PvTPS01*, *PvTPS04*, *PvTPS05*, *PvTPS06*, *PvTPS08*, *PvTPS11*, *PvTPS14*, *PvTPS16*, *PvTPS19*, *PvTPS36*, *PvTPS53*, *PvTPS56*), of which ten genes and *PvTPS12* and *PvTPS15* were equally of more highly induced upon root treatment with MeJA (Fig. 8A, Supplementary Fig. S5). SA-induced expression exceeding that in response to FAW and MeJA treatment was observed for *PvTPS04*, *PvTPS13*, *PvTPS16*, and *PvTPS53*. Highest induction of *TPS* transcript levels in roots was found for 11 genes in response to the application of MeJA (*PvTPS05*, *PvTPS06*, *PvTPS10*, *PvTPS11*, *PvTPS14*, *PvTPS17*, *PvTPS19*, *PvTPS20*, *PvTPS36*, *PvTPS53*, *PvTPS56*) or both MeJA and SA (*PvTPS53*) (Fig. 8A, Supplementary Fig. S5).

For nine TPS genes we were able to identify their enzymatic products as components of the induced volatile blends of leaves and/or roots: The genes encoding (*E*)-β-caryophyllene synthases (*PvTPS11*, *PvTPS14*, *PvTPS19*) showed highest transcript abundance in leaves and roots upon treatment with MeJA (Fig. 8A). (*E*)-β-caryophyllene emissions from both tissues are most likely associated with the activity of these TPSs. Expression of *PvTPS16*, whose recombinant protein produced (*E*)-β-farnesene, was strongly induced by SA treatment in leaves and is likely be associated with (*E*)-β-farnesene emission from this tissue (Figs. 1 and 8A). The gene encoding *Pv*TPS05, which was found to produce *α*-selinene volatiles, was most strongly expressed in roots by application of MeJA matching the detection of this compound from root tissue (Fig. 2, Supplementary Fig. S5). *Pv*TPS06 and *Pv*TPS09 both catalyze the formation of δ-cadinene, which was emitted from roots and leaves upon MeJA treatment (Figs. 1, 2, and 5). Only TPS06 was highly induced by MeJA indicating its likely function in planta (Fig. 8A). Moreover, transcript levels of *PvTPS36* were substantially induced in leaves in response to MeJA application although emission of one of the primary products of the *Pv*TPS36 enzyme, limonene, occurred only at low levels (Figs. 1 and 8). Interestingly, terpene products (cycloisosativene, borneol) associated with two genes (*PvTPS01*, *PvTPS04*), which showed highest expression in leaves upon FAW and/or phytohormone treatment, could only be detected in roots (Fig. 2, Supplementary Figs. S1 and S5).

Some TPS genes with lower levels of induction may contribute to the emission of particular terpenes (e.g. the (*E*)-β-farnesene synthase *PvTPS02*). Compounds produced in vitro by several other TPSs could not be detected or occurred only at trace levels in leaves and roots despite a strong induction of their corresponding genes. For example, 1,8-cineole produced by *Pv*TPS02 was only detected in trace amounts in root tissues. Linalool, the single product of enzymes encoded by *PvTPS12*, *PvTPS13*, *PvTPS15*, and *PvTPS56* (Fig. 6) was neither detected in emissions from

leaves and roots and may be further metabolized upon stress treatment. Other TPSs for which no associations could be established between their enzymatic products and volatile emissions include *Pv*TPS03 ((*E*)-γ-bisabolene synthase), *Pv*TPS10 (α-patchoulene synthase), *Pv*TPS17 and *Pv*TPS20 ((*E*)-β-bisabolene synthases) and TPS53 (geraniol synthase). Emission of germacrene D from roots may be associated with *PvTPS55*, the expression of which was not determined.

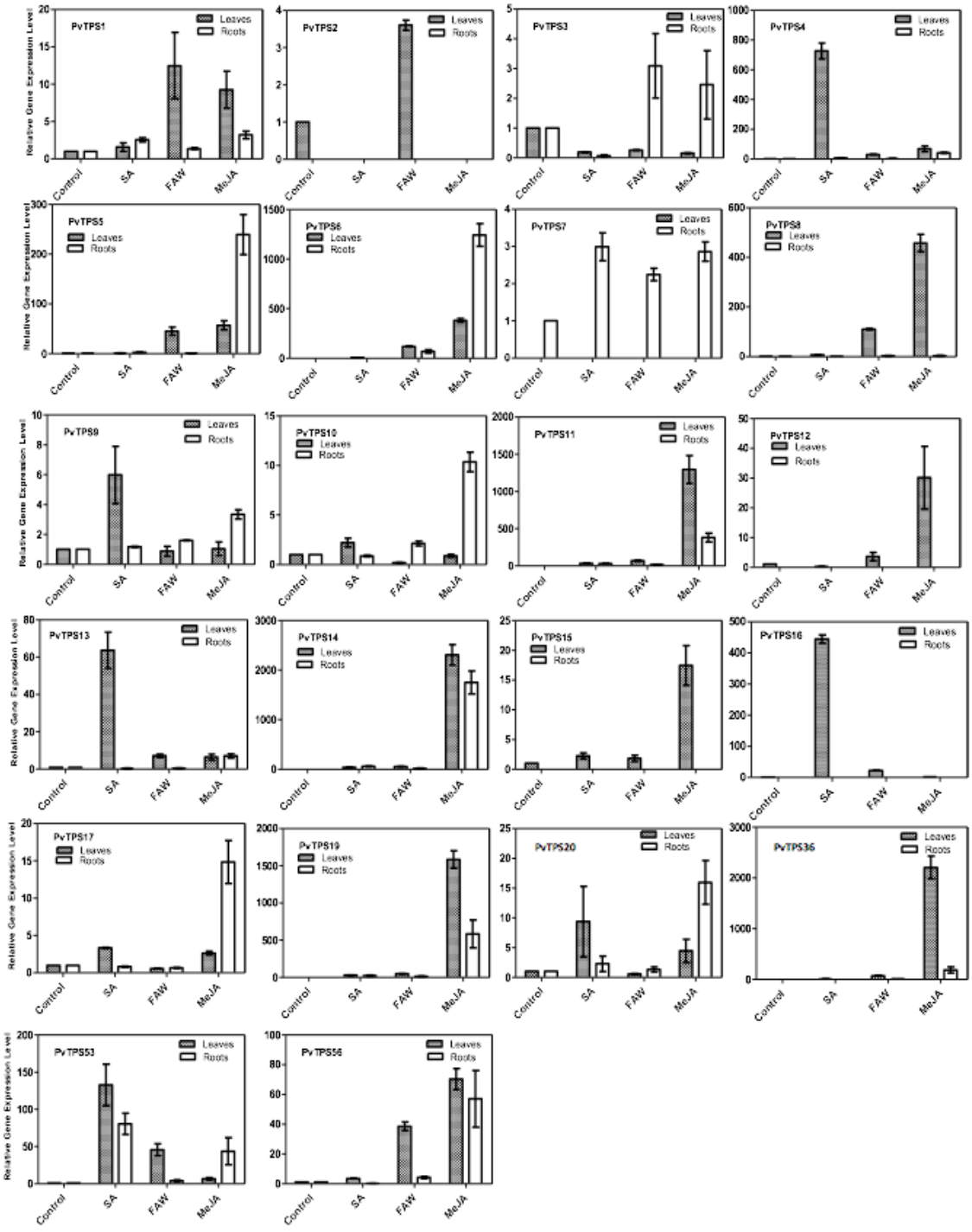


Figure. 8. Gene expression analysis of selected *TPS* genes in switchgrass. Samples were analyzed in biological and technical triplicate and normalized to the expression of the control gene actin. Control expression levels were scaled to 1 for comparison of treatment effects.

Discussion

The switchgrass genome contains a large family of 44 predicted full length mono-TPS and sesqui-TPS genes, of which 32 genes encode functionally active proteins. Sesqui-TPSs belonging to the type-a subfamily make up the majority of this TPS group, while only few mono-TPSs have emerged in the type-a clade or are distributed over the type-b, type-g, and type-e/f clades (Fig. 3). Similar distributions have been shown to occur in the TPS families of rice and sorghum (Chen et al., 2011). Expansions of type-a clades are also common in dicots although typically a higher proportion of mono-TPSs can be found in the type-b and type-g clades of dicot species (Chen et al., 2011; Kulheim et al., 2015)

The size of the switchgrass mono-TPSs and sesqui-TPSs family is almost twice as large as the number of characterized or predicted proteins with mono- or sesqui-TPS activity in maize (Springer et al., 2018). Polyploidy likely contributed to the expansion of the switchgrass TPS gene family, which is in agreement with studies by Hofberger et al. (2015) demonstrating the role of polyploidy events in the diversification and expansion of terpene secondary metabolism. Gene duplication through polyploidization generates gene redundancy eventually increasing functional divergence and allowing species adaption (Wendel, 2000). As an allotetraploid, switchgrass evolved from two diploid ancestors giving rise to two complete sub-genomes (N and K) and functional divergence of *TPS* genes. In *Panicum hallii*, a diploid relative of switchgrass, approximately 32 putative full length TPS genes are annotated (https://phytozome.jgi.doe.gov/), indicating that polyploidization of switchgrass more than doubled the number of *TPS* genes. Polyploidy events in domesticated grasses may not always result in large TPS gene families as has been suggested for wheat (Schmelz et al., 2014). However, in switchgrass, obligate outcrossing and limited breeding have maintained massive phenotypic and adaptive polymorphisms (Casler et

al., 2007), in line with a higher level of diversification in TPS genes. Nevertheless, one-third of the TPS genes we characterized appear to be functionally inactive while several other TPSs might have limited in vivo activity due to their subcellular localization suggesting inactivation and loss of in vivo function for a substantial fraction of the gene family.

A comparison between the switchgrass sub-genomes found that only 35% or 50% of the TPS genes on sub-genome K and N, respectively, have syntenic orthologs on the other sub-genome. This limited synteny indicates sub-genome divergence in TPS gene organization. Syntenic regions include TPS genes with identical functions (*PvTPS14* and *PvTPS19* – (*E*)-β-caryophyllene synthases; *PvTPS17* and *PvTPS20* – (*E*)-β-bisabolene synthases), while other orthologs adopted different functional activities. Further comparison with the genome of the closely related diploid species *Panicum hallii* revealed syntenic orthologs for more than 15 switchgrass TPS genes on six of the nine *P. hallii* chromosomes. Corresponding syntenic orthologs could also be identified for several of these genes on the genomes of the close relative *Setaria viridis* and on *Sorghum bicolor*. These findings are consistent with the observed collinearity between the switchgrass, *Setaria* and sorghum genomes (Casler et al., 2011), and suggest the presence of ancestral TPS genes in the common progenitor of sorghum and switchgrass more than 20 million years ago. Syntenic regions on the sorghum genome include a cluster of TPS genes on chromosome 7, which were found to encode insect-induced sesquiterpene synthases and share (*E*)-β-farnesene synthase activity (*Sorbic.007G055600*, *PvTPS109*) (Zhuang et al., 2012).

Most mono-TPS and sesqui-TPS genes of switchgrass exhibit tissue-specific expression patterns (Fig. 7). With the exception of the root-accumulated monoterpene borneol, the products associated with these TPSs could not be found in leaves and roots under constitutive conditions and became in part only detectable in response to stress treatment when gene expression was

induced. It is possible that under non-treatment conditions enzyme activity or substrate levels are too low to result in detectable amounts of product. In roots, microbial activity may also metabolize terpene compounds as has been shown in vetiver grass (Del Giudice et al., 2008). It may also be possible that the enzymatic products are further metabolized to non-volatile derivatives. While we have not extensively searched for such compounds, attempts to release volatile terpenes from possible glycosylated forms by acid treatment have not been successful.

Twelve TPS genes were found to be induced in switchgrass leaves upon feeding by FAW larvae. At least half of these genes are likely to contribute to the production of the volatile terpenes released upon FAW feeding based on the activity of their corresponding enzymes. The majority of the FAW-induced genes also responded to belowground treatment with MeJA and two genes were induced by root treatment with SA indicating bottom-up systemic responses in de novo terpene biosynthesis (Fig. 8B). While these effects are likely to be less pronounced with the application of lower concentrations of MeJA and SA or in response to actual root herbivory or pathogen infection, several studies have reported similar root induced systemic responses in the metabolism of terpenoids and other secondary metabolites in photosynthetic tissues (Bezemer et al., 2003; Bezemer et al., 2004; Rasmann and Turlings, 2007; Erb et al., 2008; Kaplan et al., 2008). By contrast, much weaker systemic effects have been observed on root defensive metabolites including terpenes in maize upon shoot treatments or foliar feeding (Bezemer et al., 2003; Bezemer et al., 2004; Rasmann and Turlings, 2007; Erb et al., 2008; Kaplan et al., 2008). Our findings support this notion since FAW feeding did not cause a major increase of TPS gene expression in switchgrass roots and only a local treatment with MeJA could elicit such a response (Fig. 8B).

The terpene olefins released by switchgrass leaves and roots upon insect or hormone treatment are frequently found in stress-induced volatile blends of other monocots and dicots

(Unsicker et al., 2009; Massalha et al., 2017). While determining the function of these compounds is beyond the scope of this study, we assume they play roles in direct and indirect defenses similar to those described previously in maize, rice, or other plants (Degenhardt et al., 2009; Hare and Sun, 2011; Taniguchi et al., 2014; Chen et al., 2018). A common constituent of herbivore-induced volatile blends in many plants including grasses is (*E*)-β-caryophyllene (Fricke et al., 1995; Sabulal et al., 2006; Köllner et al., 2008; Huang et al., 2012). This sesquiterpene, when released from damaged leaves of maize and rice plants, has been implicated in recruiting parasitoids of herbivores (Cheng et al., 2007; Köllner et al., 2008; Yuan et al., 2008). We identified three (*E*)-β-caryophyllene synthase genes (*PvTPS11, PvTPS14* and *PvTPS19*) (Fig. 5), all of which are located on chromosome 2 and induced upon FAW feeding and treatment with MeJA. By contrast, in maize, rice, and sorghum, only single genes (*ZmTPS23*, Os08g04500, *SbTPS4*) have been associated with the synthesis of (*E*)-β-caryophyllene upon herbivore feeding (Köllner et al., 2008; Zhuang et al., 2012; Chen et al., 2014). In MeJA-treated root tissue, *PvTPS14* was found to be induced approximately four fold higher than *PvTPS11* and *PvTPS19* and likely contributes to the emission of (*E*)-β-caryophyllene belowground. Induced root expression of (*E*)-β-caryophyllene synthases is common among grasses and has been implicated with recruitment of entomopathogenic nematodes for indirect defense against belowground herbivory (Köllner et al., 2008).

(*E*)-β-farnesene is another sesquiterpene that is released by many plant species and plays, among other volatiles, a role in indirect defense in maize (Schnee et al., 2006; Degenhardt, 2009). We found four TPS genes that encode functionally active (*E*)-β-farnesene synthases (Fig. 5). However, only *PvTPS02* expression correlated with compound emission as a result of herbivore damage (Fig. 1 and Supplementary Fig. S5). Another gene, *PvTPS16*, was highly expressed in leaves following SA treatment and strongly correlated with (*E*)-β-farnesene emission under this

condition (Figs. 1 and 8A). Despite limited and controversial evidence (Gibson and Pickett, 1983; Kunert et al., 2010), this response could potentially affect aphids, since (*E*)-β-farnesene serves as an alarm pheromone for many aphid taxa (Bowers et al., 1977; Pickett, 1983) and aphids are known to elicit both SA- and JA-dependent signaling pathways (Moran et al., 2002). A recent study by Donze-Reiner et al. (2017) found several TPS genes to be induced upon feeding by the grain aphid *Schizaphis graminum.* However, none of the (*E*)-β-farnesene synthase genes was among those induced by *S. graminum* indicating that their expression might be suppressed. Instead, genes induced by aphid feeding included the limonene synthase *PvTPS42* and the (*E*)-β-bisabolene synthases *PvTPS17* and *PvTPS20* among other genes that belong to the type c and e/f families and have in part be characterized as diterpene synthases (Pelot et al., 2018). Whether these terpene compounds are produced upon *S. graminum* feeding is currently unknown.

We found only two monoterpenes (limonene, β-ocimene) to be emitted at low levels from treated switchgrass leaves (Fig. 1). Except of *PvTPS36*, which was induced in leaves by MeJA treatment and makes limonene as an enzymatic product (Figs. 6 and 8), no terpene products of the other induced mono-TPS genes could be detected possibly because of the reasons addressed earlier. Interestingly, enzymatic products of two TPSs, the cycloisosativene synthase *Pv*TPS01 and the borneol synthase, *Pv*TPS04, could only be observed in emissions from roots although the corresponding genes were most highly expressed in leaves upon FAW, MeJA, or SA treatment (Fig. 1, Supplementary Figs. S1 and S5). Whether the absence of the compounds in leaf tissue is due to limited enzymatic activity, metabolization of the product or transport from shoots to roots remains to be determined.

In summary, our study has provided a genetic road map for investigating the biosynthesis and function of volatile terpenoids in switchgrass. We have shown that the switchgrass genome

contains an extended family of mono-TPS and sesqui-TPS genes, several of which share syntenic orthologs in other grasses and exhibit tissue specific expression and respond to herbivory and phytohormone treatment above- and belowground. The volatiles associated with these genes may exhibit functions in above- and belowground direct and indirect defense similar to those described for other grasses. Further studies involving the generation of switchgrass mutants will evaluate these ecological roles in greater detail.

MATERIALS AND METHODS

Plant materials

Seeds from the lowland allotetraploid switchgrass cv. Alamo were purchased from Bamert Seed Company (Muleshoe, TX) and used throughout this study. The seeds were sowed into potting substrate in 200 ml aluminum cans or 2.5" pots, and grown for five weeks at 26°C (16 h day) and 24°C (8 h night) in a Percival growth chamber. After germination, 15 seedlings were selected in each can or pot and grown for five weeks.

Plant treatments

Five week old seedlings were treated with larvae of *S. frugiperda* as described by Zhuang et al. (2012) with some modifications. Cans with 15 seedlings were each placed into a collection chamber and ten 2nd instar larvae were released inside the chamber for overnight feeding. For treatment with MeJA and SA, 25 ml of 5 mM MeJA or SA dissolved in ethanol were added per can or pot as a soil drench and left for 24 h respectively. For physical wounding, a surgical scalpel

was used to wound leaves and stems. Untreated plants and mock treated plants (ethanol only) were used as controls. Three replicates were performed for each treatment.

Volatile collection and identification

Volatiles emitted from leaves of the treated switchgrass and control plants placed in glass chambers were collected with an open headspace sampling system (Analytical Research Systems, Gainesville, FL, USA) in the light from 9:00 am to 1:00 pm. FAW larvae were removed prior to volatile trapping. The volatiles were collected with volatile collection traps (Poropak-Q, http://www.ars.com) and eluted with 100 µl methylene chloride containing 0.003% nonyl acetate (v/v). The collected volatiles were analyzed on a Shimadzu 17A gas chromatograph coupled to a Shimadzu QP5050A (http://www.shimadzu.com). Root volatiles were analyzed by SPME (solid-phase microextraction, AOC-5000 Shimadzu) through incubation of detached pooled root tissue (1 g) in a 20 ml glass vial containing 2 ml of water and 20 ng of the standard 1-bromodecane. The SPME fiber consisting of 100-µM polydimethylsiloxane (Supelco) was placed into the headspace of the vial for 30 min following treatment. Collected volatiles where thermally desorbed for 4 min and analyzed using a gas chromatograph (240^{o}C injector port) coupled with a quadrupole mass spectrometer (GC-MS-QP2010S, Shimadzu). Extracts were separated with a 2:1 split on a 30 m × 0.25 mm i.d. × 0.25 µm film thickness Zebron capillary column (Phenomenex) using Helium as the carrier gas (1.4 ml min^{-1} flow rate) and a temperature gradient of 5^{o}C min^{-1} from 40^{o}C (hold 2 min) to 220^{o}C. Compound identification was based on similarity to library matches (NIST, WILEY), authentic standards (Sigma-Aldrich, (*E*)-β-caryophyllene, germacrene-D) and comparison to Opopanax essential oil (Floracopeia, α-humulene, δ-cadinene).

Identification of TPS genes from the switchgrass genome and phylogeny reconstruction

Putative switchgrass TPS genes were retrieved from Phytozome (www.phytozome.jgi.doe.gov) through an annotation based keyword search of genome versions v.1 and v.4. In addition, RNA-seq data kindly provided by the Noble Foundation (https://www.noble.org) for above- and belowground tissues were assembled de novo using Trinity (Grabherr et al., 2011). Assembled transcriptomes were queried with a representative switchgrass TPS sequence (*PvTPS01*) using NCBI's TBLASTX. Resulting BLAST hits were manually curated for putative functionality based on length and presence of the conserved aspartate rich motif (DDxxD) necessary for ionization of the prenyldiphosphate substrate (Class I TPSs). Class I and II diterpene synthases identified in this study were not further pursued based on previous reporting by Pelot et al. (2018). Gene models were refined further by comparing transcripts to genome sequences available in Phytozome. Putative N-terminal plastidic transit peptides were predicted using multiple sequence alignments and analysis of each sequence with the transit peptide prediction software ChloroP (Emanuelsson et al., 1999). Phylogeny reconstruction was based on protein sequence alignments, which were performed using MAFFT (Katoh et al., 2002). Maximum likelihood (ML) trees were then built from the MAFFT alignment using PhyML (Guindon et al., 2010) with 500 bootstrap replicates as previously described (Pelot et al., 2018). Final phylogeny annotation and design were performed in iTOL (Letunic and Bork, 2007). Heat map analysis was based on publically available expression data at http://www.phytozome.net/ following previously described methods (Pelot et al., 2018).

Synteny analysis and identification of orthologous TPS genes

P. virgatum (v4.1)*, Setaria italica* (v2.2), *Sorghum bicolor* (v3.1), *and P. hallii* var. *hallii* (v2.1) genome annotations were downloaded from phytozome (phytozome.jgi.doe.gov). Syntenic blocks were generated following Lovell et al. (2018) via the GENESPACE pipeline. Orthofinder was run on synteny-constrained BLASTp results to build orthologous gene networks.

Gene expression analysis

Total RNA was isolated from switchgrass leaves and roots using the RNeasy Plant Mini Kit according to the manufacturer's protocol (http://www.qiagen.com). cDNA was synthesized using the General Electric first strand synthesis kit according to the manufacturer's protocol (http://www.gelifesciences.com). Gene expression analysis was carried out using quantitative RT-PCR, which was described previously (Chen et al., 2018). Sequences of primers used for RT-qPCR are listed in Supplementary Table 2.

Protein expression in E. coli and terpene synthase activity assay

Full-length and truncated genes (predicted transit peptide removed) were synthesized and cloned (*Nde*I) into the pET-28b(+) prokaryotic expression vector. Constructs were transformed into *E. coli* BL21-CodonPLus(DE3) cells (Stratagene) and grown at 37^{o} C in 100 ml Luria-Bertani (LB) media supplemented with 50 µM kanamycin until an optical density at 600 nm (OD_{600}) of 0.5 to 0.7. Protein production was then induced with 0.5 mM isopropyl 1-thio-ß-D-galactopyranoside (IPTG) and incubated with shaking at 18^{o} C for 16 h. Recombinant protein extraction and partial purification were performed as described by Tholl et al. (2005), with the modification that N-terminal His-tags were implemented for partial purification. Enzyme reactions (125 µl total

volume) were prepared in a 10-ml screw cap vial (Supelco) by combing partially purified protein with 20 mM $MgCl_2$ and 60 μM commercially available prenyl diphosphate substrates GDP and (*E,E*)-FDP (Echelon Biosciences). Assay mixtures were incubated for 5 min at 30° C in the presence of a 100-μM polydimethylsiloxane fiber (Supelco). Collected volatiles where thermally desorbed for 4 min and analyzed using a gas chromatograph (240°C injector port) coupled with a quadrupole mass spectrometer (GC-MS-QP2010S, Shimadzu). Extracts were separated with a 5:1 split under the same conditions described above. Compound identification, in addition to those compounds described above, was based on similarity to library matches (NIST, WILEY, cycloisosativene, β-elemene, α-selinene, valencene, α-patchoulene, copaene), authentic standards (Sigma-Aldrich, α-terpinolene, borneol, 1,8-cineol, linalool, geraniol) and comparison to Opopanax oil (Floracopeia, α-pinene, limonene, sabinene, a-terpineol, (*E*)-β-farnesene, (*E*)-γ-bisabolene, α-santalene, γ-curcumene, β-bisabolene).

Acknowledgements

We would like to thank Dr. Nan Zhao for his assistance with volatile profiling and Dr. Qidong Jia for his assistant with sequence analysis. Debbie Wiley for assistance with plant maintenance.

REFERENCES

Bezemer, T.M., Wagenaar, R., Van Dam, N.M., Van Der Putten, W.H., and Wackers, F.L. (2004). Above- and below-ground terpenoid aldehyde induction in cotton, *Gossypium herbaceum*, following root and leaf injury. *J. Chem. Ecol.* 30**,** 53-67. doi: 10.1023/B:JOEC.0000013182.50662.2a.

Bezemer, T.M., Wagenaar, R., Van Dam, N.M., and Wackers, F.L. (2003). Interactions between above- and belowground insect herbivores as mediated by the plant defense system. *Oikos* 101**,** 555-562. doi: 10.1034/j.1600-0706.2003.12424.x.

Block, A.K., Vaughan, M.M., Schmelz, E.A., and Christensen, S.A. (2019). Biosynthesis and function of terpenoid defense compounds in maize (*Zea mays*). *Planta* 249**,** 21-30. doi: 10.1007/s00425-018-2999-2.

Boutanaev, A.M., Moses, T., Zi, J.C., Nelson, D.R., Mugford, S.T., Peters, R.J., and Osbourn, A. (2015). Investigation of terpene diversification across multiple sequenced plant genomes. *Proc. Natl. Acad. Sci. U.S.A.* 112**,** E81-E88. doi: 10.1073/pnas.1419547112.

Bowers, W.S., Nishino, C., Montgomery, M.E., and Nault, L.R. (1977). Structure-activity relationships of analogs of the aphid alarm pheromone, (*E*)-β-farnesene. *J. Insect Physiol.* 23**,** 697-701.

Casler, M.D., Tobias, C.M., Kaeppler, S.M., Buell, C.R., Wang, Z.Y., Cao, P.J., Schmutz, J., and Ronald, P. (2011). The switchgrass genome: tools and strategies. *Plant Genome* 4**,** 273-282. doi: 10.3835/plantgenome2011.10.0026.

Casler, M.D., Vogel, K.P., Taliaferro, C.M., Ehlke, N.J., Berdahl, J.D., Brummer, E.C., Kallenbach, R.L., West, C.P., and Mitchell, R.B. (2007). Latitudinal and longitudinal adaptation of switchgrass populations. *Crop Sci.* 47**,** 2249-2260. doi: 10.2135/cropsci2006.12.0780.

Chen, F., Tholl, D., Bohlmann, J., and 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. doi: 10.1111/j.1365-313X.2011.04520.x.

Chen, H., Li, G.L., Köllner, T.G., Jia, Q.D., Gershenzon, J., and Chen, F. (2014). Positive Darwinian selection is a driving force for the diversification of terpenoid biosynthesis in the genus Oryza. *BMC Plant Biol.* 14. doi: 10.1186/s12870-014-0239-x.

Chen, X.J., Chen, H., Yuan, J.S., Köllner, T.G., Chen, Y.Y., Guo, Y.F., Zhuang, X.F., Chen, X.L., Zhang, Y.J., Fu, J.Y., Nebenfuhr, A., Guo, Z.J., and Chen, F. (2018). The rice terpene synthase gene OsTPS19 functions as an (*S*)-limonene synthase in planta, and its overexpression leads to enhanced resistance to the blast fungus *Magnaporthe oryzae*. *Plant Biotechnol. J.* 16**,** 1778-1787. doi: 10.1111/pbi.12914.

Cheng, A.X., Xiang, C.Y., Li, J.X., Yang, C.Q., Hu, W.L., Wang, L.J., Lou, Y.G., and Chen, X.Y. (2007). The rice (*E*)-beta-caryophyllene synthase (OsTPS3) accounts for the major inducible volatile sesquiterpenes. *Phytochemistry* 68**,** 1632-1641. doi: 10.1016/j.phytochem.2007.04.008.

Degenhardt, J. (2009). Indirect defense responses to herbivory in grasses. *Plant Physiol.* 149**,** 96-102. doi: 10.1104/pp.108.128975.

Degenhardt, J., Hiltpold, I., Köllner, T.G., Frey, M., Gierl, A., Gershenzon, J., Hibbard, B.E., Ellersieck, M.R., and Turlings, T.C.J. (2009). Restoring a maize root signal that attracts insect-killing nematodes to control a major pest. *Proc. Natl. Acad. Sci. U.S.A.* 106**,** 13213-13218. doi: 10.1073/pnas.0906365106.

Del Giudice, L., Massardo, D.R., Pontieri, P., Bertea, C.M., Mombello, D., Carata, E., Tredici, S.M., Tala, A., Mucciarelli, M., Groudeva, V.I., De Stefano, M., Vigliotta, G., Maffei, M.E., and Alifano, P. (2008). The microbial community of Vetiver root and its involvement into essential oil biogenesis. *Environ. Microbiol.* 10**,** 2824-2841. doi: 10.1111/j.1462-2920.2008.01703.x.

Dicke, M. (1994). Local and systemic production of volatile herbivore-induced terpenoids - their role in plant-carnivore mutualism. *J. Plant Physiol.* 143**,** 465-472. doi: 10.1016/S0176-1617(11)81808-0.

Donze-Reiner, T., Palmer, N.A., Scully, E.D., Prochaska, T.J., Koch, K.G., Heng-Moss, T., Bradshaw, J.D., Twigg, P., Amundsen, K., Sattler, S.E., and Sarath, G. (2017). Transcriptional analysis of defense mechanisms in upland tetraploid switchgrass to greenbugs. *BMC Plant Biol.* 17. doi: 10.1186/s12870-017-0998-2.

Dudareva, N., Pichersky, E., and Gershenzon, J. (2004). Biochemistry of plant volatiles. *Plant Physiol.* 135**,** 1893-1902. doi: 10.1104/pp.104.049981.

Emanuelsson, O., Nielsen, H., and Von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8**,** 978-984. doi: 10.1110/ps.8.5.978.

Erb, M., Ton, J., Degenhardt, J., and Turlings, T.C.J. (2008). Interactions between arthropod-induced aboveground and belowground defenses in plants. *Plant Physiol.* 146**,** 867-874. doi: 10.1104/pp.107.112169.

Erb, M., Veyrat, N., Robert, C.A.M., Xu, H., Frey, M., Ton, J., and Turlings, T.C.J. (2015). Indole is an essential herbivore-induced volatile priming signal in maize. *Nature Communications* 6. doi: 10.1038/ncomms7273.

Fricke, C., Rieck, A., Hardt, I.H., Konig, W.A., and Muhle, H. (1995). Identification of (+)-beta-caryophyllene in essential oils of liverworts by enantioselective gas-chromatography. *Phytochemistry* 39**,** 1119-1121. doi: 10.1016/0031-9422(95)00184-9.

Gibson, R.W., and Pickett, J.A. (1983). Wild potato repels aphids by release of aphid alarm pheromone. *Nature* 302**,** 608-609. doi: 10.1038/302608a0.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q.D., Chen, Z.H., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29**,** 644-U130. doi: 10.1038/nbt.1883.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59**,** 307-321. doi: 10.1093/sysbio/syq010.

Hare, J.D., and Sun, J.J. (2011). Production of induced volatiles by *Datura wrightii* in response to damage by insects: effect of herbivore species and time. *J. Chem. Ecol.* 37**,** 751-764. doi: 10.1007/s10886-011-9985-5.

Hofberger, J.A., Ramirez, A.M., van den Bergh, E., Zhu, X.G., Bouwmeester, H.J., Schuurink, R.C., and Schranz, M.E. (2015). Large-scale evolutionary analysis of genes and supergene clusters from terpenoid modular pathways provides insights into metabolic diversification in flowering plants. *PLoS One* 10. doi: 10.1371/journal.pone.0128808.

Huang, M., Sanchez-Moreiras, A.M., Abel, C., Sohrabi, R., Lee, S., Gershenzon, J., and Tholl, D. (2012). The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (*E*)-beta-caryophyllene, is a defense against a bacterial pathogen. *New Phytol.* 193**,** 997-1008. doi: 10.1111/j.1469-8137.2011.04001.x.

Kaplan, I., Halitschke, R., Kessler, A., Sardanelli, S., and Denno, R.F. (2008). Constitutive and induced defenses to herbivory in above- and belowground plant tissues. *Ecology* 89**,** 392-406. doi: 10.1890/07-0471.1.

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30**,** 3059-3066. doi: 10.1093/nar/gkf436.

Köllner, T.G., Held, M., Lenk, C., Hiltpold, I., Turlings, T.C.J., Gershenzon, J., and Degenhardt, J. (2008). A maize (*E*)-beta-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most American maize varieties. *Plant Cell* 20**,** 482-494. doi: 10.1105/tpc.107.051672.

Kost, C., and Heil, M. (2006). Herbivore-induced plant volatiles induce an indirect defence in neighbouring plants. *J. Ecol.* 94**,** 619-628. doi: 10.1111/j.1365-2745.2006.01120.x.

Kulheim, C., Padovan, A., Hefer, C., Krause, S.T., Köllner, T.G., Myburg, A.A., Degenhardt, J., and Foley, W.J. (2015). The Eucalyptus terpene synthase gene family. *BMC Genomics* 16. doi: 10.1186/s12864-015-1598-x.

Kunert, G., Reinhold, C., and Gershenzon, J. (2010). Constitutive emission of the aphid alarm pheromone, (*E*)-beta-farnesene, from plants does not serve as a direct defense against aphids. *BMC Ecol.* 10**,** 23. doi: 10.1186/1472-6785-10-23.

Lee, S.T., Mitchell, R.B., Wang, Z., Heiss, C., Gardner, D.R., and Azadi, P. (2009). Isolation, characterization, and quantification of steroidal saponins in switchgrass (*Panicum virgatum* L.). *J. Agric. Food Chem.* 57**,** 2599-2604. doi: 10.1021/jf803907y.

Letunic, I., and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23**,** 127-128. doi: 10.1093/bioinformatics/btl529.

Lovell, J.T., Jenkins, J., Lowry, D.B., Mamidi, S., Sreedasyam, A., Weng, X.Y., Barry, K., Bonnette, J., Campitelli, B., Daum, C., Gordon, S.P., Gould, B.A., Khasanova, A., Lipzen, A., MacQueen, A., Palacio-Mejia, J.D., Plott, C., Shakirov, E.V., Shu, S.Q., Yoshinaga, Y., Zane, M., Kudrna, D., Talag, J.D., Rokhsar, D., Grimwood, J., Schmutz, J., and Juenger, T.E. (2018). The genomic landscape of molecular responses to natural drought stress in *Panicum hallii*. *Nature Communications* 9. doi: 10.1038/s41467-018-07669-x.

Massalha, H., Korenblum, E., Tholl, D., and Aharoni, A. (2017). Small molecules below-ground: the role of specialized metabolites in the rhizosphere. *Plant J.* 90**,** 788-807. doi: 10.1111/tpj.13543.

Moran, P.J., Cheng, Y., Cassell, J.L., and Thompson, G.A. (2002). Gene expression profiling of Arabidopsis thaliana in compatible plant-aphid interactions. *Arch. Insect Biochem. Physiol.* 51**,** 182-203. doi: 10.1002/arch.10064.

Osbourn, A.E. (1996). Preformed antimicrobial compounds and plant defense against fungal attack. *Plant Cell* 8**,** 1821-1831. doi: 10.1105/tpc.8.10.1821.

Parrish, D.J., and Fike, J.H. (2005). The biology and agronomy of switchgrass for biofuels. *Crit. Rev. Plant Sci.* 24**,** 423-459. doi: 10.1080/07352680500316433.

Pelot, K.A., Chen, R., Hagelthorn, D.M., Young, C.A., Addison, J.B., Muchlinski, A., Tholl, D., and Zerbe, P. (2018). Functional diversity of diterpene synthases in the biofuel crop switchgrass. *Plant Physiol.* 178**,** 54-71. doi: 10.1104/pp.18.00590.

Pichersky, E., and Gang, D.R. (2000). Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci.* 5**,** 439-445.

Pickett, R.W.G.J.A. (1983). Wild potato repels aphids by release of aphid alarm pheromone. *Nature* 608-609.

Rasmann, S., Köllner, T.G., Degenhardt, J., Hiltpold, I., Toepfer, S., Kuhlmann, U., Gershenzon, J., and Turlings, T.C.J. (2005). Recruitment of entomopathogenic nematodes by insect-damaged maize roots. *Nature* 434**,** 732-737. doi: 10.1038/nature03451.

Rasmann, S., and Turlings, T.C.J. (2007). Simultaneous feeding by aboveground and belowground herbivores attenuates plant-mediated attraction of their respective natural enemies. *Ecol. Lett.* 10**,** 926-936. doi: 10.1111/j.1461-0248.2007.01084.x.

Sabulal, B., Dan, M., Anil, J.J., Kurup, R., Pradeep, N.S., Valsamma, R.K., and George, V. (2006). Caryophyllene-rich rhizome oil of *Zingiber nimmonii* from South India: chemical

characterization and antimicrobial activity. *Phytochemistry* 67**,** 2469-2473. doi: 10.1016/j.phytochem.2006.08.003.

Sanderson, M.A., Adler, P.R., Boateng, A.A., Casler, M.D., and Sarath, G. (2006). Switchgrass as a biofuels feedstock in the USA. *Can. J. Plant Sci.* 86**,** 1315-1325. doi: 10.4141/P06-136.

Schmelz, E.A., Huffaker, A., Sims, J.W., Christensen, S.A., Lu, X., Okada, K., and Peters, R.J. (2014). Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant J.* 79**,** 659-678. doi: 10.1111/tpj.12436.

Schnee, C., Köllner, T.G., Held, M., Turlings, T.C.J., Gershenzon, J., and Degenhardt, J. (2006). The products of a single maize sesquiterpene synthase form a volatile defense signal that attracts natural enemies of maize herbivores. *Proc. Natl. Acad. Sci. U.S.A.* 103**,** 1129-1134. doi: 10.1073/pnas.0508027103.

Springer, N.M., Anderson, S.N., Andorf, C.M., Ahern, K.R., Bai, F., Barad, O., Barbazuk, W.B., Bass, H.W., Baruch, K., Ben-Zvi, G., Buckler, E.S., Bukowski, R., Campbell, M.S., Cannon, E.K.S., Chomet, P., Dawe, R.K., Davenport, R., Dooner, H.K., Du, L.H., Du, C.G., Easterling, K.A., Gault, C., Guan, J.C., Hunter, C.T., Jander, G., Jiao, Y.P., Koch, K.E., Kol, G., Köllner, T.G., Kudo, T., Li, Q., Lu, F., Mayfield-Jones, D., Mei, W.B., McCarty, D.R., Noshay, J.M., Portwood, J.L., Ronen, G., Settles, A.M., Shem-Tov, D., Shi, J.H., Soifer, I., Stein, J.C., Stitzer, M.C., Suzuki, M., Vera, D.L., Vollbrecht, E., Vrebalov, J.T., Ware, D., Wei, S.R., Wimalanathan, K., Woodhouse, M.R., Xiong, W.W., and Brutnell, T.P. (2018). The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat. Genet.* 50. doi: 10.1038/s41588-018-0158-0.

Taniguchi, S., Hosokawa-Shinonaga, Y., Tamaoki, D., Yamada, S., Akimitsu, K., and Gomi, K. (2014). Jasmonate induction of the monoterpene linalool confers resistance to rice bacterial blight and its biosynthesis is regulated by JAZ protein in rice. *Plant, Cell Environ.* 37**,** 451-461. doi: 10.1111/pce.12169.

Tholl, D., Chen, F., Petri, J., Gershenzon, J., and Pichersky, E. (2005). Two sesquiterpene synthases are responsible for the complex mixture of sesquiterpenes emitted from Arabidopsis flowers. *Plant J.* 42**,** 757-771. doi: 10.1111/j.1365-313X.2005.02417.x.

Tholl, D., and Lee, S. (2011). Terpene specialized metabolism in *Arabidopsis thaliana*. *Arabidopsis Book* 9. doi: 10.1199/tab.0143.

Tholl, D., Sohrabi, R., Huh, J.H., and Lee, S. (2011). The biochemistry of homoterpenes - common constituents of floral and herbivore-induced plant volatile bouquets. *Phytochemistry* 72**,** 1635-1646. doi: 10.1016/j.phytochem.2011.01.019.

Turlings, T.C.J., Tumlinson, J.H., and Lewis, W.J. (1990). Exploitation of herbivore-induced plant odors by host-seeking parasitic wasps. *Science* 250**,** 1251-1253. doi: 10.1126/science.250.4985.1251.

Unsicker, S.B., Kunert, G., and Gershenzon, J. (2009). Protective perfumes: the role of vegetative volatiles in plant defense against herbivores. *Curr. Opin. Plant Biol.* 12**,** 479-485. doi: 10.1016/j.pbi.2009.04.001.

Vaughan, M.M., Wang, Q., Webster, F.X., Kiemle, D., Hong, Y.J., Tantillo, D.J., Coates, R.M., Wray, A.T., Askew, W., O'Donnell, C., Tokuhisa, J.G., and Tholl, D. (2013). Formation of the unusual semivolatile diterpene rhizathalene by the Arabidopsis class I terpene synthase TPS08 in the root stele is involved in defense against belowground herbivory. *Plant Cell* 25**,** 1108-1125. doi: 10.1105/tpc.112.100057.

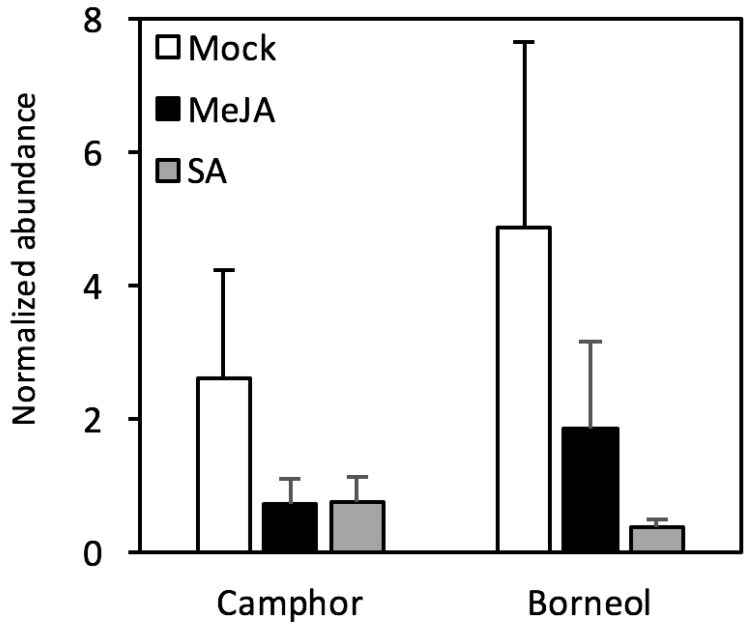
Vogel, K.P. (2004). *Warm-season (C4) grasses.* American Society of Agronomy, Crop Science Society of America, Soil Science Society of America.

Wendel, J.F. (2000). Genome evolution in polyploids. *Plant Mol. Biol.* 42**,** 225-249. doi: 10.1023/A:1006392424384.

Yuan, J.S., Köllner, T.G., Wiggins, G., Grant, J., Degenhardt, J., and Chen, F. (2008). Molecular and genomic basis of volatile-mediated indirect defense against insects in rice. *Plant J.* 55**,** 491-503. doi: 10.1111/j.1365-313X.2008.03524.x.

Zhuang, X.F., Köllner, T.G., Zhao, N., Li, G.L., Jiang, Y.F., Zhu, L.C., Ma, J.X., Degenhardt, J., and Chen, F. (2012). Dynamic evolution of herbivore-induced sesquiterpene biosynthesis in sorghum and related grass crops. *Plant J.* 69**,** 70-80. doi: 10.1111/j.1365-313X.2011.04771.x.

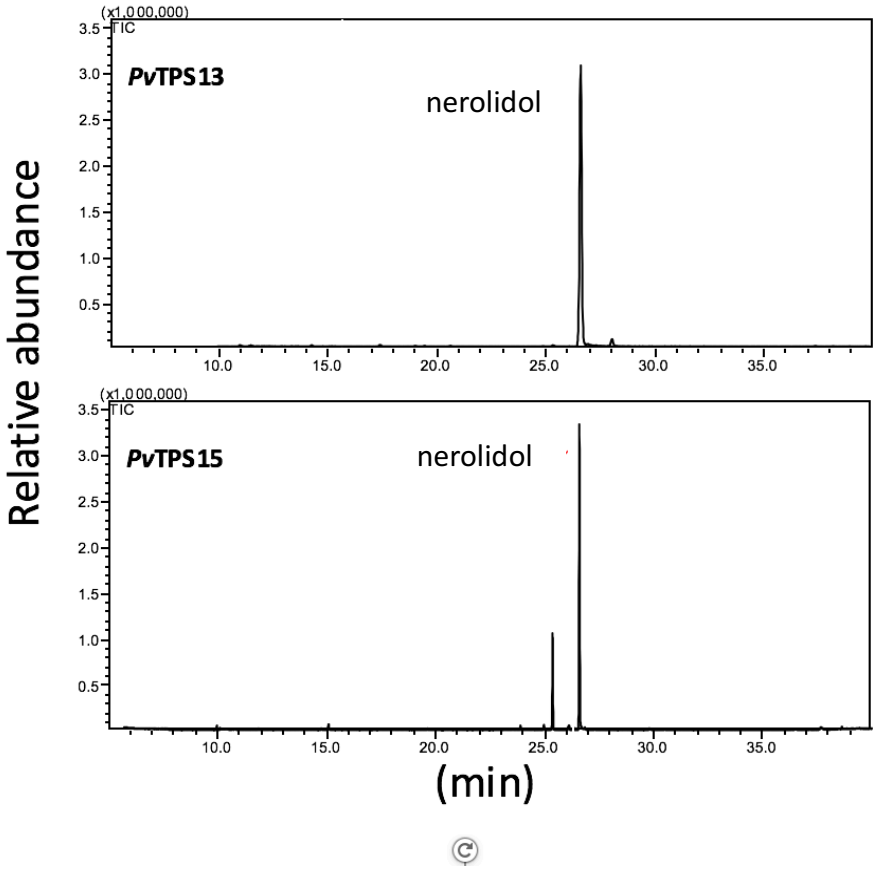
SUPPLEMENTAL FIGURES AND TABLES



Supplemental Figure 1. SPME-GC-MS analysis of camphor and borneol emission following MeJA and SA treatment. Volatile compounds were analyzed in triplicate from detached pooled root material of 6-week old plants. Samples were normalized to an internal standard (1-bromodecane) and grams fresh. Compound identification was based on similarity to library matches (NIST, WILEY) and authentic standards (Sigma).

Supplemental Table 1. RT-qPCR primers used in this study.

	FWD	REV
PvTPS1	GCCCCGCAGTCACAGTGC	CAGGTCGGAGGAGGCAAAGCTTTCC
PvTPS2	GGATCACGGTAGAGGAGGCCAATG	CTTTGACACTGTATGGGTAGGTATACGC
PvTPS3	GCACTTCATTCCAACAGTG	CGTTTATAATTTTCCTGACTTC
PvTPS4	GGTTGAACCCTGATAATTCAGTGCCTAG	CAAGTGGATGGTATCCTTGAGGTTTTGAC
PvTPS5	GCGGATGCGAGAGCGGG	CTGGCCCTGAATTGACGACGC
PvTPS6	CTACTTCCGTTGCTACGACCATTCC	GGTTCATTGTCTCCACGACATTATTCTTGG
PvTPS7	GTTTGGCTATCTTACCTGTTGG	GTCGGAGCTGCTGATG
PvTPS8	GAGGTGGCTGGAGCAGCGG	CTCTCATGCTGGCCCCTAAAAGTGGAAC
PvTPS9	CACCACCAAGGACTTTGTGGTTGTC	GCCATTCTTCGGACACATTAACCAAAGG
PvTPS10	GCGATGGCGTACCAG	CCACGGATACACAGTC
PvTPS11	GCCCTGGCCTCGCAACAAG	CCTGCTTGATCCGTGCACTCTCC
PvTPS12	GTCACAGCCGGTTGGCCG	CATGGCTGAATTGGCACACCAGATC
PvTPS13	CTGCACGCACCTTCTCCGTGC	CGGGCGCTGTTCCTACGTTC
PvTPS14	GATCTCCGGCAGCCCTG	CAGCTTGCGAACCAC
PvTPS15	CAGGGCCGTGAAGC	GACGGGACGCTCTG
PvTPS16	CGCCTCCAGTTTGCTC	GCCGCTGTAGCGTG
PvTPS17	GCCAGGAGGAGATCAAG	GGCAAAGTCGGCGG
PvTPS19	GGACCTGCAGAGGCAG	CTCCGTGCCCTCTCC
PvTPS20	GCGTCGGGGCATGAAG	GAATTTCCTGAAGTGAAGTTG
PvTPS36	CTTCAAACCACCTCCTC	GGCGTTCTGACTGAG
PvTPS53	CTATGTCACATGGAATGTGGC	GCCCTCTCTGTCTCTG
PvTPS56	GAGCGAGGCTGCG	GCATCCGCTTCTCAG
ELF1a	GGAGAAGTCCCACATCAACATCGTGG	CGCTCCGCCTTGAGCTTGTC



Supplemental Figure 2. SPME-GC-MS analysis of volatile products produced by recombinant TPSs with (*E,E*)-FPP as a substrate. Compound identification was based on similarity to library matches (NIST, WILEY), authentic standards and comparison to Opopanax oil.

CHAPTER III. BIOSYNTHESIS AND FUNCTION OF VOLATILE TERPENOIDS IN SWITCHGRASS ROOT-MICROBIOME INTERACTIONS

INTRODUCTION

Inter- and intraspecific interactions between organisms occur ubiquitously throughout the domains of life. Volatile compounds are often used by organisms to facilitate these interactions (Junker et al., 2018). Plants, for example, can produce fragrant compounds to recruit pollinators as well as noxious compounds to deter insect feeding (Lucas-Barbosa et al., 2011). Emissions of volatiles do not only occur from plant tissues aboveground but also from roots (Penuelas et al., 2014). The accumulation and release of volatile compounds from roots has been implicated with direct or indirect chemical defenses against pathogens and root herbivores (Rasmann et al., 2005; Kollner et al., 2008; Huang et al., 2012; Vaughan et al., 2013; Sohrabi et al., 2015). However, much less attention has been paid to possible functions of volatile (and non-volatile) compounds in interaction of plants with root-associated microbes.

Like many organisms, plants host a structured microbial community which supports plant growth and resilience to stress (Yang et al., 2009; Bulgarelli et al., 2013). The mechanisms for how plants assemble and maintain a "core" microbiome have received increased attention due to the potential applications for increasing plant yields and hardiness (Zhang et al., 2015). Throughout major plant organs (e.g. flowers, leaves, roots), microbiome structure can be heavily influenced by environmental factors and local release of nutrients (Lindow and Brandl, 2003; Berg and Smalla, 2009; Junker et al., 2011). In addition, plant pollinators can mediate transmission of microbes to new niche spaces influencing microbial community dispersals and composition (Beisner et al., 2006; Ushio et al., 2015). In roots, microbial communities show significant differences between root interior (endosphere) and the root–soil interface (rhizosphere) across

monocots (rice, maize) and dicots (*Arabidopsis*; Bulgarelli et al., 2012; Edwards et al., 2015; Hu et al., 2018). Though soil type has a major influence on root-microbe interactions (Lundberg et al., 2012), plant genotype and sloughing of plant material from the root tip also drive microbial community assembly and maintenance (Bulgarelli et al., 2012; Lundberg et al., 2012; Badri et al., 2013).

Although plants produce a large number of secondary or specialized metabolites that mediate interactions with other organisms (Rasmann et al., 2005; Kollner et al., 2008; Huang et al., 2012; Vaughan et al., 2013; Sohrabi et al., 2015), there is limited knowledge on whether these chemical compounds serve as host specific factors that affect root microbial community composition. In *Arabidopsis*, the multifaceted defense hormone salicylic acid has been shown to modulate root colonization by serving as both an immune signal and carbon source for specific bacterial taxa (Lebeis et al., 2015). In addition, volatile specialized molecules are known to accumulate in roots or be released from root tissues to serve in belowground chemical interactions (Penuelas et al., 2014). Volatile metabolites such as terpenoids can modify microbial colonization or pathogen growth in the phyllosphere and anthosphere (Huang et al., 2012; Junker and Tholl, 2013; Junker et al., 2018); but, very little information is available on how these molecules function in the rhizosphere or endosphere. In fragrant grasses such as vetiver grass (*Chrysopogon zizanioides*), a grass native to India, high levels of 15-carbon sesquiterpenes are stored in roots (Belhassen et al., 2015). Del Giudice et al. (2018) could show that colonization of sterile roots of *C. zizanioides* with rhizobacteria stimulated the accumulation of terpenes and that specific root endophytes transformed these sesquiterpene volatiles into a broader range of terpene compounds. Following these findings, we hypothesized that volatile compounds such as terpenes might play significant roles in establishing root-microbial associations in other grasses or plants in general.

Here, we have investigated terpene biosynthesis and dynamics in switchgrass roots to address this question. As presented in Chapter 1, we have shown that switchgrass can produce a large number of different volatile terpenes in response to biotic stress (Muchlinski et al., in review). Moreover, we have found that switchgrass roots accumulate the 10-carbon monoterpene borneol and its oxidation product camphor. Both compounds are found commonly in essential oil producing plants such as sage, where the biosynthesis of borneol diphosphate was first described (Wise et al., 1998). Although borneol and camphor have antimicrobial activity (Tabanca et al., 2001), common soil bacteria (e.g. *Pseudomonas putida*) can overcome this effect and use the compound as a sole nutrient source (Prasad et al., 2013). We, therefore, assumed that switchgrass can use borneol to recruit beneficial microorganisms to the root endosphere while simultaneously defending against harmful soil microbes. We show that switchgrass cultivars of diverse genetic and ecological backgrounds constitutively accumulate borneol and camphor in root tissues. We have investigated the formation of borneol and the dynamics of borneol and camphor in the plant in response to treatments with rhizobacterial communities and salicylic acid. We have further initiated experiments to characterize endophytic isolates for their ability to use the monoterpenes as carbon source. Finally, RNAi-based switchgrass lines have been generated to further investigate the effect of borneol and camphor on rhizosphere and endophytic bacterial (and possibly fungal) communities in situ. Results from this study contribute to the growing knowledge of biological function of specialized metabolites in belowground interactions.

RESULTS

Borneol and camphor accumulation is conserved across upland and lowland switchgrass ecotypes

To determine major volatile compounds produced in switchgrass roots, we analyzed hexane extracts of roots from six cultivars by GC/MS. Briefly, plants were grown from seed to the reproductive 1 growth stage (R1, ~3 months old) in plastic pots under greenhouse conditions containing soil collected from a local switchgrass field site (Moore Farm, Blacksburg, VA; see MATERIALS AND METHODS). We found that the monoterpenoids borneol and the borneol oxidation product camphor are accumulated in roots of all cultivars, albeit at different ratios, regardless of ecotype (upland/lowland) or level of polyploidy (Fig. 1). Additionally, domesticated cultivars (i.e. Alamo, Pathfinder, Sunburst, Trailblazer) showed no significant differences in borneol levels from natural accessions (i.e. Dacotah and Kanlow), suggesting borneol metabolism has been retained irrespective of breeding efforts (Supplemental Fig. S1).

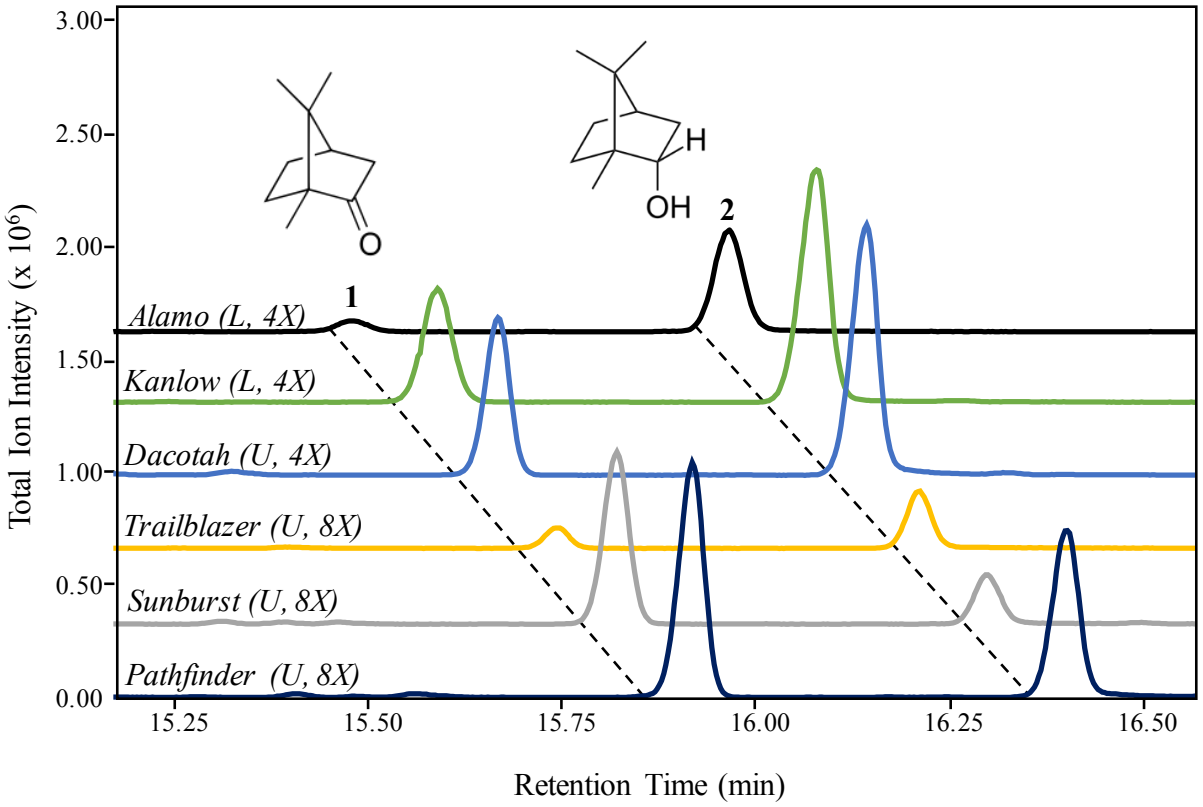
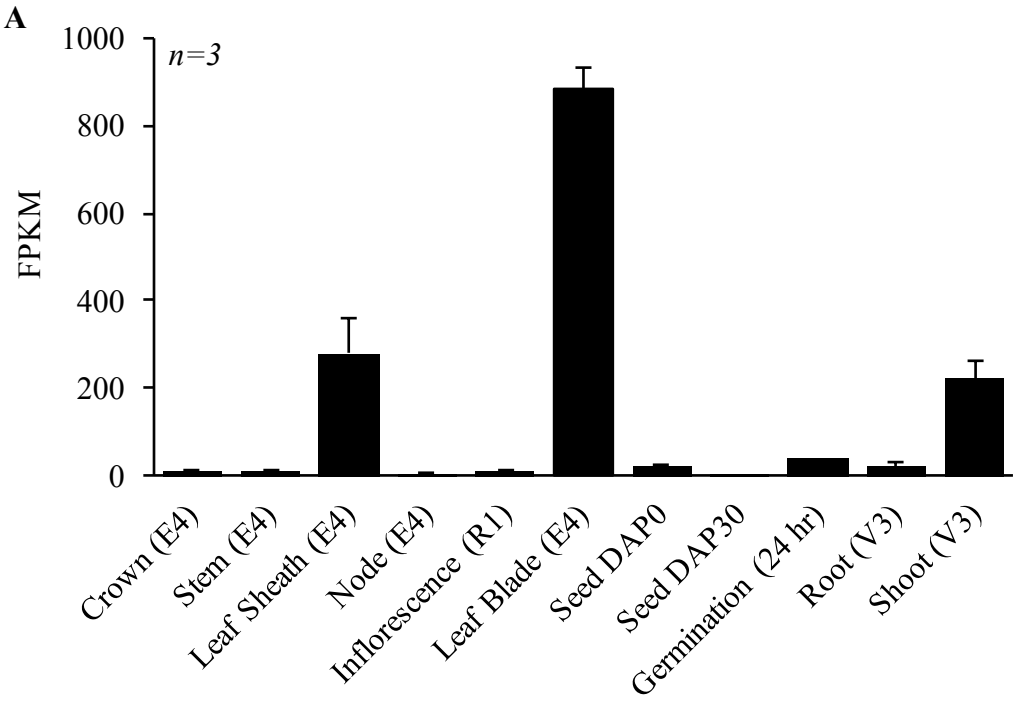


Figure 1. GC/MS analysis of camphor and borneol from roots of six switchgrass cultivars. Marked peaks represent camphor (1) and borneol (2) based on comparisons with authentic standards. Ecotype: upland (U) or lowland (L) and ploidy level: tetraploid (2n=4x=36) or octoploid (2n=8x=72), are shown.

PvTPS04 is expressed in leaf tissues

To date, a single borneol synthase gene (*PvTPS04*) has been identified and biochemically characterized from switchgrass (Muchlinski et al., in review). To better understand the formation and function of borneol in switchgrass, we first analyzed tissue specific expression patterns of *PvTPS04*. Analysis of transcript levels in genotype AP13 (Alamo-related), across publically available RNA-seq data (www.phytozome.jgi.doe.gov), revealed that *PvTPS04* is expressed in above- and belowground tissues including root, leaf, stem and inflorescence (Fig. 2A). Highest expression was detected in leaf blade tissue at the Elongation 4 growth stage (E4). Quantitative

RT-PCR revealed significant differences in expression of *PvTPS04* between leaf and other major switchgrass tissues (p=0.013; ANOVA, Fig. 2B). Specifically, transcript levels of *PvTPS04* in leaf tissue were found to be significantly higher than those in inflorescences, stem or root tissue (p<0.01, Tukey-HSD, Fig. 2B).



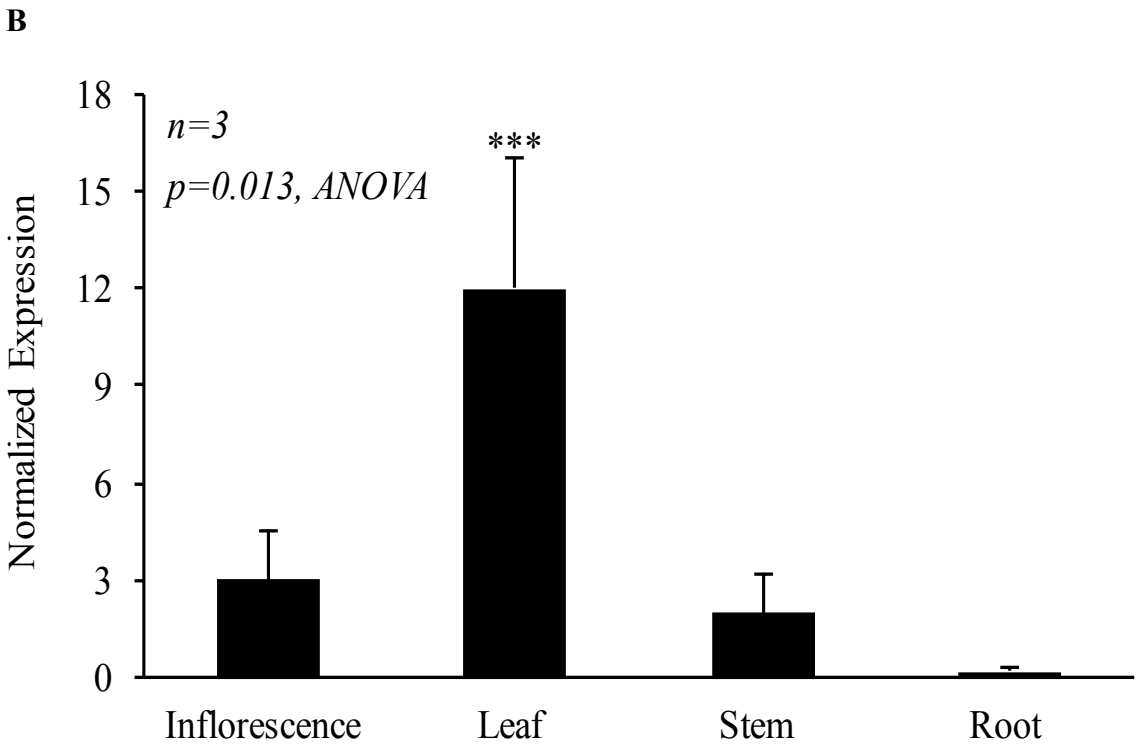


Figure 2. Expression analysis of *PvTPS04* across switchgrass (cv. AP13) tissues. A) In silico expression of *PvTPS04* in different tissues and growth stages reported as FPKM. E: elongation, R: reproductive, V: vegetative, DAP: days after pollination. B) RT-qPCR of *PvTPS04* across in above and belowground tissues. Expression was calculated following the $2^{-\Delta\Delta CT}$ method and normalized to actin.

Pv*TPS04 *produces (-)-borneol in leaf tissues

Borneol is commonly found in nature in two different enantiomers, (+/-)-borneol. To determine which stereoisomer is produced and accumulated in switchgrass, we examined the stereospecificity of borneol produced by the recombinant *Pv*TPS04 enzyme and in root hexane extracts in comparison to authentic borneol standards using GC chiral chromatography combined with mass spectrometry. We found that the *Pv*TPS04 protein produced exclusively (-)-borneol in vitro when GPP was provided as a substrate (Fig. 3). Additionally, only (-)-borneol was detected in root tissues suggesting *Pv*TPS04 is likely responsible for the formation of (-)-borneol in planta.

Although *PvTPS04* is most highly expressed in leaves, borneol is only accumulated in root tissues (Fig. 1; Muchlinski et al., in review). To determine whether *Pv*TPS04 enzyme activity

correlates with the transcript abundance of *PvTPS04*, we prepared crude protein extracts from both tissues for in vitro enzyme assays providing GPP as a substrate. We found that borneol is only produced in protein extracts of leaf tissues and not roots, indicating borneol formation likely occurs in leaf tissues in planta (Supplemental Fig. S2)

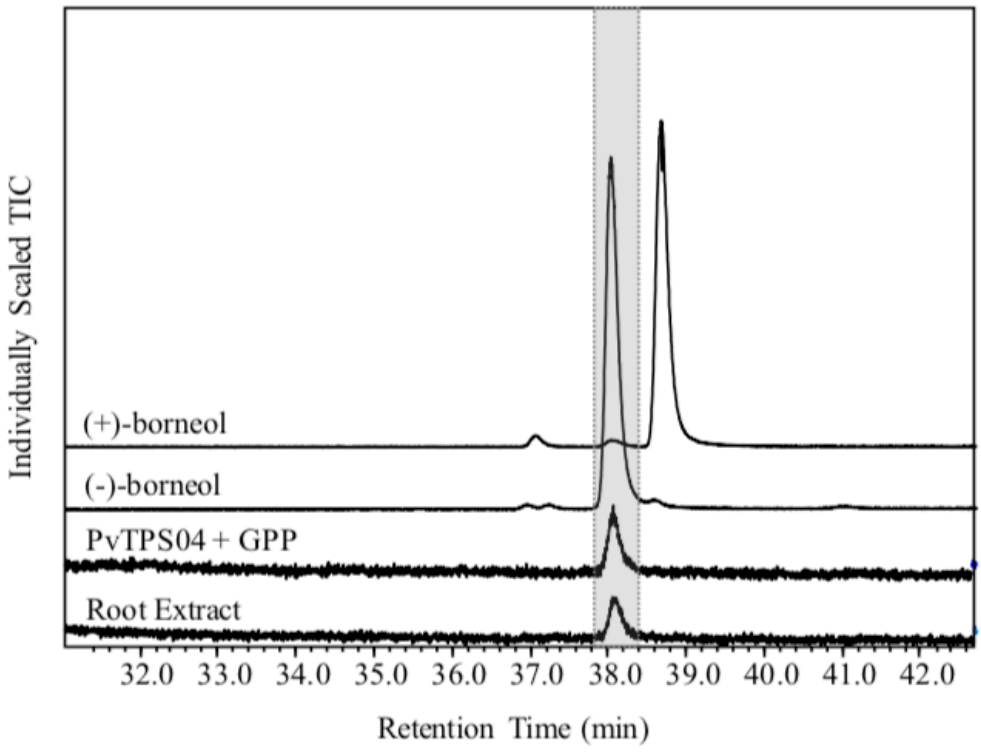


Figure 3. Chiral GC/MS analysis of (-) borneol produced in switchgrass roots and by the recombinant enzyme *Pv*TPS04. (-)-borneol is highlighted in gray. GPP, geranyl diphosphate. Authentic borneol standards were obtained from Sigma-Aldrich.

Salicylic acid modulates borneol and camphor levels in switchgrass roots

The phytohormone salicylic acid (SA) has multiple roles in plants such as in the activation of defensive pathways in response to pathogen infection and in regulating plant growth and development (Vicente and Plasencia, 2011). As borneol is known to have antimicrobial properties (Tabanca et al., 2001), we hypothesized that the compound could be involved in root defense

against pathogens and regulated by SA. To test this hypothesis, we grew switchgrass plants (cv. Alamo) in growth chambers for six weeks using an aeroponic culture system with calcined clay substrate adapted from Vaughan et al. (2011) (see MATERIALS AND METHODS). Plants were watered with 100 ml of a 50 µM solution of SA or mock solution (ethanol and pH adjusted) and incubated for 18 h. Root tissues were then rinsed, flash frozen in liquid nitrogen, and stored at -80^{o}C for chemical analyses. We found that treatment of plants with SA resulted in an overall increase in borneol levels, however this trend was not found to be significant ($p>0.05$, Student's T-Test, Fig. 4A). Interestingly, camphor levels significantly decreased in roots following SA treatment, suggesting conversion of borneol to camphor is downregulated or inhibited in response to SA treatment ($p=0.012$, Student's T-Test, Fig. 4B). The results indicate that the observed changes in root terpene profiles might in part be associated with a SA-dependent signaling response.

Switchgrass rhizosphere microorganisms alter terpenoid levels in roots of naïve plants

Based on the results from SA treatment, we sought to better understand the relationship between soil microorganisms and switchgrass root terpenoid levels. To this end, we isolated culturable microbes from the switchgrass rhizosphere (from plants grown in field soil under greenhouse conditions as described above), and applied an inoculant solution to non-soil grown plants identical to the experimental system used for SA treatments. We found that roots of inoculated plants had significantly higher levels of borneol compared to the mock control ($p=0.009$, Student's T-Test, Fig. 5A). Similar to results from treatments with SA, camphor levels significantly decreased in inoculum-treated roots ($p=0.05$, Student's t-test, Fig. 5B). In general, we observed that plants grown in the calcined clay substrate accumulate higher levels of camphor

compared to soil grown plants. However, when exposed to the microbial inoculum, camphor levels in calcined clay grown plants returned to levels typically observed for soil grown plants (Supplemental Fig. S3).

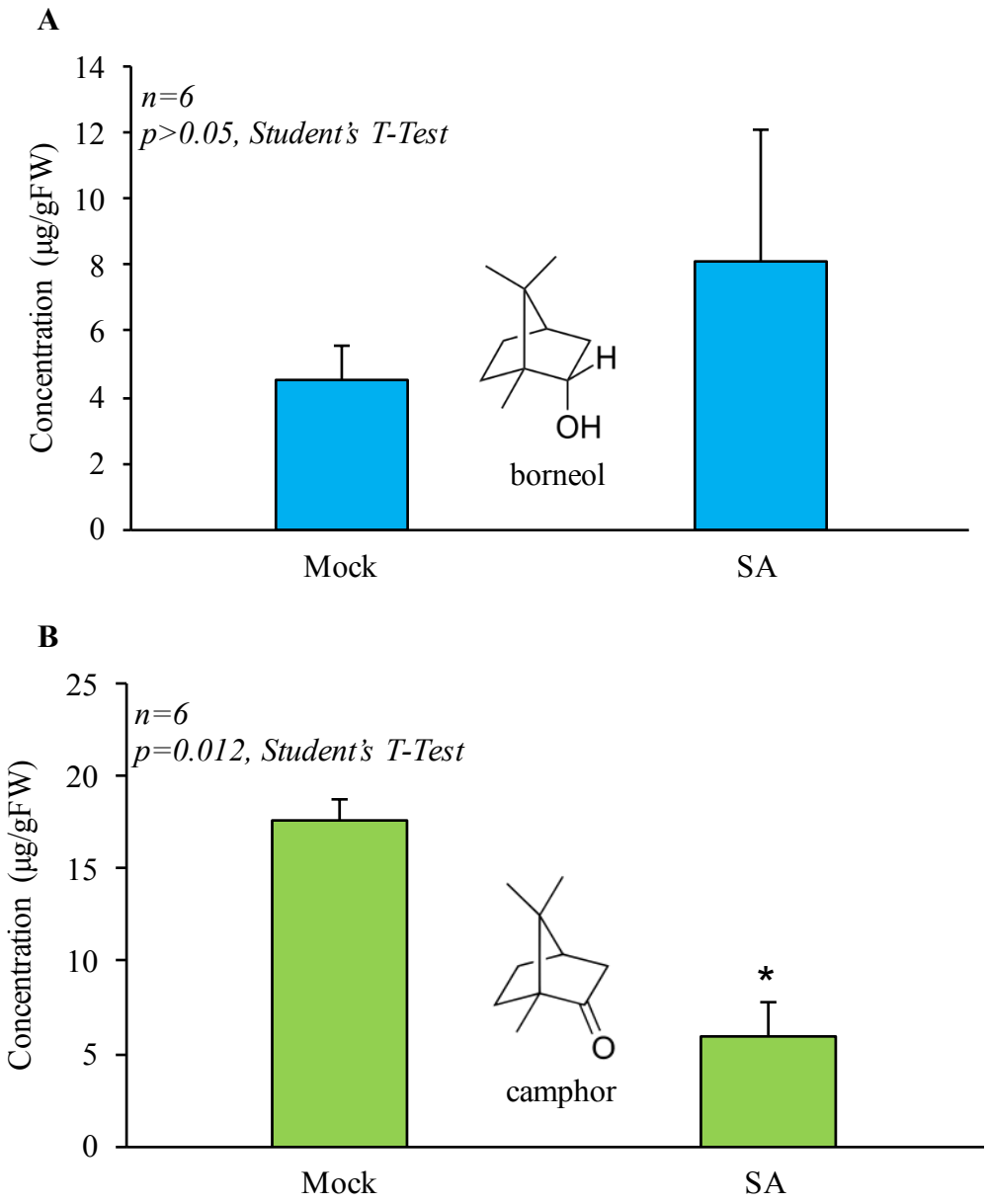


Figure 4. Effect of salicylic acid (SA) treatment on borneol (A) and camphor (B) levels in switchgrass roots (cv. Alamo) grown in calcined clay substrate. Plants were treated for 18 h with 100 ml SA (50 µM) or acidified water with ethanol as a mock control. Statistical significance was considered when $p \leq 0.05$ where $\alpha = 0.05$. Asterisks indicate level of significance.

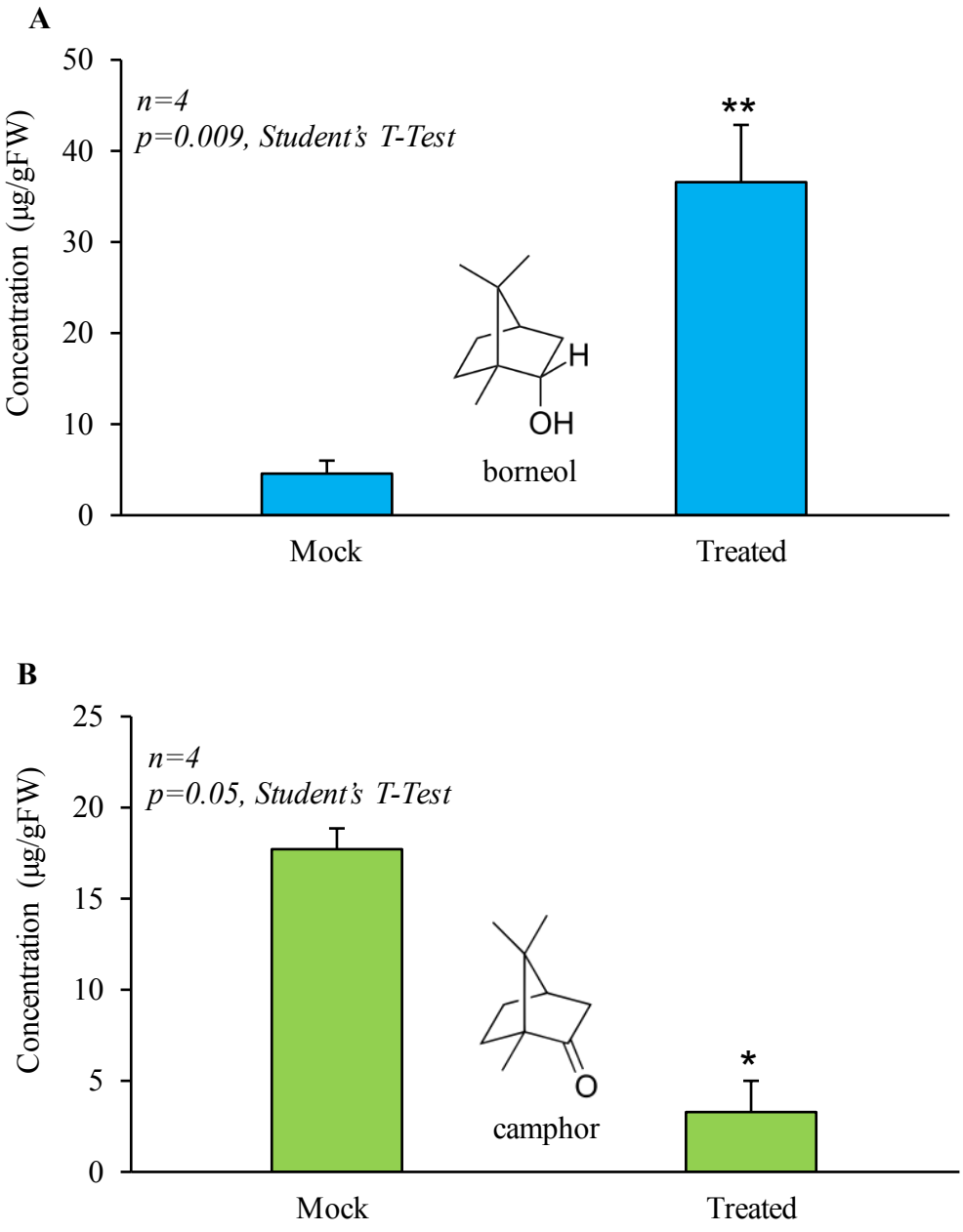


Figure 5. Effect of microbial inoculation on borneol (A) and camphor (B) levels in switchgrass roots (cv. Alamo) grown in calcined clay substrate. Plants were treated for 6 h with 100 ml of resuspended rhizosphere microorganisms or phosphate buffered saline solution as a mock control. Statistical significance was considered when $p \leq 0.05$ where $\alpha = 0.05$. Asterisks indicate level of significance.

The switchgrass root endosphere is enriched with Proteobacteria

Following our observation that treatment with rhizobacteria can influence terpenoid biosynthesis and accumulation in switchgrass roots, we asked the question whether the accumulation of terpenoids in the roots has in turn effects on the root bacterial community, especially in the endophytic compartment. To begin to answer this question, we compared the bacterial communities in the endosphere, rhizosphere and bulk soil (bulk) of switchgrass (cv. AP13) by using partial 16S rRNA gene sequencing and taxonomic diversity and abundance analysis. Plants were propagated from tillers provided by the Noble Foundation and cultivated under greenhouse conditions in field soil as described above. Samples were collected once plants reached the R1 growth stage (~3 months). We found that Proteobacteria are enriched in the switchgrass endosphere compared to the rhizosphere and bulk soil communities (Fig. 6; Table 1). In addition, Acidobacteria, Planctomycetes, and Verrucomicrobia show trends for decreased abundance in the endosphere compartment compared to the rhizosphere and bulk soil.

Principle Coordinate Analysis (PCoA) with a Bray-Curtis distance matrix based on operational taxonomic units OTUs was used to visualize community wide differences (Fig. 7). The first two axes are shown, where the first axis explains over half the variance (54.4%), and there is a strong separation between endosphere and the other root associated zones (rhizosphere, bulk soil). Statistical significance was tested using PERMANOVA, and the results indicated that the bulk soil and rhizosphere samples are indistinguishable ($p>0.05$, PERMANOVA). However, the endosphere bacterial community is significantly different from both bulk soil ($p=0.015$, PERMANOVA) and rhizosphere ($p=0.012$, PERMANOVA). We additionally found a significant difference in dispersion among the zones ($p=0.004$, PERMANOVA), indicating that endosphere communities are overall more variable than bulk soil and rhizosphere communities.

To determine if specific taxa (genera) were enriched in specific zones, we used Indicator Species Analysis. Indicator value was calculated as: Indicator $Value_{ij} = A_{ij} * B_{ij} * 100$; where i is the species (genus) and j is the group (zone), Aij is the average number of genera (i) within zone (j) divided by the sum of the average of the genus across all zones. Bij is the number of sites for which the species is present within the zone divided by the number of sites in the zone (Neuhäuslová-Novotná, 1976). Results from Indicator Species Analysis are presented in Table 1 for significantly enriched endosphere taxa, where significance was calculated by PERMANOVA (n=999). We found that 27 genera were enriched in the root endosphere, over half of which belonged to the Proteobacteria. We found that Verrucomicrobia and Actinobacteria were significantly enriched in the rhizosphere and bulk soil samples as compared to the endosphere, respectively (Supplemental Table. S1). Collectively these results indicate that the endosphere of switchgrass roots represents a distinct bacterial microenvironment compared to other root-associated zones.

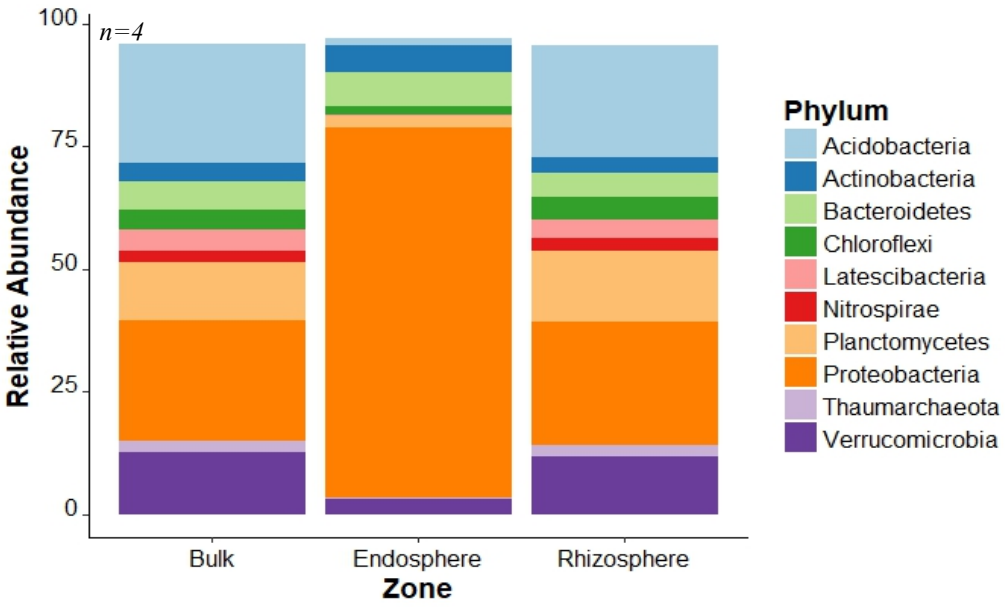


Figure 6. Stacked bar plot showing average relative abundance of phyla across bulk soil (Bulk), endosphere and rhizosphere zones of soil grown switchgrass plants.

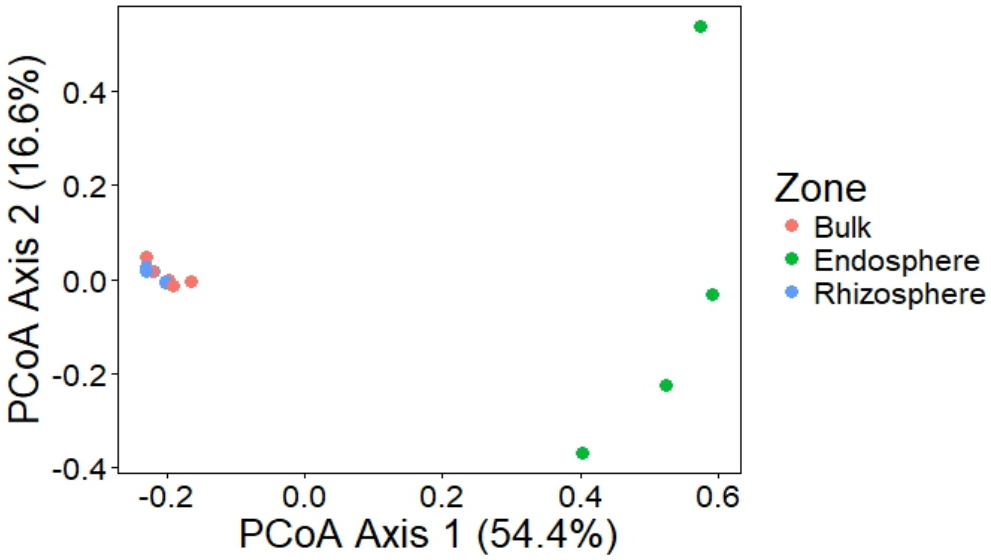


Figure 7. Principle coordinate analysis (PCoA) based on the Bray-Curtis distance matrix of OTUs from bulk soil (Bulk), endosphere and rhizosphere samples of soil grown switchgrass plants.

Table 1. Analysis of indicator species enriched in the endosphere compartment of switchgrass roots. Calculated indicator value is reported under the Stat column. Significance was determined by PERMANOVA where alpha = 0.05.

Putative Taxonomic Rank	Stat	p-value
Proteobacteria_Betaproteobacteria_Burkholderiales_Comamonadaceae_Pelomonas	0.996	0.005**
Proteobacteria_Alphaproteobacteria_Caulobacterales_Caulobacteraceae_Asticcacaulis	0.994	0.006**
Proteobacteria_Gammaproteobacteria_Pseudomonadales_Moraxellaceae_Acinetobacter	0.991	0.005**
Verrucomicrobia_Verrucomicrobiae_Verrucomicrobiales_Verrucomicrobiaceae_Brevifollis	0.988	0.006**
Bacteroidetes_Cytophagia_Cytophagales_Cytophagaceae_Emticicia	0.986	0.003**
Proteobacteria_Betaproteobacteria_Burkholderiales_Oxalobacteraceae_Janthinobacterium	0.985	0.003**
Bacteroidetes_Cytophagia_Cytophagales_Cytophagaceae_Flexibacter	0.977	0.003**
Proteobacteria_Betaproteobacteria_Burkholderiales_Comamonadaceae_Paucibacter	0.974	0.005**
Fibrobacteres_Fibrobacteria_Fibrobacterales_Fibrobacteraceae_uncultured	0.972	0.013*
Proteobacteria_Alphaproteobacteria_Rhodospirillales_Rhodospirillaceae_Thalassospira	0.97	0.003**
Proteobacteria_Gammaproteobacteria_Enterobacteriales_Enterobacteriaceae_Escherichia	0.963	0.008**
Proteobacteria_Alphaproteobacteria_Caulobacterales_Caulobacteraceae_Caulobacter	0.947	0.049*
Proteobacteria_Gammaproteobacteria_Xanthomonadales_Xanthomonadaceae_Tahibacter	0.945	0.024*
Spirochaetae_Spirochaetes_Spirochaetales_Leptospiraceae_Turneriella	0.938	0.024*
Proteobacteria_Alphaproteobacteria_Sphingomonadales_Sphingomonadaceae_Sphingopyxis	0.889	0.006**
Bacteroidetes_Sphingobacteriia_Sphingobacteriales_Chitinophagaceae_Taibaiella	0.866	0.011*
Deinococcus-Thermus_Deinococci_Deinococcales_Deinococcaceae_Deinococcus	0.866	0.011*
Proteobacteria_Alphaproteobacteria_Rhizobiales_Methylobacteriaceae_Methylobacterium	0.866	0.011*
Proteobacteria_Betaproteobacteria_Neisseriales_Neisseriaceae_Vogesella	0.866	0.011*
Proteobacteria_Betaproteobacteria_Nitrosomonadales_Gallionellaceae_uncultured	0.866	0.011*
Proteobacteria_Gammaproteobacteria_Enterobacteriales_Enterobacteriaceae_Enterobacter	0.865	0.05*
Proteobacteria_Betaproteobacteria_Burkholderiales_Oxalobacteraceae_Duganella	0.864	0.031*
Bacteroidetes_Flavobacteriia_Flavobacteriales_Flavobacteriaceae_Chryseobacterium	0.861	0.019*
Bacteroidetes_Sphingobacteriia_Sphingobacteriales_Sphingobacteriaceae_Mucilaginibacter	0.859	0.019*
Proteobacteria_Alphaproteobacteria_Caulobacterales_Caulobacteraceae_Brevundimonas	0.842	0.047*
Actinobacteria_Actinobacteria_Micrococcales_Microbacteriaceae_Lysinimonas	0.82	0.043*
Proteobacteria_Betaproteobacteria_Methylophilales_Methylophilaceae_Methylophilus	0.793	0.036*

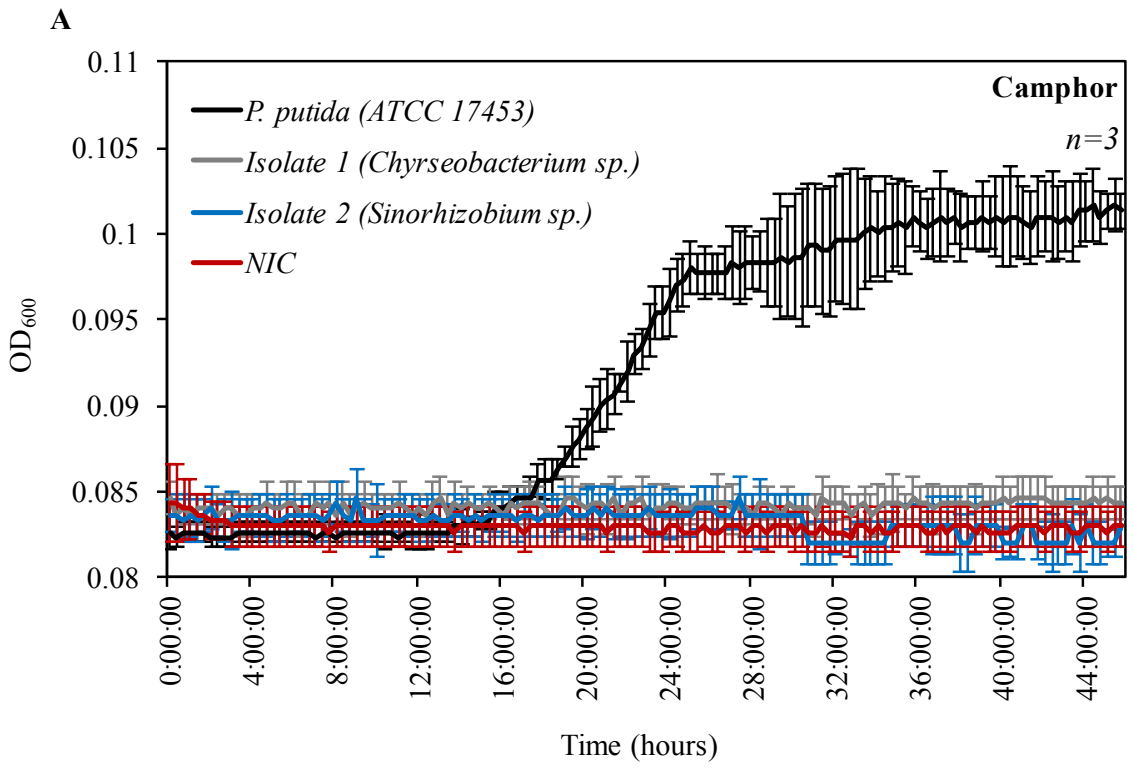
Isolation of root endophytic bacteria on monoterpene containing medium

To better understand a possible relationship between root bacterial endophytes and monoterpenoid production, we hypothesized that selected root endophytes may be able to metabolize and/or detoxify terpenoids, as has been observed for camphor metabolism by isolates

of the common soil bacterium *Pseudomons putida* (Prasad et al., 2013; Tsang et al., 2016). To this end, we selectivity enriched, isolated, and characterized culturable root endophytes providing only terpenoids (camphor or borneol) as a sole carbon source (see MATERIALS AND METHODS). Briefly, surface sterilized root samples were placed in liquid culture with minimal medium supplemented with borneol or camphor at natural and ~10x above and below natural levels. Cultures were then incubated at 28^{o}C with gentle shaking (120 RPM). Following a one-week incubation, 100 µl of culture was transferred to a fresh preparation of terpenoid supplemented medium and incubated for a second week. Selection via subculturing continued for two additional weeks prior to aliquots being plated on solidified minimal medium supplemented with pure terpenoid compound. Bacterial colonies were identified by 16S rRNA amplification, Sanger sequencing, and assigned putative annotation based on comparisons to sequences databases (NCBI).

In total, we identified ~40 isolates, and putatively classified them into seven genera based on partial 16S rRNA sequence comparisons to publically available data (NCBI BLAST, Supplemental Table S2). Some of the species in these genera such as *P. putida* have been previously reported to utilize terpenoids as a sole carbon source (Tsang et al., 2016). In addition, many species of the predicted genera have known plant growth promoting roles, suggesting a possible terpene-dependent recruitment of N-fixing and growth promoting bacteria to the rhizosphere and endosphere (Ker et al., 2012; 2014; Wang et al., 2015). To further determine if these isolates were able to metabolize terpenoid compounds, we selected and cultured two endosphere isolates (*Chryseobacterium sp.* and *Sinorhizobium sp.*) and monitored their growth (OD_{600}) in liquid medium while providing only camphor or borneol as the sole carbon source. These isolates were specifically selected based on significant enrichment in the switchgrass root

endospore (*Chryseobacterium*, Table. 1, p=0.019, PERMANOVA) or their known role in N-fixation such as *Sinorhizobium* (Coronado et al., 1995). We found that neither isolate showed enhanced growth over approximately two days on borneol or camphor as the sole carbon source when supplied at concentrations naturally occurring in the root (Fig. 8A and B). We continued to monitor for one week with no growth observed. We concluded, therefore, that these two isolates may not be specialists and simply tolerate borneol and camphor based on the original enrichment conditions. Specifically, these bacteria may show growth only under mixed culture conditions and are capable of inactivating borneol and camphor but not necessarily depend on these compounds as sole carbon source.



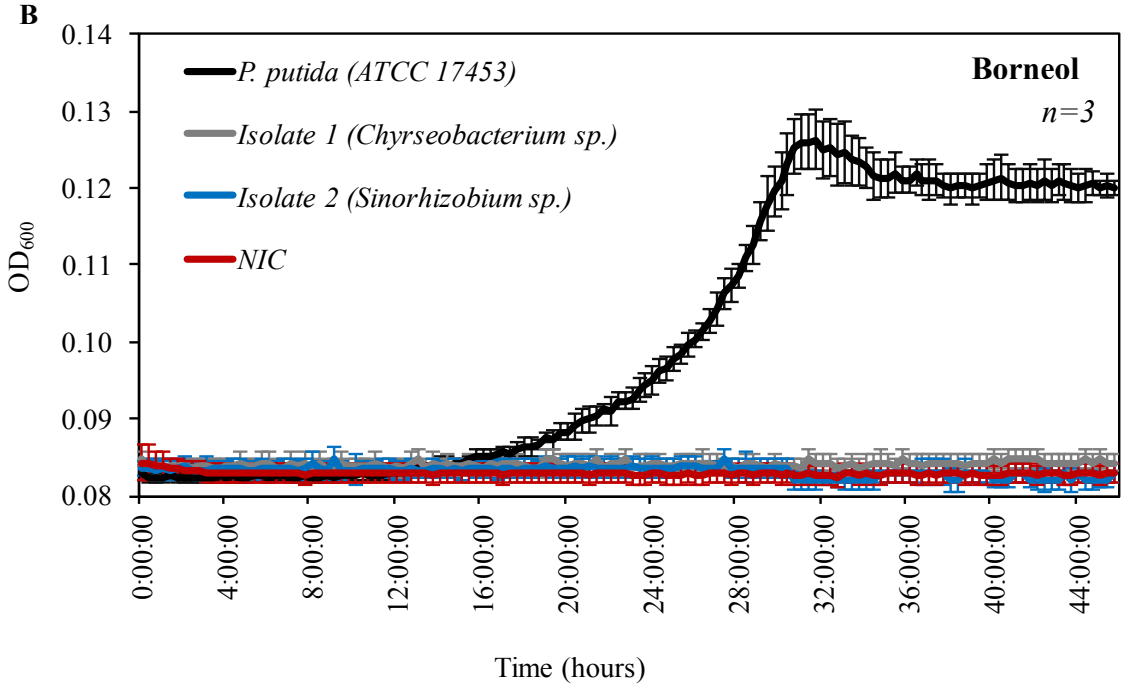


Figure 8. Growth curve analysis of endosphere isolated bacteria. Bacteria were inoculated in triplicate in a minimal medium containing camphor (A) or borneol (B) as the only carbon source. Growth measurements were taken every 20 min for 45 h. NIC: non-inoculated control. *P. putida* strain ATCC 17453 was used as a positive control.

TPS04 RNAi plants produce lower borneol levels compared to specific wild-type lines

To further evaluate if monoterpenoids are necessary for maintaining a defined root microbiome, we used an RNA interference based gene silencing approach (RNAi) to reduce borneol levels in roots of switchgrass. Briefly, a 400 bp fragment from the 3'-UTR of *PvTPS04* was amplified and cloned into the monocot specific RNAi vector pANIC12a (see MATERIALS AND METHODS). The vector construct was then introduced into switchgrass callus by

Agrobacterium-mediated plant transformation in the lab of Dr. Neal Stewart (UT Knoxville) following established methods (Xi et al., 2009). Regenerated plants were provided as tillers for two wild-type (WT) lines (A, B) and six transgenic lines (TL) and were used throughout this study. WTA and TLs were derived from the same callus, whereas WTB was derived from different callus material. Based on initial RT-PCR screening, we selected three transgenic lines for downstream analyses (TL2, 4 and 6). Comparisons of the two wild-type and three transgenic lines by quantitative RT-PCR indicated significant reduction in *PvTPS04* expression (Fig. 8A, p=0.0034, ANOVA). However, post-hoc comparisons showed that *PvTPS04* expression was only significantly higher in one wild-type line (B) compared to the transgenic lines (Fig. 8A, tps04-2: p=0.0074, tps04-4: p=0.0073, tps04-6: p=0.0072, Tukey's HSD). No significant difference was detected between wild-type line A and transgenic lines (Fig. 8A, p>0.05, Tukey's HSD).

To determine if the root terpenoid profile had been modified as a result of the *PvTPS04* gene knockdown, borneol levels were quantified from extracts of wild-type and transgenic switchgrass lines (Supplemental Table. 3). Consistent with the RT-qPCR results, we found significantly higher levels of borneol in wild-type line B compared to the three transgenic lines (Fig. 8B, p=0.0004, ANOVA). However, we also found a significant difference in borneol levels between wild type lines (Fig. 8B, p=0.001, Tukey's HSD), and no significant differences in borneol levels occurring between wild-type line A and the transgenic lines. Despite borneol levels being slightly reduced in some RNAi lines, the high variability in borneol levels in wild-type line B, and the differences between wild-type lines resulted in halting further microbial analyses. Additionally, wild type line B and the knockdown lines were regenerated from different calli, which raises questions about a uniform genetic and biochemical background. We did, however, backup our switchgrass RNAi work by utilizing a closely related model system *Setaria viridis*

which also accumulates borneol (albeit at lower levels) in root tissues. Details of this work are described in Appendix B.

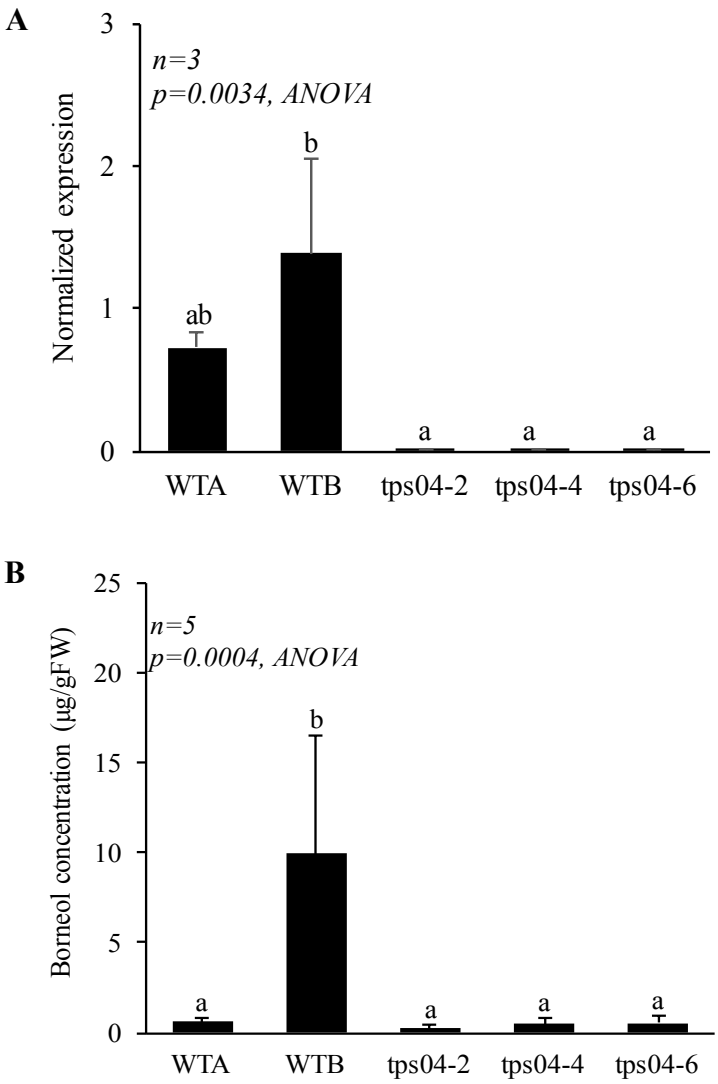


Figure 9. Analysis of *PvTPS04* expression and borneol production in wild-type and transgenic switchgrass. A) RT-qPCR of *PvTPS04* in leaves of wild-type (WT) and transgenic lines (2, 4 and 6). Expression was calculated following the $2^{-\Delta\Delta CT}$ method and normalized to actin. B) GC/MS analysis of borneol levels in roots of wild-type (WT) and transgenic lines (2, 4 and 6). Statistical significance was considered when $p \leq 0.05$ where $\alpha=0.05$. Letters indicate significant differences based on post-hoc Tukey HSD comparisons.

DISCUSSION

Plants produce a bouquet of volatile organic compounds (VOCs) to facilitate interactions with a wide range of beneficial and harmful organisms. In this study, we aimed to better understand the biological function of monoterpenoids in grasses, specifically borneol, which accumulates at high (μg) levels in switchgrass roots. Here we showed that domesticated and natural switchgrass cultivars produce borneol (and camphor) irrespective of ploidy level or ecotype (Fig. 1). The high conservation of these compounds (especially in domesticated cultivars) indicated that these compounds are likely serving important biological functions in switchgrass. However, little knowledge exists for the roles of monoterpenoids in belowground interactions. Formation of the monoterpene 1,8-cineole has been demonstrated in *A. thaliana* roots (Chen et al., 2004), and levels were shown to increase in response to bacterial and fungal pathogens in vitro (Steeghs et al., 2004). In addition, 1,8-cineole has been shown to deter insect herbivores and exhibit negative allelopathic effects on a variety of plant species (Tripathi et al., 2001; Singh et al., 2002). Although structurally similar, far less is known about borneol and camphor biosynthesis and function in plants. Studies of borneol and camphor production have mostly been restricted to the mint family (Lamiaceae), where both compounds occur at high levels in trichomes as constituents of essential oils (Zaks et al., 2008). Unlike *Arabidopsis* where 1,8-cineole is produced and accumulated in roots, we found that biosynthesis of borneol in switchgrass occurs in leaf tissues (Fig. 2). Although at first surprising, an identical mechanism exists in sage (*Salvia officialalis*, Lamiaceae) where borneol is formed in leaves and then transported to roots where it is further metabolized to camphor (Croteau and Karp, 1979; Croteau and Shaskus, 1985; Croteau et al., 1987). The proposed reason for this is to recover mobile carbon from senescing oil glands (Croteau et al., 1987), however it is unclear if similar transport and recovery is occurring in

switchgrass. As borneol and camphor only accumulate in root tissues, we speculate that borneol formation occurs in leaves due to an effective C4 photosynthesis and formation of higher pools of GPP in photosynthetic chloroplasts than non-photosynthetic leucoplasts. Perhaps a borneol derivative such as bornyl acetate is then transported via the phloem into roots, where borneol is released and stored. In addition, the exact site of storage is not known but might be similar to parenchyma cells in the cortex of vetiver grass roots (Del Giudice et al., 2008). Further tests (e.g. labeling studies, phloem sap analysis) are needed to conclusively resolve borneol transport in switchgrass.

Following treatments with microbes from the switchgrass rhizosphere, we found that "naïve" switchgrass plants increase borneol production in their roots (Fig. 5A). This stimulating effect might lead to higher "constitutive" levels of borneol at an early stage of seedling growth to establish chemical defenses against herbivores or fungal and oomycete pathogens. Similar observations have been made in vetiver grass, where bacterial inoculation is necessary for terpene accumulation in roots (Adams et al., 2008; Del Giudice et al., 2008). Although not statistically significant, SA treatment of switchgrass grown in calcined clay led to increased borneol accumulation in roots (Fig. 4A) indicating that the response observed upon bacterial inoculation may be mediated in an SA-dependent manner. As we observed previously that SA treatment of roots causes increased transcript abundance of *PvTPS04* in leaves (Muchlinski et al., in review), it is possible that borneol biosynthesis is inducd in leaves and the compound is transported to the roots as discussed above. In addition, the accumulation of borneol in roots might be caused in part by the reduced conversion to camphor. It should be pointed out that treatments of switchgrass seedlings grown in potting substrate with a 5 mM SA solution did not lead to an increase in the emission of borneol from roots (see Chapter 2). While in this study no internal levels of borneol

were measured, it remains to be determined whether the response in terpene levels is dependent on the applied concentration of SA.

Since root-associated bacteria have been shown to metabolize volatile terpenes (Del Guice et al., 2008) and soil-borne bacteria such as Pseudomonads, can break down a wide range of carbon substrates including monoterpenoids (Prasad et al., 2013; Tsang et al., 2016) we assumed that bacteria specialized in using borneol or camphor as sole carbon source would reside in the switchgrass root endosphere. Although we identified and cultured many root endophytes, we found no clear evidence under our culture conditions that these endophytes metabolize monoterpenoids as their sole carbon source (Fig. 8). Therefore, we concluded that culturable switchgrass endophytes may only tolerate borneol and/or camphor or co-metabolize them with other carbon sources. It is also possible that the bacterial endophytes of switchgrass are not in immediate contact with the stored terpenes and, hence, might use these compounds only as a co-substrate.

In this study we show that the formation of borneol occurs by a similar mechanism as that known from essential oil producing plants in the mint family. In addition, we found that both microbial inoculations and SA led to increased borneol levels in roots. We predicted that borneol might be necessary for shaping the root microbiome. So far, we have been unable to provide sufficient evidence for this effect using *Setaria* as an alternative system (Appendix B). While generating *Setaria* RNAi lines that do not accumulate borneol has been successful, borneol levels in wild type Setaria roots are approximately 100 fold lower than those in switchgrass and, hence, might have limited activity on microbes in the rhizosphere and endosphere. Despite challenges in our transgenic approach in switchgrass, comparison of a "high borneol" producing wild type line with the "low borneol" producing RNAi or even wild-type lines may still be sufficient to determine

the effect of borneol on the root microbial community. However, more confirmation would be needed to demonstrate that only terpenoid levels are altered between the different genetic backgrounds. Additional studies could use these lines to further investigate defensive activities of borneol against specific microbial pathogens. More recently, borneol has been shown to affect root growth in both *Arabidopsis* and *Artemisia* (Horiuchi et al., 2007; Tian et al., 2014), supporting an alternative role in allelopathy and/or root development, which could also be addressed in switchgrass. Results from this study highlight the intricacies of investigating biological function of specialized metabolites in plants. We hope the groundwork laid here provides a baseline for continued investigation of terpenoid metabolism and function in complex polyploid model systems.

MATERIALS AND METHODS

Plant growth conditions (greenhouse)

Seed stocks for switchgrass cultivars were purchased from ERNST seeds (www.ernstseed.com/), and grown under greenhouse conditions at Virginia Tech (Blacksburg, VA, USA). Unless otherwise stated, plants were maintained in plastic pots (17 cm diam., 20 cm high, with four holes at the bottom for drainage) and filled with 3.7 kg of a soil and sand mixture (soil:sand = 2:1 v/v, sand: 0.1-1.0 mm diam.) collected from an established switchgrass field in Blacksburg, VA, USA (Moore Farm). Plants were irrigated every two days and grown at temperatures of 30±1^{o}C/25±1^{o}C (day/night), a 14 h photoperiod, 75% relative humidity, and with photosynthetically active radiation of approximately 500 μmol $m^{-2}s^{-1}$. Plant material for the switchgrass genomic model AP13 were kindly provided as tillers from the Noble Foundation and maintained under greenhouse conditions as described above.

Plant SA treatment and growth conditions

Plants (cv. Alamo) were germinated from seed and cultivated to the Elongation 2 (E2) growth stage (Hardin et al., 2013), in red Solo© cups in calcined clay substrate adapted from Vaughan et al. (2013). Plants were maintained at 30^{o}C in a Percival growth chamber under 16 h light/8 h dark. For SA treatments, plants (6-mock EtOH only and 6-treated) were watered with 100 ml of a 50 μM solution of SA (Sigma-Aldrich) and incubated for 18 h. Root tissues were rinsed, flash frozen in liquid nitrogen, and stored at -80^{o}C for chemical analyses.

Plant treatment with rhizobacteria

For microbial inoculations, plants of the cultivar Alamo were germinated from seed and cultivated to the E2 growth stage in red Solo© cups in calcined clay substrate as described above. Established greenhouse grown Alamo plants in field soil were used as a source of rhizobacteria. Soil (1 g) was collected by physically removing adhered particles from roots and suspended in 1 ml of PBS buffer (pH=7). The solution was gently mixed on a rocking table and pelleted by light centrifugation. The resulting supernatant was serially diluted and plated on LB/agar and grown at 28^{o}C for one week. Treatments were applied to plants at the E2 growth stage by a sterile transfer of the colonies on the microbial plate and suspension into 100 ml of PBS buffer (pH=7). The resulting inoculant solution (or PBS mock) was applied as a 25 ml volume to four plants and incubated for six hours. Following incubation, root material was harvested and analyzed for terpenoid compounds by GC/MS as described below.

Analysis of volatile terpene compounds in roots of switchgrass

Six cultivars of switchgrass (Alamo, Dacotah, Trailblazer, Sunburst, Pathfinder and Greenville) were grown as described above in a randomized block design with three biological replicates. Total root mass was collected from mature plants once they reached the R1 growth stage (Hardin et al., 2013). Tissues were rinsed, flash frozen in liquid nitrogen, and stored at -80^{o}C for chemical analyses. Samples (1 g) were ground to a fine powder for 2 min in the presence of liquid nitrogen, weighed, transferred to 5 ml hexane and mixed by vortex for 20 s. The ground material was placed in an ultrasonic bath (Fisher Scientific) for 10 min then pelleted by centrifugation. Following the collection of two 5 ml fractions, 1-bromodecane was added for a final concentration of 20 ng/µl as an internal standard and extracts were dried over a $MgSO_4$ mock column and concentrated on ice to ~40 µl under a gentle stream of nitrogen. Extracts were analyzed using a gas chromatograph (240^{o}C injector port) coupled with a quadrupole mass spectrometer (GC-MS-QP2010S, Shimadzu). Extracts were separated with a 2:1 split on a 30 m x 0.25 mm ID x 0.25 µm Zebron capillary column (Phenomenex) using Helium as the carrier gas (1.4 ml min^{-1} flow rate) and a temperature gradient of 5^{o}C min^{-1} from 40^{o}C (hold 2 min) to 220^{o}C. Identification of major volatile compounds was confirmed by comparisons of retention times and mass spectra to authentic standards when available and mass spectral libraries (Wiley and NIST). Absolute amounts of borneol were determined based on calibration with authentic borneol (Sigma-Aldrich) and normalization to gram fresh weight.

Determination of borneol stereochemistry

Root hexane extracts of cv. Alamo described above were analyzed by chiral separation on a Machery-Nagel fused silica capillary column (HYDRODEX-ß-3P) using a temperature program

from 40°C to 230°C with a ramp speed of 2°C/min with a QP2010S GC/MS as described above (Shimadzu). Borneol was produced from recombinant *Pv*TPS04 enzyme by incubating partially purified protein as described in Chapter II with 60 μM GPP and 10 mM $MgCl_2$ for 1 hour at 30°C with the addition of a 1 ml hexane overlay. Following incubation, hexane fractions were dried over $MgSO_4$, concentrated to ~40 μl, and 1 μl was injected into the GC/MS using the methods described above. Identification of borneol enantiomers was confirmed by comparisons of retention times to authentic standards (Sigma-Aldrich).

In silico evaluation of PvTPS04 expression

RNA-seq data were kindly provided by the Noble foundation (https://www.noble.org) and were assembled in house using Trinity (Grabherr et al., 2011). Assembled transcriptomes were queried for TPS sequences using NCBI tBLASTx. Resulting hits were manually curated for matches to *Pv*TPS04. Raw reads from each assembly were mapped to the respective *Pv*TPS04 contig using bowtie (bowtie-bio.sourceforge.net/) and indexed to retrieve read counts.

Bacterial 16S rRNA analysis from the switchgrass endosphere and rhizosphere

Plants (cv. AP13) for this experiment were cultivated as described above under greenhouse conditions (Blacksburg, Virginia Tech) and harvested at the elongation 4 growth stage (Hardin et al., 2013). DNA samples were extracted from five biological replicates of bulk soil and rhizosphere soil using the PowerSoil DNA extraction kit (MoBio). For endosphere DNA sample collection, roots were first rinsed and surface sterilized following Del Giudice et al. (2008). To prepare samples for Illumina MiSeq, a conserved region of the 16S ribosomal gene was amplified in triplicate using barcoded 8F and 1492R primers and cleaned reaction products were verified by

agarose gel electrophoresis and quantified using a NanoDrop UV-Vis spectrophotometer (Thermo Scientific). Samples were sequenced at the Biocomplexity Institute (Virginia Tech) on a MiSeq platform generating 250 million PE reads. Data analysis was done in house following standard protocols using the QIIME data analysis pipeline (Appendix C, www.qiime.org).

Isolation and characterization of switchgrass root endophytes

Root tissue from switchgrass (cv. AP13) described above was surface sterilized following Del Giudice et al. (2008), pulverized to a uniform powder, and added to 10 ml of a minimal medium containing: 1X M9 minimal salts (Sigma-Aldrich), trace metal solution (13.4 mM EDTA, 3.1 mM $FeCl_3$-$6H_2O$, 0.62 mM $ZnCl_2$, 76 µM $CuCl_2$-$2H_2O$, 42 µM $CoCl_2$-$2H_20$, 162 µM H_3BO_3, 8.1 µM $MnCl_2$-$4H_2O$, 0.1 mM $NiCl_2$-$6H_2O$, 0.14 mM Na_2MoO_4-$2H_2O$), vitamin solution (Folic acid 2.0 mg/L, Pyridoxine hydrochloride 10.0 mg/L, Riboflavin 5.0 mg/L, Biotin 2.0 mg/L, Thiamine 5.0 mg/L, Nicotinic acid 5.0 mg/L, Calcium Pantothenate 5.0 mg/L, Vitamin B_{12} 0.1 mg/L, p-Aminobenzoic acid 5.0 mg/L, Monopotassium phosphate 900.0 mg/L), 1 mM $MgSO_4$, 0.3 mM $CaCl_2$, 3 mM $(NH_4)_2SO_4$ and supplemented with either 4, 40 or 400 µg/ml (-)-camphor or (-)-borneol (Sigma-Aldrich) dissolved in methanol. Cultures were then incubated at 28^{o}C with gentle shaking (120 RPM). Following a one-week incubation, 100 µl of culture was transferred to a fresh preparation of terpenoid supplemented medium and incubated for one week. Selection via subculturing continued for an additional three weeks prior to aliquots being plated on solidified minimal medium (0.03% R2-A agar, Sigma-Aldrich) supplemented with pure terpenoid compound. Bacterial colonies were identified by 16S rRNA amplification (8F, 1496R primers), Sanger sequencing at the Biocomplexity Institute (Virginia Tech, Blacksburg VA) and assigned putative annotation based on comparisons to the NCBI sequence database. Growth curve analysis for selected isolates was achieved by culturing a single colony in medium for 24 h, adjusting OD 600 to 0.002, and inoculating sterile 96-well cultures of minimal medium supplemented with terpenoids or solvent in triplicate. Measurements of optical density were taken every 20 min for one-week using a BioTEK Synergy plate reader.

Generation of RNAi lines

A 400 bp region of the 3'-UTR of the *TPS04* gene was amplified with Q5 High-Fidelity DNA polymerase (ThermoFisher Scientific) in a 25 µl reaction volume with the following PCR conditions: 98^{o}C for 30 s, followed by 30 cycles of 98^{o}C for 30 s, 55^{o}C for 30 s, 72^{o}C for 1 min and a final extension at 72^{o}C for 2 minutes. The amplified fragment was gel purified using a NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel, MN) and concentrated to ~10 µl. A 10 µl A-tailing reaction was prepared with 3 µl of purified PCR product incubated in the presence of 10 mM dATPs and Taq polymerase (ThermoFisher Scientific) at 72^{o}C for 30 min. The resulting product was ligated overnight into the pGEM-T Easy vector (Promega) and Sanger sequenced to verify the insert. Plasmid (2 µg) was digested with restriction enzymes (BamHI, XhoI) and the resulting ~400 bp fragment was purified as described above and ligated overnight into the corresponding restriction sites of the entry vector pENTR11 (ThermoFisher Scientific). An LR clonase reaction (ThermoFisher Scientific) was then performed following the manufactures protocol with the monocot specific RNAi vector pANIC12a (ABRC Stock #1720) to generate the final construct. Agrobacterium-mediated plant transformation and regeneration was carried out in the lab of Dr. Neal Stewart following established methods (Xi et al., 2009). Provided tillers from two WT lines (A, B) and six transgenic lines were used throughout this study.

Total RNA extraction and RT-qPCR

Total RNA was extracted from the root tissue described above in biological triplicate using the TRIzol reagent (Life Technologies, Carlsbad, CA) and the Qiagen Plant RNAeasy kit (Qiagen) in accordance with the manufacturers protocol. RNA was treated on column for DNA contamination

with DNAse (Qiagen), and used for first strand cDNA synthesis using SuperScriptII reverse transcriptase and oligo(dT) 18 primer (Invitrogen) according to the manufacturer's instructions. RNA was first normalized between samples and replicates to 2.5 µg based on denaturing gel electrophoresis and spectrophotometer measurements at 260 nm. The resulting cDNA was diluted to 100 ng/µl. Reactions were performed with 1 µl cDNA in a 20 µl reaction using Power SYBR Green PCR master mix (Applied Biosystems) and gene specific primers (Supplemental Table S1). PCR amplifications were done in biological and technical replicates with a CFX96 Touch real-time PCR detection system (Bio-Rad) with the following cycles: 95^{o}C for 10 min, followed by 40 cycles of 95^{o}C for 15 s, 50^{o}C 30 s and 60^{o}C for 1 min. Melt curve analysis was performed at the end of amplifications to ensure specificity of each primer pair. Relative expression levels across tissues were calculated using the relative quantification method and normalized to actin (Livak and Schmittgen, 2001).

REFERENCES

Adams, R.P., Nguyen, S., Johnston, D.A., Park, S., Provin, T.L., and Habte, M. (2008). Comparison of vetiver root essential oils from cleansed (bacteria- and fungus-free) vs. non-cleansed (normal) vetiver plants. *Biochem. Syst. Ecol.* 36**,** 177-182. doi: 10.1016/j.bse.2007.10.004.

Badri, D.V., Chaparro, J.M., Zhang, R.F., Shen, Q.R., and Vivanco, J.M. (2013). Application of natural blends of phytochemicals derived from the root exudates of *Arabidopsis* to the soil reveal that phenolic-related compounds predominantly modulate the soil microbiome *J. Biol. Chem.* 288**,** 30503-30503. doi: 10.1074/jbc.A112.433300.

Beisner, B.E., Peres, P.R., Lindstrom, E.S., Barnett, A., and Longhi, M.L. (2006). The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology* 87**,** 2985-2991. doi: 10.1890/0012-9658.

Belhassen, E., Filippi, J.J., Brevard, H., Joulain, D., and Baldovini, N. (2015). Volatile constituents of vetiver: a review. *Flavour Fragrance J.* 30**,** 26-82. doi: 10.1002/ffj.3227.

Berg, G., and Smalla, K. (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol. Ecol.* 68**,** 1-13. doi: 10.1111/j.1574-6941.2009.00654.x.

Bulgarelli, D., Rott, M., Schlaeppi, K., van Themaat, E.V.L., Ahmadinejad, N., Assenza, F., Rauf, P., Huettel, B., Reinhardt, R., Schmelzer, E., Peplies, J., Gloeckner, F.O., Amann, R., Eickhorst, T., and Schulze-Lefert, P. (2012). Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* 488**,** 91-95. doi: 10.1038/nature11336.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E.V.L., and Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. *Ann. Rev. Plant Bio., Vol 64* 64**,** 807-838. doi: 10.1146/annurev-arplant-050312-120106.

Chen, F., Ro, D.K., Petri, J., Gershenzon, J., Bohlmann, J., Pichersky, E., and Tholl, D. (2004). Characterization of a root-specific *Arabidopsis* terpene synthase responsible for the formation of the volatile monoterpene 1,8-cineole. *Plant Physiol.* 135**,** 1956-1966. doi: 10.1104/pp.104.044388.

Coronado, C., Zuanazzi, J.A.S., Sallaud, C., Quirion, J.C., Esnault, R., Husson, H.P., Kondorosi, A., and Ratet, P. (1995). Alfalfa root flavonoid production is nitrogen regulated. *Plant Physiol.* 108**,** 533-542. doi: 10.1104/pp.108.2.533.

Croteau, R., Elbialy, H., and Dehal, S.S. (1987). Metabolism of monoterpenes - metabolic-fate of (+)-camphor in sage (*Salvia officinalis*). *Plant Physiol.* 84**,** 643-648. doi: 10.1104/pp.84.3.643.

Croteau, R., and Karp, F. (1979). Biosynthesis of monoterpenes - preliminary characterization of bornyl pyrophosphate synthetase from sage (*Salvia officinalis*) and demonstration that geranyl pyrophosphate is the preferred substrate for cyclization. *Arch. Biochem. Biophys.* 198**,** 512-522. doi: 10.1016/0003-9861(79)90526-5.

Croteau, R., and Shaskus, J. (1985). Biosynthesis of monoterpenes - demonstration of a geranyl pyrophosphate - (-)-bornyl pyrophosphate cyclase in soluble enzyme preparations from Tansy (*Tanacetum vulgare*). *Arch. Biochem. Biophys.* 236**,** 535-543. doi: 10.1016/0003-9861(85)90656-3.

Del Giudice, L., Massardo, D.R., Pontieri, P., Bertea, C.M., Mombello, D., Carata, E., Tredici, S.M., Tala, A., Mucciarelli, M., Groudeva, V.I., De Stefano, M., Vigliotta, G., Maffei, M.E., and Alifano, P. (2008). The microbial community of *Vetiver* root and its involvement into essential oil biogenesis. *Environ. Microbiol.* 10**,** 2824-2841. doi: 10.1111/j.1462-2920.2008.01703.x.

Edwards, J., Johnson, C., Santos-Medellin, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., Eisen, J.A., and Sundaresan, V. (2015). Structure, variation, and assembly of the root-

associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* 112**,** E911-920. doi: 10.1073/pnas.1414592112.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q.D., Chen, Z.H., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29**,** 644-U130. doi: 10.1038/nbt.1883.

Hardin, C.F., Fu, C.X., Hisano, H., Xiao, X.R., Shen, H., Stewart, C.N., Parrott, W., Dixon, R.A., and Wang, Z.Y. (2013). Standardization of switchgrass sample collection for cell wall and biomass trait analysis. *Bioenergy Research* 6**,** 755-762. doi: 10.1007/s12155-012-9292-1.

Horiuchi, J.I., Muroi, A., Takabayashi, J., and Nishioka, T. (2007). Exposing Arabidopsis seedlings to borneol and bornyl acetate affects root growth: Specificity due to the chemical and optical structures of the compounds. *J. Plant Inter.* 2**,** 101-104. doi: 10.1080/17429140701575624.

Hu, L.F., Robert, C.A.M., Cadot, S., Zhang, X., Ye, M., Li, B.B., Manzo, D., Chervet, N., Steinger, T., van der Heijden, M.G.A., Schlaeppi, K., and Erb, M. (2018). Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nat. Comm.* 9. doi: 10.1038/s41467-018-05122-7.

Huang, M., Sanchez-Moreiras, A.M., Abel, C., Sohrabi, R., Lee, S., Gershenzon, J., and Tholl, D. (2012). The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (*E*)-β-caryophyllene, is a defense against a bacterial pathogen. *New Phytol.* 193**,** 997-1008. doi: 10.1111/j.1469-8137.2011.04001.x.

Junker, R.R., Kuppler, J., Amo, L., Blande, J.D., Borges, R.M., van Dam, N.M., Dicke, M., Dotterl, S., Ehlers, B.K., Etl, F., Gershenzon, J., Glinwood, R., Gols, R., Groot, A.T., Heil, M., Hoffmeister, M., Holopainen, J.K., Jarau, S., John, L., Kessler, A., Knudsen, J.T., Kost, C., Larue-Kontic, A.A.C., Leonhardt, S.D., Lucas-Barbosa, D., Majetic, C.J., Menzel, F., Parachnowitsch, A.L., Pasquet, R.S., Poelman, E.H., Raguso, R.A., Ruther, J., Schiestl, F.P., Schmitt, T., Tholl, D., Unsicker, S.B., Verhulst, N., Visser, M.E., Weldegergis, B.T., and Kollner, T.G. (2018). Covariation and phenotypic integration in chemical communication displays: biosynthetic constraints and eco-evolutionary implications. *New Phytol.* 220**,** 739-749. doi: 10.1111/nph.14505.

Junker, R.R., Loewel, C., Gross, R., Dotter, S., Keller, A., and Bluthgen, N. (2011). Composition of epiphytic bacterial communities differs on petals and leaves. *Plant Biol.* 13**,** 918-924. doi: 10.1111/j.1438-8677.2011.00454.x.

Junker, R.R., and Tholl, D. (2013). Volatile organic compound mediated interactions at the plant-microbe interface. *J. Chem. Ecol.* 39**,** 810-825. doi: 10.1007/s10886-013-0325-9.

Ker, K., Seguin, P., Driscoll, B.T., Fyles, J.W., and Smith, D.L. (2012). Switchgrass establishment and seeding year production can be improved by inoculation with rhizosphere endophytes. *Biomass & Bioenergy* 47**,** 295-301. doi: 10.1016/j.biombioe.2012.09.031.

Ker, K., Seguin, P., Driscoll, B.T., Fyles, J.W., and Smith, D.L. (2014). Evidence for enhanced N availability during switchgrass establishment and seeding year production following inoculation with rhizosphere endophytes. *Arch. Agron. Soil Sci.* 60**,** 1553-1563. doi: 10.1080/03650340.2014.898840.

Kollner, T.G., Held, M., Lenk, C., Hiltpold, I., Turlings, T.C.J., Gershenzon, J., and Degenhardt, J. (2008). A maize (*E*)-beta-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most American maize varieties. *Plant Cell* 20**,** 482-494. doi: 10.1105/tpc.107.051672.

Lebeis, S.L., Paredes, S.H., Lundberg, D.S., Breakfield, N., Gehring, J., McDonald, M., Malfatti, S., Glavina del Rio, T., Jones, C.D., Tringe, S.G., and Dangl, J.L. (2015). Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* 349**,** 860-864. doi: 10.1126/science.aaa8764.

Lindow, S.E., and Brandl, M.T. (2003). Microbiology of the phyllosphere. *Appl. Environ. Microbiol.* 69**,** 1875-1883.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 25**,** 402-408. doi: 10.1006/meth.2001.1262.

Lucas-Barbosa, D., van Loon, J.J.A., and Dicke, M. (2011). The effects of herbivore-induced plant volatiles on interactions between plants and flower-visiting insects. *Phytochem.* 72**,** 1647-1654. doi: 10.1016/j.phytochem.2011.03.013.

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., del Rio, T.G., Edgar, R.C., Eickhorst, T., Ley, R.E., Hugenholtz, P., Tringe, S.G., and Dangl, J.L. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488**,** 86. doi: 10.1038/nature11237.

Neuhäuslová-Novotná, Z. (1976). H. Ellenberg Zeigerwerte der Gefässpflanzen Mitteleuropas Scripta Geobotanica, 9. *Folia geobot. phyto.* 11**,** 22-22. doi: 10.1007/bf02853313.

Penuelas, J., Asensio, D., Tholl, D., Wenke, K., Rosenkranz, M., Piechulla, B., and Schnitzler, J.P. (2014). Biogenic volatile emissions from the soil. *Plant Cell and Environ.* 37**,** 1866-1891. doi: 10.1111/pce.12340.

Prasad, B., Mah, D.J., Lewis, A.R., and Plettner, E. (2013). Water oxidation by a cytochrome P450. *PLoS One* 8. doi: 10.1371/journal.pone.0061897.

Rasmann, S., Kollner, T.G., Degenhardt, J., Hiltpold, I., Toepfer, S., Kuhlmann, U., Gershenzon, J., and Turlings, T.C.J. (2005). Recruitment of entomopathogenic nematodes by insect-damaged maize roots. *Nature* 434**,** 732-737. doi: 10.1038/nature03451.

Singh, H.P., Batish, D.R., Kaur, S., Ramezani, H., and Kohli, R.K. (2002). Comparative phytotoxicity of four monoterpenes against *Cassia occidentalis*. *Ann. Appl. Biol.* 141**,** 111-116. doi: 10.1111/j.1744-7348.2002.tb00202.x.

Sohrabi, R., Huh, J.H., Badieyan, S., Rakotondraibe, L.H., Kliebenstein, D.J., Sobrado, P., and Tholl, D. (2015). In planta variation of volatile biosynthesis: an alternative biosynthetic route to the formation of the pathogen-induced volatile homoterpene DMNT via triterpene degradation in *Arabidopsis* roots. *Plant Cell* 27**,** 874-890. doi: 10.1105/tpc.114.132209.

Steeghs, M., Bais, H.P., de Gouw, J., Goldan, P., Kuster, W., Northway, M., Fall, R., and Vivanco, J.M. (2004). Proton-transfer-reaction mass spectrometry as a new tool for real time analysis of root-secreted volatile organic compounds in *Arabidopsis*. *Plant Physiol.* 135**,** 47-58. doi: 10.1104/pp.104.038703.

Tabanca, N., Kirimer, N., Demirci, B., Demirci, F., and Baser, K.H.C. (2001). Composition and antimicrobial activity of the essential oils of *Micromeria cristata subsp phrygia* and the enantiomeric distribution of borneol. *J. Agric. Food Chem.* 49**,** 4300-4303. doi: 10.1021/jf0105034.

Tian, N., Liu, S., Li, J., Xu, W., Yuan, L., Huang, J., and Liu, Z. (2014). Metabolic analysis of the increased adventitious rooting mutant of *Artemisia annua* reveals a role for the plant monoterpene borneol in adventitious root formation. *Physiol. Plant.* 151**,** 522-532. doi: 10.1111/ppl.12139.

Tripathi, A.K., Prajapati, V., Aggarwal, K.K., and Kumar, S. (2001). Toxicity, feeding deterrence, and effect of activity of 1,8-cineole from *Artemisia annua* on progeny production of *Tribolium castanaeum* (Coleoptera : Tenebrionidae). *J. Econ. Entomol.* 94**,** 979-983. doi: 10.1603/0022-0493-94.4.979.

Tsang, H.L., Huang, J.L., Lin, Y.H., Huang, K.F., Lu, P.L., Lin, G.H., Khine, A.A., Hu, A., and Chen, H.P. (2016). Borneol dehydrogenase from *Pseudomonas sp* strain TCU-HL1 catalyzes the oxidation of (+)-borneol and its isomers to camphor. *Appl. Environ. Microbiol.* 82**,** 6378-6385. doi: 10.1128/Aem.01789-16.

Ushio, M., Yamasaki, E., Takasu, H., Nagano, A.J., Fujinaga, S., Honjo, M.N., Ikemoto, M., Sakai, S., and Kudoh, H. (2015). Microbial communities on flower surfaces act as signatures of pollinator visitation. *Sci Rep* 5**,** 8695. doi: 10.1038/srep08695.

Vaughan, M.M., Tholl, D., and Tokuhisa, J.G. (2011). An aeroponic culture system for the study of root herbivory on *Arabidopsis thaliana*. *Plant Methods* 7. doi: 10.1186/1746-4811-7-5.

Vaughan, M.M., Wang, Q., Webster, F.X., Kiemle, D., Hong, Y.J., Tantillo, D.J., Coates, R.M., Wray, A.T., Askew, W., O'Donnell, C., Tokuhisa, J.G., and Tholl, D. (2013). Formation of the unusual semivolatile diterpene rhizathalene by the *Arabidopsis* class I terpene synthase TPS08 in the root stele is involved in defense against belowground herbivory. *Plant Cell* 25**,** 1108-1125. doi: 10.1105/tpc.112.100057.

Vicente, M.R.S., and Plasencia, J. (2011). Salicylic acid beyond defence: its role in plant growth and development. *J. Exp. Bot.* 62**,** 3321-3338. doi: 10.1093/jxb/err031.

Wang, B.X., Seiler, J.R., and Mei, C.S. (2015). *Burkholderia phytofirmans* strain PsJN advanced development and altered leaf level physiology of switchgrass. *Biomass & Bioenergy* 83**,** 493-500. doi: 10.1016/j.biombioe.2015.10.029.

Wise, M.L., Savage, T.J., Katahira, E., and Croteau, R. (1998). Monoterpene synthases from common sage (*Salvia officinalis*) - cDNA isolation, characterization, and functional expression of (+)-sabinene synthase, 1,8-cineole synthase, and (+)-bornyl diphosphate synthase. *J. Biol. Chem.* 273**,** 14891-14899. doi: 10.1074/jbc.273.24.14891.

Xi, Y.J., Fu, C.X., Ge, Y.X., Nandakumar, R., Hisano, H., Bouton, J., and Wang, Z.Y. (2009). *Agrobacterium*-mediated transformation of switchgrass and inheritance of the transgenes. *Bioenergy Research* 2**,** 275-283. doi: 10.1007/s12155-009-9049-7.

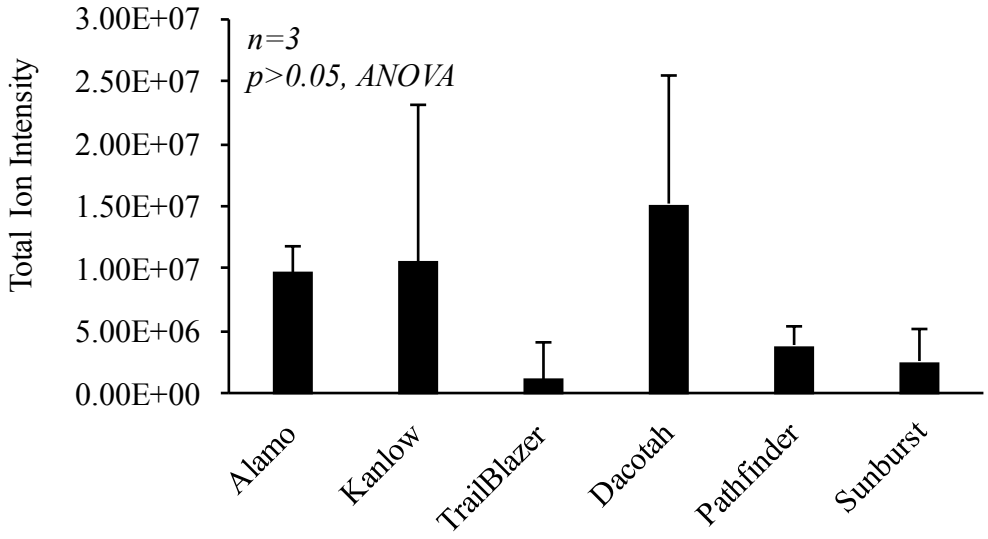
Yang, J., Kloepper, J.W., and Ryu, C.M. (2009). Rhizosphere bacteria help plants tolerate abiotic stress. *Trends Plant Sci.* 14**,** 1-4. doi: 10.1016/j.tplants.2008.10.004.

Zaks, A., Davidovich-Rikanati, R., Bar, E., Inbar, M., and Lewinsohn, E. (2008). Biosynthesis of linalyl acetate and other terpenes in lemon mint (*Mentha aquatica var. citrata*, Lamiaceae) glandular trichomes. *Isr. J. Plant Sci.* 56**,** 233-244. doi: 10.1560/Ijps.56.3.233.

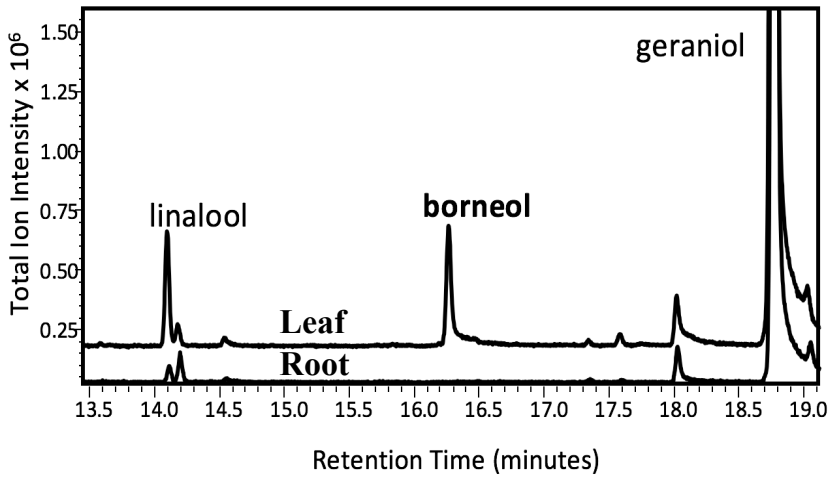
Zhang, Y.X., Ruyter-Spira, C., and Bouwmeester, H.J. (2015). Engineering the plant rhizosphere. *Curr. Opin. Biotechnol.* 32**,** 136-142. doi: 10.1016/j.copbio.2014.12.006.

Zhuang, X.F., Kollner, T.G., Zhao, N., Li, G.L., Jiang, Y.F., Zhu, L.C., Ma, J.X., Degenhardt, J., and Chen, F. (2012). Dynamic evolution of herbivore-induced sesquiterpene biosynthesis in sorghum and related grass crops. *Plant J.* 69**,** 70-80. doi: 10.1111/j.1365-313X.2011.04771.x.

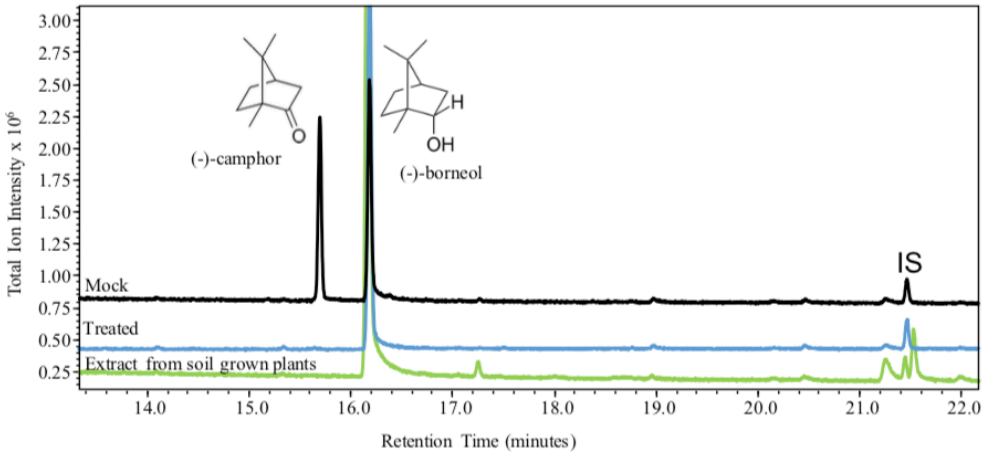
SUPPLEMENTARY FIGURES AND TABLES



Supplemental Figure S1. Quantitative analysis of borneol levels from roots of six switchgrass cultivars.



Supplemental Figure S2. Qualitative comparison of volatile products from in vitro assays with crude protein extracts from leaves and roots of switchgrass. GPP was provided as a substrate.



Supplemental Figure S3. Qualitative comparison of volatile products from switchgrass plants treated with rhizosphere microbes (Treated) or a mock control (PBS).

Supplemental Table S1. Indicator species identified from the switchgrass rhizosphere and associated bulk soil. Calculated indicator value is reported under the Stat column. Significance was determined by PERMANOVA where alpha = 0.05.

Putative Taxonomic Rank	Stat	p-value	Zone
Chloroflexi_Chloroflexia_Kallotenuales_AKIW781_unculturedbacterium	0.894	0.018*	Bulk
Actinobacteria_Acidimicrobiia_Acidimicrobiales_uncultured_unculturedactinobacterium	0.828	0.011*	Bulk
Actinobacteria_Actinobacteria_Frankiales_Sporichthyaceae_uncultured	0.828	0.007**	Bulk
Actinobacteria_Actinobacteria_Streptosporangiales_Thermomonosporaceae_Spirillospora	0.824	0.027*	Bulk
Chlamydiae_Chlamydiae_Chlamydiales_cvE6_unculturedbacterium	1	0.001	Rhizosphere
Proteobacteria_Deltaproteobacteria_Myxococcales_Vulgatibacteraceae_Vulgatibacter	0.894	0.009	Rhizosphere
Verrucomicrobia_Spartobacteria_Chthoniobacterales_DA101soilgroup_unculturedSpartobacteriabacterium	0.845	0.03	Rhizosphere
Verrucomicrobia_OPB35soilgroup_UnknownOrder_UnknownFamily_Pedosphaera	0.816	0.028	Rhizosphere

Supplemental Table S2. Summary of bacterial isolates enriched from the switchgrass endosphere on camphor or borneol.

Isolate	Carbon Source
Rhizobium sp.	Borneol
Pseudomonas sp.	Borneol
Chryseobacterium sp.	Borneol
Brevundimonas sp.	Camphor
Microbacterium sp.	Camphor
Sinorhizobium sp.	Camphor
Sphingopyxis sp.	Camphor

Supplemental Table S3. Summary of root borneol levels in wild-type and transgenic switchgrass plants. Total borneol content was calculated by standard curve analysis with authentic borneol and normalized to grams fresh weight. TL = Transgenic Line.

Sample	Borneol Peak Area	Volume	Total Area	ng Borneol	gFW	ng/gFW
WTA-1	8278010	40	331120400	291	0.53	549
WTA-2	12862759	40	514510360	453	0.54	839
WTA-3	8816578	40	352663120	310	0.57	544
WTA-4	9634253	40	385370120	339	0.49	692
WTA-5	5326642	40	213065680	187	0.51	366
WTB-1	32680074	200	6536014800	5768	0.53	10883
WTB-2	25016992	40	1000679680	882	0.52	1696
WTB-3	49932130	40	1997285200	1762	0.57	3091
WTB-4	49976378	200	9995275600	8822	0.53	16644
WTB-5	49403865	200	9880773000	8720	0.5	17441
TL2-1	3903805	40	156152200	137	0.49	279
TL2-2	8228067	40	329122680	289	0.54	536
TL2-3	266365	40	10654600	8	0.59	14
TL2-4	1243017	40	49720680	43	0.51	83
TL2-5	2746630	40	109865200	96	0.53	181
TL4-1	13848831	40	553953240	488	0.51	956
TL4-2	6852341	40	274093640	241	0.5	481
TL4-3	2053812	40	82152480	71	0.52	137
TL4-4	7313410	40	292536400	257	0.47	547
TL4-5	1615636	40	64625440	56	0.52	107
TL6-1	11936284	40	477451360	420	0.5	840
TL6-2	2067763	40	82710520	72	0.52	138
TL6-3	3037512	40	121500480	106	0.53	200
TL6-4	18019373	40	720774920	635	0.52	1221
TL6-5	3537504	40	141500160	124	0.55	225

CHAPTER IV. Characterization of Terpene Synthases Involved in the Production of Major Aroma and Flavor Compounds in Carrot Cultivars

The main objective of this study was:

I. Identification and biochemical characterization of *TPS* genes in carrot

The work for this objective has been prepared for submission to *Scientific Reports.*

Contributions:

Andrew Muchlinski: Transcriptome assembly, RNA-seq analysis, gene identification/annotation, RNA extraction and (RT)-PCR, RT-PCR, qRT-PCR, primer design, restriction based cloning, recombinant protein expression, nickel affinity chromatography, terpene synthase in vitro assays (SPME and solvent extraction), analysis and annotation of enzyme products by GC-MS, hexane extractions of colored carrots and carrot tissues and analysis by GC-MS and GC-FID, random forest analysis of colored carrot volatiles, phylogenetic tree reconstruction, amino acid alignments, analysis of diterpene formation via co-expression and GC-MS, hierarchical cluster analysis of *TPS* expression, manuscript writing and figure preparation.

Other contributions:

Mwafaq Ibdah, researcher at Newe Ya'ar Research Center: performed preliminary analysis of carrot *TPSs* and preliminary analysis of carrot volatiles; Shelby Ellison, researcher at the University of Madison-Wisconsin: performed preliminary RNA extractions and provided carrot tissues; Mossab Yahyaa, researcher at Newe Ya'ar Research Center: performed preliminary analysis of carrot *TPSs* and preliminary analysis of carrot volatiles; Bhagwat Nawade, researcher at Newe Ya'ar Research Center: performed preliminary analysis of carrot *TPSs* and preliminary analysis of carrot volatiles; Douglas Senalik, researcher at the USDA-ARS, University of

Madison-Wisconsin: performed in silico analysis of *TPS* gene expression across colored carrots and preliminary gene annotation; Phillip Simon, research leader at the USDA-ARS, University of Madison-Wisconsin: provided carrot tissues; Dorothea Tholl, Professor in the Dept. of Biology at Virginia Tech: assisted with experimental design, manuscript preparation and review.

ABSTRACT

Carrot (*Daucus carota L.*) plants produce a terpene-rich essential oil that directly contributes to aroma and flavor. Volatile constituents of carrot essential oils consist predominantly of monoterpenes and sesquiterpenes residing in highly interconnected phloem oil ducts found in above- and belowground tissue. To determine central enzymes which contribute to the terpene component of carrot volatile blends, we first analyzed tissue specific expression patterns of carrot terpene synthase genes (*TPS*) in the genomic model carrot (cv. DH1) and in roots of four aromatically unique colored carrot genotypes (orange-4943B, red-R6637, yellow-Y9244A and purple-P7262). We selected nineteen key biosynthetic enzymes involved in terpene formation and compared in vitro products from recombinant proteins with native volatile profiles obtained from DH1 and colored carrot genotypes. In addition to the (*E*)-β-caryophyllene synthase previously reported (*Dc*TPS01), we biochemically characterized several highly expressed *TPS* genes with direct correlations to major compounds of carrot flavor and aroma including germacrene-D (*Dc*TPS11), γ-terpinene (*Dc*TPS30) and α-terpinolene (*Dc*TPS03). Random forest analysis of colored carrot volatiles revealed that nine terpene compounds are sufficient for distinguishing the flavor and aroma of raw colored carrots. Interestingly, accumulation of specific terpene compounds rather than chemical diversity is responsible for differences in sensory quality traits in colored genotypes. As accumulations of specific terpene compounds can contribute to the

undesired flavor in carrot, our report provides a detailed roadmap for future breeding efforts to enhance carrot flavor and aroma.

Significance Statement (75 words max)

Carrot is a widely consumed root vegetable crop known for its high nutritional quality and characteristic aroma and flavor. The biosynthetic origins of carrot sensory quality are largely unknown. Here we report the identification and characterization of terpene synthase genes that are involved with the production of major aroma and flavor compounds of carrots. Results from this work can be directly applied for breeding and improving sensory quality of carrot.

INTRODUCTION

Aromatic plants of the Apiaceae family have been utilized for their flavor, nutritional quality, and essential oil properties since antiquity. The Apiaceae family comprises over 3,700 species across 434 genera and includes common leaf and seed crops such as celery, dill, cilantro and parsley, as well as storage root crops e.g. parsnip and carrot. The common carrot (*Daucus carota L.*) has been widely cultivated since 900 AD originating from a feral weed ("Queen Anne's lace") in Afghanistan and Central Asia [1]. Although the first use of carrots was likely for seed [1], high levels of primary and secondary metabolites in the storage root, including β-carotene the precursor to vitamin A, has led to decades of domestication and breeding for improved growth traits. With annual global production reaching 37 million metric tons (www.fao.org), carrot is nutritionally and economically one of the most important crops worldwide.

Although conventional breeding practices have focused primarily on carrot morphology, disease and pest resistance, and β-carotene content, increased attention has been focused on carrot aroma and flavor. Palatability traits can be highly genotype specific mainly attributed to the genetic background, complex blends of volatile organic compounds, and total sugar content [2-4].

Carrots produce a primary and secondary metabolite rich essential oil that directly contributes to carrot aroma and flavor. Volatile constituents of carrot essential oils consist predominantly of 10-carbon monoterpenes and 15-carbon sesquiterpenes residing in highly interconnected phloem oil ducts found in above- and belowground tissues [5,6]. Studies of flavor and aroma have been reported for orange and novel-colored carrots. For example, orange cultivars had the highest intensities of spicy and woody aroma, whereas yellow and purple cultivars have a considerable sweeter odor [7]. By contrast red cultivars have a strong bitter odor and harsh flavor attributed to overall low levels of sugars and high levels of sesquiterpenes [6]. Accumulations of specific terpene compounds, e.g. γ-terpenine, impart a harsh turpentine-like taste in carrot roots reducing overall palatability [2]. In addition, sensory analysis of carrots exposed to elevated temperature conditions (>18^{o}C) showed significant increases in harsh and bitter taste which was directly correlated with increases in terpene levels presumably masking perceptions of sweet taste [10]. As the biochemical basis of terpene related taste and aroma in carrot is not well described, understanding the formation carrot terpenes could provide a valuable resource for breeding (conventional or accelerated) more desirable carrots.

Plant essential oils are predominately composed of low-molecular weight volatile monoterpenes and sesquiterpenes [11]. With diverse roles in plant primary and secondary (or specialized) metabolism, terpenes mediate plant-pollinator interactions, defense against pathogens and herbivores, and protection against environmental stress [12-14]. Biosynthesis of the central precursors of terpenoids in plants occurs along two independent pathways resulting in the formation of the 5-carbon isoprenoid precursors, isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). The methylerythritol phosphate (MEP) pathway generates IPP and DMAPP exclusively in plastids as precursors in monoterpene and 20-carbon

diterpene biosynthesis [15] The pathway is present in all photosynthetic eukaryotes and is absent in animals, fungi and archaea [16]. The mevalonic acid (MVA) pathway occurs independently of the MEP pathway and is present in most eukaryotes including plants [16]. Unlike the MEP pathway, the six enzymatic steps of the MVA pathway occur in the endoplasmic reticulum and/or the cytosol to produce IPP as the precursor of sesquiterpenes and 30-carbon triterpene compounds [16].

Head to tail (1'-4') ionization-condensation-elimination reactions between IPP and DMAPP lead to the formation of geranyl diphosphate (GPP, C_{10}), the precursor substrate for monoterpene biosynthesis [16]. Further stereospecific condensations can occur with the addition of IPP resulting in farnesyl diphosphate (FPP, C_{15}) and geranylgeranyl diphosphate (GGPP, C_{20}) as the precursors in sesquiterpene and diterpene formation respectively [16]. Terpene synthase (TPS) enzymes form carbocations from their prenyl diphosphate substrates, which undergo numerous reactions in the catalytic center of the TPS e.g. cyclizations, hydride shifts, rearrangements and deprotonations. These reactions may lead to the formation of several linear or cyclic products by a single enzyme. Monoterpene biosynthesis occurs via dephosphorylation and ionization of GPP resulting in a geranyl carbocation, which is the precursor of cyclic and acyclic C_{10} terpenoids. The formation of sesquiterpenes begins with the ionization of FPP generating a farnesyl cation that can undergo isomerization to a nerolidyl cation. The latter cation is the primary means for cyclic terpene formation whereas the farnesyl cation is generally responsible for acyclic C_{15} terpenes. Biosynthesis of C_{20} diterpenes occurs through two mechanisms, class I enzymes form carbocations from GGPP by ionization of the diphosphate (class I) resulting in a geranylgeranyl cation that can undergo additional reactions to form various diterpene products. Substrate protonation catalyzed by class II diTPSs results in the cyclization of GGPP to ent-CPP that can then be further converted to diterpenoid products through class I TPS activity. Products produced by TPS enzymes can be

further modified by secondary reactions (e.g. CYP450) such as hydroxylations, methylations and dehydrogenations, which contribute to the structural diversity of terpene compounds. Moreover, gene duplication often leads to the evolution of large species-specific families of TPSs and is a common mechanism to generate biosynthetic complexity and structural diversity in terpene metabolism.

TPS genes can be organized into seven sub-families based on sequence homology from *TPSs* identified from multiple sequenced plant genomes [17]. The TPS-c subfamily represents the most conserved clade among land plants containing primarily diterpene synthases (di-TPSs). In contrast, the gymnosperm specific clade (TPS-d) is represented by only monoterpene and sesquiterpene synthases (mono-TPSs, sesqui-TPSs) with similar results found in the angiosperm specific clade TPS-g producing linear monoterpene and sesquiterpene products. TPS-a and TPS-b subfamilies contain primarily cyclic mono-TPSs and sesqui-TPSs identified by their lack of the N-terminal γ domain characteristic of di-TPSs found in subfamilies c, e and f. Investigations within gene families support previous evidence for tissue-specific expression and protein localization based on promoter-reporter and transit peptides analysis, respectively. Further, the presence of a conserved aspartate motif DDxxD motif, necessary for metal cofactor chelation and required for enzyme activity, is found among all functional *TPS*s [17]. Although high sequence similarly is found among many *TPS*s, sustained mutation over time and the promiscuous nature of TPSs (i.e. many products from the same substrate) may be sufficient to generate the high chemical diversity of terpene products [18].

To date, knowledge of the genetic determinants of carrot flavor and aroma has been limited [19,20]. Biochemical characterization has been performed for only two *TPS* genes, *DcTPS01 and DcTPS02* [19], leaving a majority of the biosynthetic genes responsible for the complex blend of

carrot terpene volatiles uncharacterized. Previous genomic and transcriptomic analysis of the orange Nantes type carrot (cv. DH1) estimated the size of the *TPS* gene family to contain 36 potentially functional enzymes [21]. However, recent analysis of the carrot *TPS* gene family, predicted 65 full-length *TPS*s [20]. In conjunction with this study, several QTLs containing *TPS* loci were predicted to correlate with distinct terpene compounds. To determine the major enzymes contributing to carrot aroma and flavor, we performed biochemical characterizations of carrot *TPS* based on their expression profiles in different tissues of cv. DH1 (leaves, petioles and roots) and roots of field grown colored carrot varieties (Red, Orange, Yellow and Purple) and correlated the TPS products with volatile compounds across these tissues. As terpene content strongly influences carrot flavor and aroma [4], results from this study can be applied to enhance carrot palatability and overall carrot quality.

RESULTS

Analysis of Terpene Volatiles in Leaves, Petioles and Roots

Volatile terpene constituents of the genomic model carrot cv. DH1, were analyzed from hexane extracts using GC-MS and quantified by GC-FID. We found that major carrot tissues (leaves, petioles and roots) contain a diverse blend of terpene compounds including 19 major monoterpenes and sesquiterpenes (Fig. 1). Leaf tissues contain high levels of the monoterpenes α-pinene, β-myrcene and (*E*)-β-ocimene, and the sesquiterpenes δ-elemene, (*E*)-β-caryophyllene and germacrene-D (Fig. 1; Supplemental Table S1). Comparable profiles were obtained from petioles with the exception of lower levels of β-myrcene and germacrene-D (Fig. 1; Supplemental Table S1). Root tissues showed reduced levels of α-pinene, and increased levels of γ-terpinene and α-terpinolene compared to above ground tissues (Fig. 1; Supplemental Table S1). Putative

sesquiterpene volatiles were not reported due to low levels of abundance and lack of authentic standards or oils for verification.

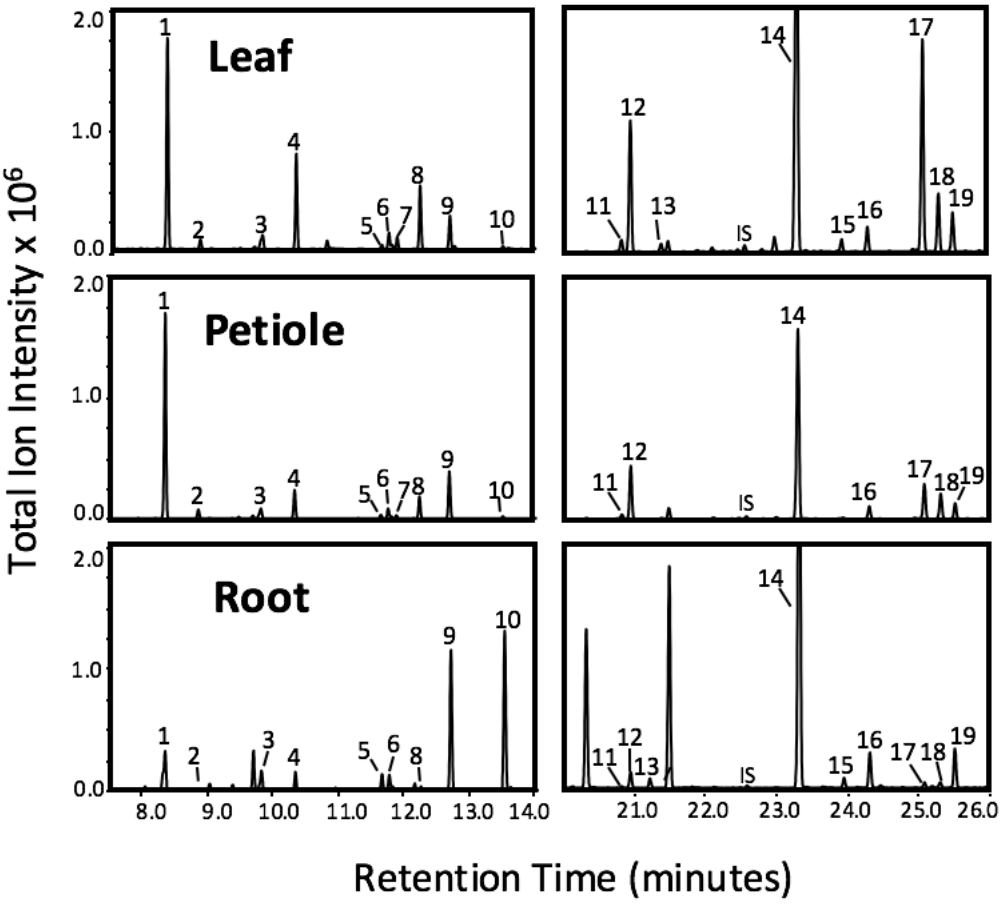


Figure 1. GC-MS analysis of hexane extracts from 11-week old carrot (cv. DH1) leaves, petioles and roots. 1: α-pinene, 2: camphene, 3: β-pinene, 4: β-myrcene, 5: cymene, 6: limonene, 7: *Z*-β-ocimene, 8: *E*-β-ocimene, 9: γ-terpinene, 10: α-terpinolene, 11: bicyclogermacrene, 12: Δ-elemene, 13: longipinene, 14: *E*-β-caryophyllene, 15: β-farnesene, 16: α-humulene, 17: germacrene-D, 18: α-farnesene, 19: β-bisabolene, IS: internal standard 1-bromodecane, Compound identification is based on comparisons to authentic standards and essential oils.

Identification of TPS Gene Models in the Carrot Genome

The carrot reference genome (Phytozome *v12*, *Daucus carota v2.0*, cv. DH1), and publically available RNA-seq data sets (SRA SAMN03216637, cv. DH1) were queried for *TPS*

genes using NCBI TBLASTX (see "Materials and Methods"). We identified 52 putative *TPS* gene models including the 36 *TPS* genes previously predicted from the DH1 cultivar by Iorizzo, et al.[21]. Although Iorizzo, et al.[21] previously generated a TPS nomenclature based on chromosomal positioning, we adopted the most recent TPS naming system for *D. carota* proposed by Keilwagen, et al.[20]. Comparisons of the 52 *TPS* gene models against the reference genome revealed 43 unique full-length open reading frames (Table 1; *DcTPS01, DcTPS02, DcTPS03, DcTPS04, DcTPS05, DcTPS07, DcTPS09, DcTPS10, DcTPS11, DcTPS12, DcTPS13, DcTPS14, DcTPS15, DcTPS16, DcTPS17, DcTPS19, DcTPS21, DcTPS23, DcTPS25, DcTPS26, DcTPS27, DcTPS28, DcTPS29, DcTPS30, DcTPS32, DcTPS33, DcTPS38, DcTPS42, DcTPS43, DcTPS44, DcTPS45, DcTPS46, DcTPS47, DcTPS48, DcTPS52, DcTPS53, DcTPS54, DcTPS55, DcTPS56, DcTPS57, DcTPS59, DcTPS60, DcTPS62).* Several *TPS* genes are located in biochemical gene clusters on chromosomes 1, 3, 4, 5, 7 and 8, including a dense 5 gene (*DcTPS14-18*) cluster on chromosome 4 (Table 1). Additional *TPS* gene models predicted by Keilwagen, et al.[20] in a genome-wide association study (GWAS) were not pursued further due to low transcript levels in roots, inability to amplify a full-length transcript or identity with previously annotated *TPSs* (Supplemental Figs. S1 and S2).

Table 1. Characteristics of the 43 terpene synthase genes analyzed in this study organized by genomic cluster. [1]Unique transcript identified by Keilwagen et al. (2017); [2]Yahyaa et al. (2015).

TPS	Locus ID	Genomic Location	Genomic Cluster	No. of Exons	cDNA Constructed	TPS Sub-family
*Dc*TPS01	*DC*AR_023152	Chr6:1181665..1185241	None	7	Yes[2]	a
*Dc*TPS32	*DC*AR_002080	Chr1:24861393..2486203	None	7	No	b
*Dc*TPS45	*DC*AR_002829	Chr1:33280604..33282540	1	7	No	g
*Dc*TPS46	*DC*AR_002830	Chr1:33286015..33288129		7	No	g
*Dc*TPS19	*DC*AR_002831	Chr1:33293414..33302528		7	Yes	g
*Dc*TPS47	*DC*AR_004091	Chr1:44627888..44628091	None	7	No	b
*Dc*TPS25	*DC*AR_012483	Chr3:47468861..47475243	None	15	Yes	c
*Dc*TPS52	*DC*AR_012537	Chr3:48081099..48082521	2	7	No	b
*Dc*TPS30	*DC*AR_012538	Chr3:48088855..48092222		7	Yes	b
*Dc*TPS09	*DC*AR_012965	Chr4:33893835..33896155	3	7	No	b
*Dc*TPS02	*DC*AR_012963	Chr4:33914246..33916610		7	Yes[2]	b
*Dc*TPS27	*DC*AR_013293	Chr4:31249549..31251992	4	8	Yes	b
*Dc*TPS55	*DC*AR_013294	Chr4:31244459..31247374		7	Yes	b
*Dc*TPS54	*DC*AR_013297	Chr4:31227164..31230361		7	Yes	b
*Dc*TPS04	*DC*AR_013298	Chr4:31217904..31220266		7	Yes	b
*Dc*TPS26	*DC*AR_013310	Chr4:31144998..31147390		7	Yes	b
*Dc*TPS56	*DC*AR_016843	Chr5:8253832..8257662	5	14	No	e
*Dc*TPS28	*DC*AR_016844	Chr5:8267895..8275147		13	Yes	e
*Dc*TPS14	*DC*AR_017536	Chr5:20668670..20671917	None	7	Yes	b
*Dc*TPS17	*DC*AR_018214	Chr5:27521963..27529973	None	7	No	b
*Dc*TPS57	*DC*AR_018422	Chr5:29664251..29668971	None	14	No	c
*Dc*TPS33	*DC*AR_019208	Chr5:37087498..37094271	None	7	No	b
*Dc*TPS59	*DC*AR_019490	Chr5:39497726..39502226	None	15	No	c
*Dc*TPS23	*DC*AR_024752	Chr7:18911173..18913238	6	7	Yes	g
*Dc*TPS60	*DC*AR_024753	Chr7:18917227..18919574		7	No	g
*Dc*TPS43	*DC*AR_026971	Chr8:27108437..27111829	7	7	No	-
*Dc*TPS44	*DC*AR_026972	Chr8:27097599..27100665		7	No	-
*Dc*TPS29	*DC*AR_027915	Chr8:17430080..17434674	None	12	No	f
*Dc*TPS62	*DC*AR_028138	Chr8:14626722..14629317	None	7	No	b
*Dc*TPS16	*DC*AR_032119	S3773:14141..17230	None	7	No	b
*Dc*TPS15	None[1]	Chr3:2698521..2703290	None	7	Yes	a
*Dc*TPS38	None[1]	Chr4:15499493..15500247	None	7	No	a
*Dc*TPS42	None[1]	Chr2:1678067..1678357	None	7	Yes	a
*Dc*TPS10	None[1]	Chr1:44680386..44685155	None	7	Yes	b
*Dc*TPS11	None[1]	Chr1:28341531..28346300	None	7	Yes	a
*Dc*TPS05	None[1]	Chr3:45432441..45440095	None	7	No	b
*Dc*TPS03	None[1]	Chr2:39586545..39589031	None	7	Yes	b
*Dc*TPS07	None[1]	Chr9:8999311..9003484	None	7	Yes	a
*Dc*TPS53	None[1]	Chr3:48692713..48694881	None	7	Yes	a
*Dc*TPS12	None[1]	Chr3:45451840..45455295	None	7	No	b
*Dc*TPS48	None[1]	Chr1:44677421..44686660	None	7	Yes	b
*Dc*TPS13	None[1]	Chr4:25547281..25566360	None	7	No	a
*Dc*TPS21	None[1]	Chr1:45337229..45352534	None	7	No	b

Amino acid alignment and phylogenetic analysis of the 43 TPS proteins indicated that carrot TPSs are organized in six TPS sub-families according to the classification by Chen, et al.[17] (Fig. 2; Supplemental Figs. S3, S4, S5, S6 and S7). We found that eight members cluster into the TPS-a sub-family (*Dc*TPS01, *Dc*TPS07, *Dc*TPS11, *Dc*TPS13, *Dc*TPS15, *Dc*TPS38, *Dc*TPS42 and *Dc*TPS53) including the previously characterized (*E*)-β-caryophyllene synthase *Dc*TPS01[19]. ChloroP analysis of subcellular localization indicated no putative transit peptides across the TPS-a clade, suggesting putative activity as sesqui-TPSs converting (*E,E*)-FPP in the cytosol (Supplemental Table S2). The TPS-b clade spans 22 members (*Dc*TPS02, *Dc*TPS03, *Dc*TPS04, *Dc*TPS05, *Dc*TPS09, *Dc*TPS10, *Dc*TPS12, *Dc*TPS14, *Dc*TPS16, *Dc*TPS17, *Dc*TPS21, *Dc*TPS26, *Dc*TPS27, *Dc*TPS30, *Dc*TPS32, *Dc*TPS33, *Dc*TPS47, *Dc*TPS48, *Dc*TPS52, *Dc*TPS54, *Dc*TPS55 and *Dc*TPS62), of which 12 were predicted to carry plastidial transit peptide sequences (*Dc*TPS02, *Dc*TPS03, *Dc*TPS04, *Dc*TPS09, *Dc*TPS10, *Dc*TPS27, *Dc*TPS30, *Dc*TPS33, *Dc*TPS48, *Dc*TPS52, *Dc*TPS54, *Dc*TPS55) suggesting these proteins are targeted to plastids where they convert GPP into monoterpenes (Fig. 2; Supplemental Table S2). We identified five type-g TPSs (*Dc*TPS19, *Dc*TPS23, *Dc*TPS45, *Dc*TPS46 and *DcT*PS60), of which only *Dc*TPS19 was predicted to function as a mono-TPS based on a putative plastidic transit peptide (Fig. 2 and Supplemental Table S2). The three members of the TPS-c clade (*Dc*TPS25, *Dc*TPS57, and *Dc*TPS59) were predicted to encode class II diterpene synthases based on the presence of the conserved DxDD motif required for the protonation-initiated cyclization of GGPP into bicyclic prenyl diphosphates including copalyl diphosphate [22]. The TPS-e/f subfamily contains 3 members (*Dc*TPS28, *Dc*TPS29, and *Dc*TPS56) and generally includes predicted class I diterpene synthases and mono-/sesquiterpene synthases. *Dc*TPS43 and *Dc*TPS44 did not cluster in any of the TPS subfamilies (Fig. 2).

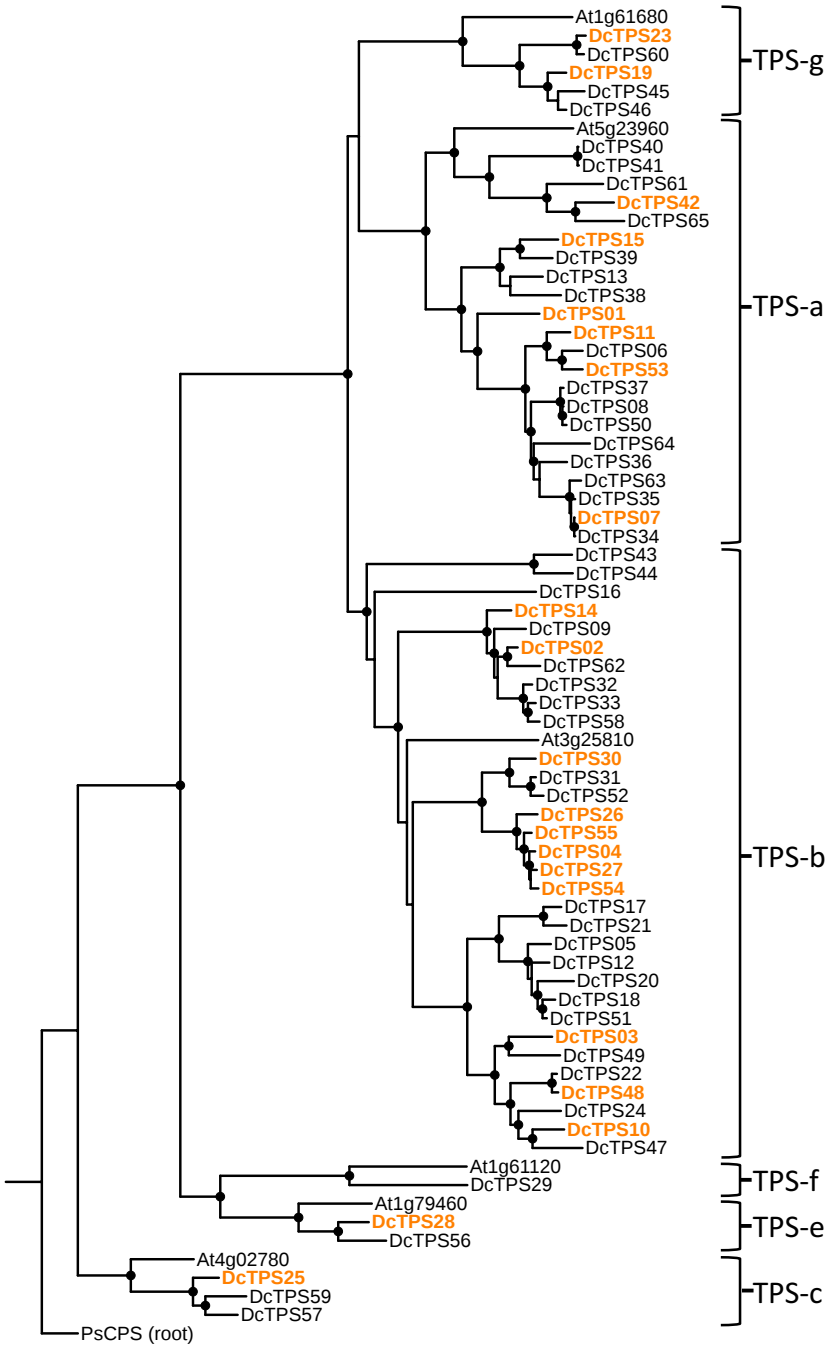


Figure 2. Maximum-likelihood phylogenetic tree of predicted TPSs in *Daucus carota (*cv. DH1). Enzymes analyzed in this study and by Yahyaa et al. 2015 are highlighted in orange. Shaded circles indicate bootstrap support of > 80% where bootstrap replicates = 500. The tree was rooted with the gymnosperm *ent*-CPP synthase from *Picea sitchensis* (PsCPS).

Gene Candidate Selection

Gene candidates for biochemical characterization were first screened by tissue specific RNA-seq analysis of DH1 and root specific RNA-seq analysis of colored carrots (Supplemental Figs. S1 and S2). *TPS* gene candidates with high in silico transcript levels were further selected based on the ability to obtain full-length transcripts and real time (RT)-qPCR amplicons across multiple tissues (Fig. 3; Supplemental Figs. S8 and S9). Full-length cDNAs or cDNAs with truncated plastidial transit peptides (19 in total) were constructed for all root-expressed *TPS* genes (*DcTPS03*, *Dc*T*PS10*, *DcTPS11*, *Dc*T*PS14*, *DcTPS15*, *Dc*T*PS25*, *DcTPS26*, *DcTPS28* and *DcTPS30*), genes with high expression in above ground tissues (*DcTPS04*, *DcTPS07*, *DcTPS19*, *DcTPS23*, *DcTPS42*, *DcTPS48*, *DcTPS53*) and any additional *TPS* genes associated with QTLs (*DcTPS27*, *DcTPS54* and *DcTPS55*) identified by Keilwagen, et al. (2017). In vitro TPS assays with the recombinant partially purified TPS proteins were performed using common TPS substrates (GPP, NPP, (*E*,*E*)-FPP, (*Z*,*Z*)-FPP and GGPP) and terpene products were analyzed by headspace SPME-GC-MS.

Characterization of TPS-a Clade Genes

In addition to *DcTPS01*, which was previously reported as an (*E*)-β-caryophyllene synthase[19], five full-length cDNAs were constructed for TPS-a type genes *DcTPS07, DcTPS11, DcTPS15, DcTPS42*, and *DcTPS53* based on expression profiling as described above. In vitro enzyme assays with recombinant partially purified *Dc*TPS42 cloned from roots indicated that the enzyme produces several putative sesquiterpene products from (*E,E*)-FPP with α-patchoulene and β-patchoulene representing the major products (Fig. 4). In addition, *Dc*TPS42 can accept GPP, NPP and (*Z*,*Z*)-FPP to produce the monoterpene products β-myrcene and β-ocimene, limonene and α-terpinolene and α- and γ-bisabolene isomers, respectively (Fig. 4; Supplemental Fig. S10).

Major and minor products from *Dc*TPS42 can be detected in carrot roots, and the multifunctionality suggests the ability to use alternative substrates in vivo depending on subcellular availability of the substrate. No diterpene synthase activity was detected when GGPP was provided as a substrate. *Dc*TPS11, a predicted cytosolic localized protein cloned from roots, was found to be highly expressed in aboveground tissues including young leaves, matures leaves and petioles (Fig. 3). The recombinant partially purified enzyme converts (*E,E*)-FPP into germacrene-D and bicyclogermacrene as the major enzymatic products in vitro (Fig. 4). Similar monoterpene products to *Dc*TPS42 were also observed with *Dc*TPS11 (limonene, α-terpinolene) as well as the formation of bisabolene isomers with (*Z,Z*)-FPP (Fig. 4; Supplemental Fig. S10). Interestingly, *Dc*TPS11 can also convert GGPP into a cembrene-like diterpene (Supplemental Fig. S11), however it may be unlikely that this reaction occurs in planta due to the limited availability of GGPP in the cytosol. *Dc*TPS07, which is highly expressed in petioles (Fig. 3), only accepts (*E,E*)-FPP as a substrate and exclusively forms germacrene-D (Fig. 4). As germacrene-D is a major component of the carrot essential oil in above- and belowground tissues, it is likely that both *Dc*TPS11 and *Dc*TPS07 contribute to the formation of this compound in vivo. *Dc*TPS53 is expressed in mature leaves and petioles and accepts all tested substrates producing δ-elemene as the major product with (*E,E*)-FPP in vitro (Figs. 3 and 4). *Dc*TPS53 converts (*Z,Z*)-FPP into γ-curcumene and β-bisabolene, and both GGP, NPP into β-myrcene and limonene (Fig. 4; Supplemental Fig. S10). In addition, *Dc*TPS53 can convert GGPP into a cubitene-like diterpene, although the limited availability of GGPP in the cytosol makes this reaction less likely in planta. All terpenes made by *Dc*TPS53 were detected in carrot tissues, suggesting that *Dc*TPS53 directly contributes to the volatile blend in vivo. *Dc*TPS15, though highly expressed in roots, had limited activity with all tested substrates (Fig. 4; Supplemental Fig. S10). Additional members of the TPS-

a clade were not tested based on previous characterization (*Dc*TPS01, [19]), low levels of constitutive expression or inability to amplify a full-length transcript (*Dc*TPS13 and *Dc*TPS38).

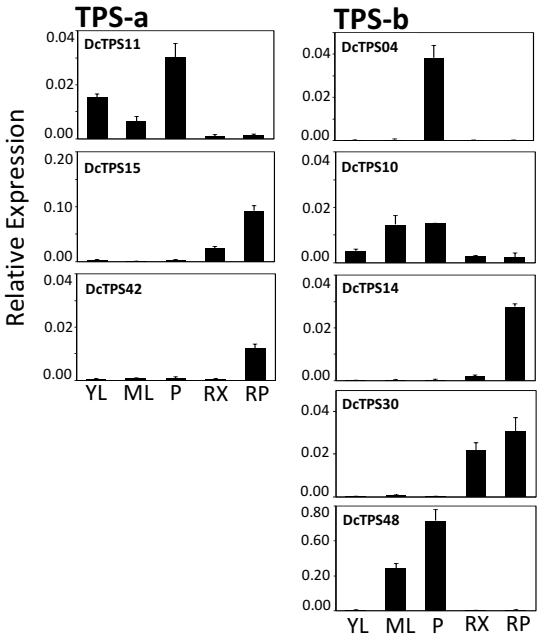


Figure 3. Analysis of gene expression by RT-qPCR for selected gene candidates. Relative expression levels across tissues for each TPS was calculated using the relative quantification method and normalized to actin. Samples were run in biological and technical triplicate and error bars indicate standard deviation from the mean. YL: young leaf, ML: mature leaf, P: petiole, RX: root xylem and RP: root phloem.

Characterization of TPS-b Clade Genes

Putative substrate preference of TPSs in the TPS-b subfamily can be generally assigned based on presence or absence of a transit peptide sequence in the mature protein. Of the 22 genes in the TPS-b subfamily, including the previously characterized *Dc*TPS02 which converts GPP into β-myrcene and geraniol in vitro [19], ten were functionally characterized (*Dc*TPS03, *Dc*TPS04, *Dc*TPS10, *Dc*TPS14, *Dc*TPS26, *Dc*TPS27, *Dc*TPS30, *Dc*TPS48, *Dc*TPS54 and *Dc*TPS55). Of these, *Dc*TPS14 and *Dc*TPS26 lack putative transit peptides suggesting the remaining 8 enzymes

are targeted to plastids (Supplemental Table S2). In vitro enzyme assays with a truncated version (43 amino acids) of *Dc*TPS30 led to the conversion of GPP (and NPP) into γ-terpinene as the major product (Fig. 4; Supplemental Fig. S10). *Dc*TPS04 and *Dc*TPS26 share ~88% sequence identity with a major difference attributed to the presence of a putative 44 amino acid transit peptide in *Dc*TPS04 (Supplemental Fig. S4). Truncated *Dc*TPS04 and full-length *Dc*TPS26 produce identical volatile profiles with sabinene and limonene from GPP (and NPP) and β-bisabolene with (*E,E*)-FPP (and (*Z,Z*)-FPP) as the major products in vitro (Fig. 4; Supplemental Fig. S10). Only *Dc*TPS26 can convert GGPP into an unidentified diterpene hydrocarbon (Supplemental Fig. S11), suggesting both enzymes might be able to use alternative substrates in vivo depending on subcellular availability. Additional full-length TPS-b type TPSs (*Dc*TPS27, *Dc*TPS54 and *Dc*TPS55) were reported by GWAS as part of a dense TPS gene cluster on chromosome 4 which includes *Dc*TPS04 and *Dc*TPS26 [20]. All five genes were significantly correlated with sabinene production in roots [20] which we confirmed by biochemical characterizations of *Dc*TPS54 and *Dc*TPS55 (Fig. 4). Although a full-length transcript was obtained for *Dc*TPS27, the presence of a ~1 kb intron after the first exon introduced a premature stop codon rendering the gene non-functional. *Dc*TPS10, expressed in leaves, petioles and roots, converts GPP into multiple monoterpene products including α-pinene, β-myrcene and limonene which can all be detected in above- and belowground tissues (Figs. 3 and 4). *Dc*TPS10 can also convert *cis*-prenyl diphosphate substrates into multiple products including the monoterpene alcohol α-terpineol (Fig. 4; Supplemental Fig. S10). However, these compounds are not readily detected in planta. *Dc*TPS03, predicted to be a root expressed mono-TPS based on transcriptome analysis, was found to be expressed in all tested tissues (Fig. 3). *Dc*TPS03 produces α-terpinolene as a major product with GPP which is a dominant component of the carrot essential oil. In addition,

*Dc*TPS03 can also produce the monoterpenes α-pinene, β-myrcene, α-phellandrene and limonene with GPP (and NPP) which likely contributes to the accumulation of these compounds in planta (Fig. 4; Supplemental Fig. S10). Limited sesquiterpene production was detected with (*E,E*)-FPP, however several putative bisabolene isomers were formed with (*Z,Z*)-FPP suggesting the potential to use this substrate in vivo (Supplemental Fig. S10). *Dc*TPS48 expression was only detected in aboveground tissues and transcripts were highly enriched in mature leaves and petioles (Fig. 3). The partially purified enzyme converted GPP (and NPP) into linalool in vitro, which can only be detected at low levels in mature leaves (Fig 4; Supplemental Fig. S10). Similar sesquiterpene products were detected as with *Dc*TPS03, and neither enzyme could convert GGPP into diterpene hydrocarbons. *Dc*TPS14, although highly expressed in root phloem tissue, did not have any activity with any tested substrates (Fig. 3; Supplemental Fig. S10). Additional members of the TPS-b clade were not tested based on previous characterization (*Dc*TPS02) [19], low levels of constitutive expression, or inability to amplify a full-length transcript (Supplemental Fig. S1; *Dc*TPS05, *Dc*TPS09, *Dc*TPS12, *Dc*TPS16, *Dc*TPS17, *Dc*TPS21, *Dc*TPS32, *Dc*TPS33, *Dc*TPS47, *Dc*TPS52, and *Dc*TPS62).

_Dc_TPS19 and _Dc_TPS23 are Members of the TPS-g Subfamily

The TPS-g subfamily in plants is closely related to the TPS-b subfamily and generally includes mono-TPSs that produce acyclic products. However, the family can also include sesqui-TPSs and di-TPSs and, therefore, enzymes in this clade are better characterized by their lack of the conserved RRX_8W motif found near the N-terminal region of monoterpene synthases in the TPS-b subfamily (Supplemental Fig. S5). *Dc*TPS19 and *Dc*TPS23 both lack the RRX_8W motif and based on sequence similarity to characterized TPSs in snapdragon flowers [23], were predicted to be produce monoterpenes (Supplemental Fig. S5). *Dc*TPS19 was found to be lowly expressed

in all tested tissues except young leaves (Fig. 3), where it likely converts GPP and (*E,E*)-FPP into linalool and nerolidol respectively, based on recombinant enzyme assays (Fig. 4). However, linalool is only detected at low levels in leaves, and nerolidol was not detected in any tested tissue. As both molecules are terpene alcohols, it is possible that these molecules are modified further e.g. glycosylation and therefore not readily observed by GC-MS. *Dc*TPS23 was found to be expressed in all tissues with the highest expression found in petioles and roots (Fig. 3). Overall enzymatic activity was limited with all substrates, however trace amounts of linalool could be detected with GPP (Fig. 4; Supplemental Fig. S10). The remaining genes in the TPS-g subfamily (*DcTPS45*, *DcTPS46*, and *DcTPS60*) were not characterized based on low levels of expression in roots or inability to amplify a full-length transcript.

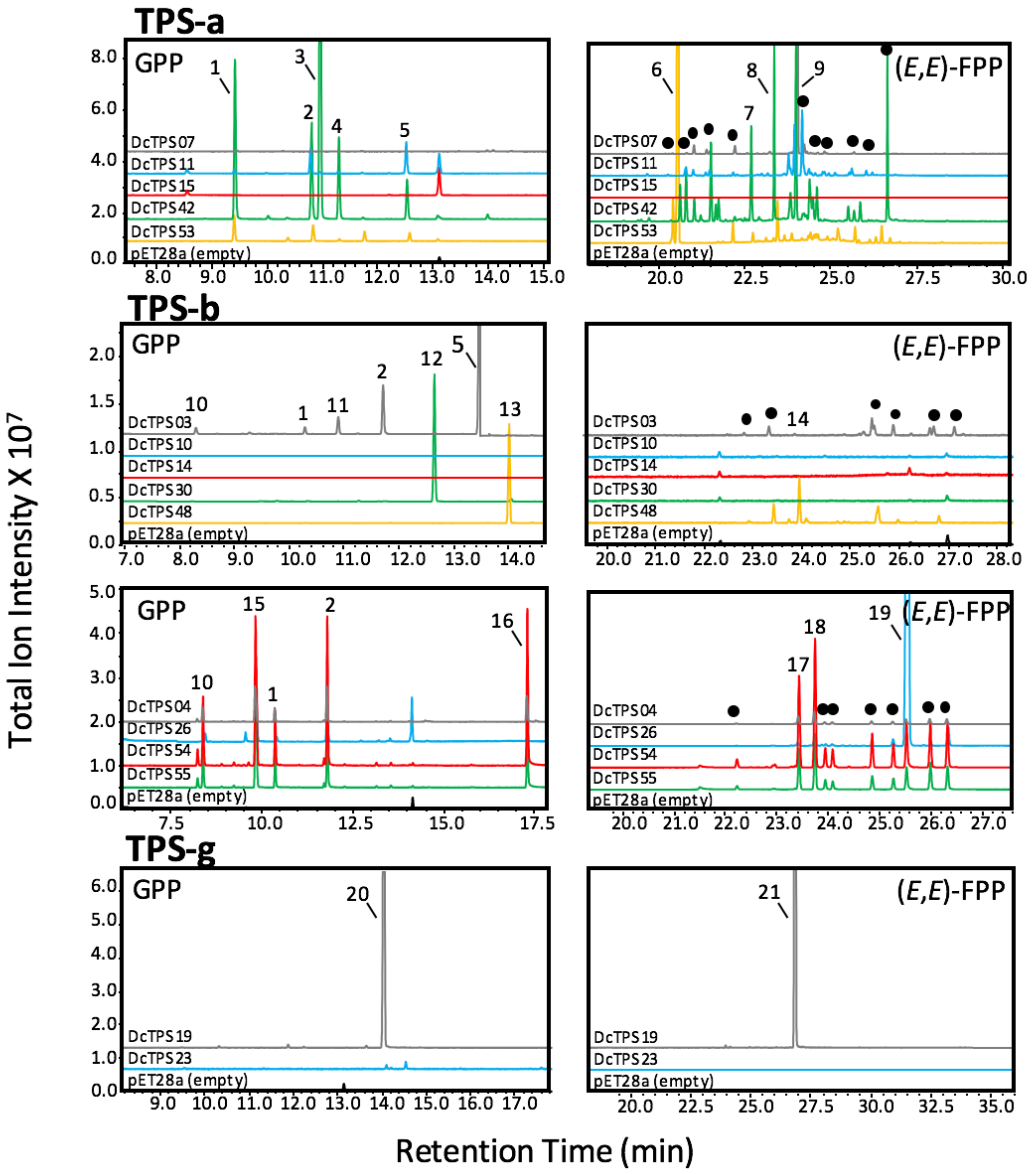


Figure 4. SPME-GC-MS analysis of the products formed in vitro by partially purified TPS enzymes. Assays were incubated for 5 minutes in the presence of a SPME fiber prior to thermal desorption as described in "Material and Methods". 1: β-myrcene, 2: limonene, 3: *Z*-β-ocimene, 4: *E*-β-ocimene, 5: α-terpinolene, 6: Δ-elemene, 7: α-patchoulene, 8: β-patchoulene, 9: germacrene-D, 10: α-pinene, 11: α-phellandrene, 12: γ-terpinene, 13: linalool, 14: β-farnesene, 15: sabinene, 16: α-terpineol, 17: bergamotene, 18: sesquisabinene, 19: β-bisabolene, 20: linalool, 21: nerolidol. Compound identification is based on comparisons to authentic standards and essential oils provided by Dr. Tobias Köllner (MPI-CE). GPP: geranyl diphosphate, *E,E*-FPP: all-*trans*-farnesyl diphosphate.

*Dc*TPS25 Belongs to the TPS-c Clade

The plant TPS-c subfamily contains enzymes with an N-terminal γ-domain characteristic of diterpene synthases involved in primary and secondary metabolism. In carrot, we identified three TPSs that belong to the TPS-c subfamily. Of these, we generated a full-length cDNA for the constitutively expressed *Dc*TPS25 based on high expression in roots compared to the other members of the TPS-c subfamily *Dc*TPS57 and *Dc*TPS59 (Fig. 3; Supplemental Figs. S8 and S9). We found that *Dc*TPS25 is a class II diterpene cyclase converting GGPP into *ent*-copalyl diphosphate (CPP) in vitro based on mass spectral comparison of the acid hydrolyzed product *ent*-copalol (Fig. 5A), and to the acid hydrolyzed product from the known *CPS* gene from *Arabidopsis*

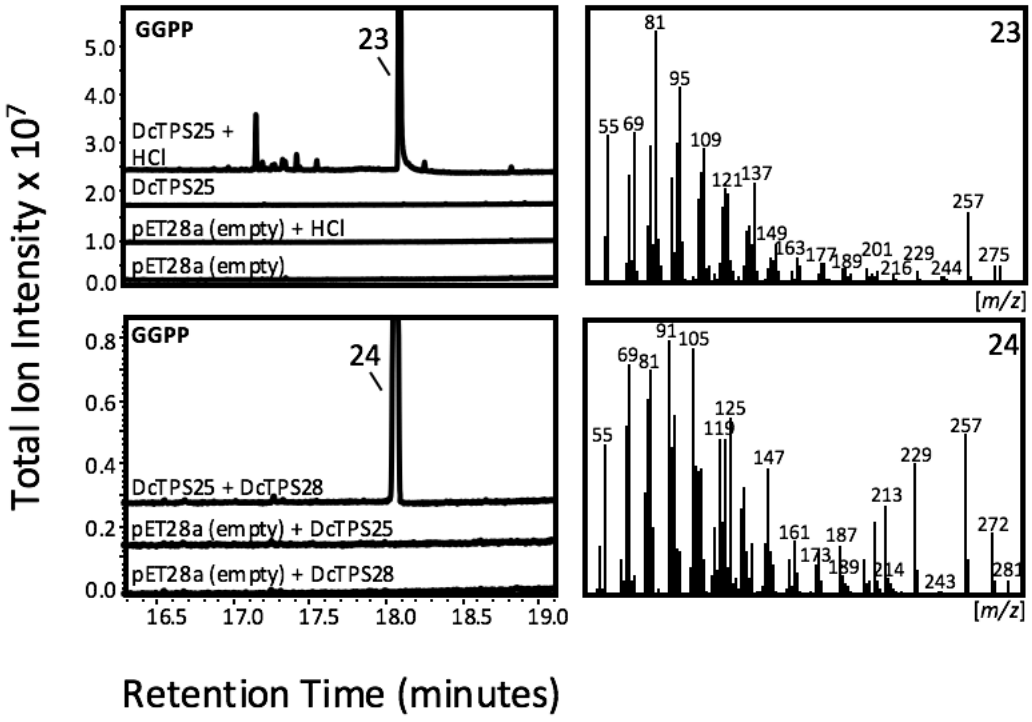


Figure 5. Functional characterization of recombinant DcTPS25 (Class II) and DcTPS28 (Class I). 23: *ent*-copalol (dephosphorylated *ent*-CPP); 24: *ent*-kaurene. Compound identification is based on library comparisons (NIST/WILEY) and comparisons to enzymatic products from known *ent*-CPP and *ent*-karuene synthases provided by Dr. Reuben Peters (Iowa State).

thaliana (kindly provided by Dr. Reuben Peters). No enzymatic activity was detected with any other substrate tested.

_Dc_TPS28 is an _ent_-kaurene Synthase in the TPS-e/f Subfamily

Of the three TPS-e/f type genes identified by RNA-seq analysis (*TPS28, TPS29* and *TPS56*), we constructed a full-length cDNA for the predicted class I diterpene synthase *DcTPS28* based on high expression in roots (Fig. 3). In vitro enzyme assays with the recombinant protein incubated with GPP, NPP, (*E,E*)-FPP, (*Z,Z*)-FPP or GGPP produced no detectable hydrocarbon product. However, when co-expressed with a pGGeC plasmid (provided by Dr. Reuben Peters, Iowa State) which carries a *GGPPS* gene from *Abies* grandis and a *CPS* gene from *Arabidopsis thaliana* [24], *Dc*TPS28 successfully converted *ent*-CPP into *ent*-kaurene (Fig. 5B). Additionally, *ent*-kaurene could be produced by co-incubating partially purified *Dc*TPS25, *Dc*TPS28 and GGPP confirming activities for both enzymes (Fig. 5B) Confirmation of *ent*-kaurene production was verified by mass spectral comparison to products from a known *ent*-kaurene synthase gene from *Bradyrhizobium japonicum* (pDEST15-BjKS courtesy of Dr. Reuben Peters).

Colored Carrot Varieties Have Distinct Volatile Profiles

Carrot varieties of different color can be distinguished by distinct sensory qualities. Do determine whether these differences correlate with differences in terpene profiles, we performed a random forest analysis of 14 major monoterpene and sesquiterpene compounds (Fig. S12). This analysis revealed a strong separation of the colored genotypes (Fig. 6). Boruta variable selection identified 9 terpene factors as important in distinguishing the colored varieties (Table. S5). We found that orange carrot roots in this study (cv. B493B) accumulated significantly higher levels of (*E*)-β-caryophyllene (ANOVA; p=2.95e-05) and α-humulene (ANOVA; p=1.03e-04) compared

to red, purple and yellow (Fig. 7). In addition, yellow carrots (cv. Y9244A), accumulated high levels of β-bisabolene (ANOVA; p=2.02e-03) and and γ-bisabolene (ANOVA; p=7.51e-03) in comparison the other tested varieties (Fig. 7). Although α-terpinolene significantly contributes to cultivar differences (Table. S5; ANOVA; p=0.046), no significant pairwise difference was detected among cultivars (Fig.7). To determine if observed cultivar specific terpene differences correlated with *TPS* gene expression, we analyzed transcript levels from RNA-seq data using the Bioconducter package Limma (Fig. S1). We found that *DcTPS01* expression patterns strongly correlate with the major volatiles produced by the enzyme (i.e. (*E*)-β-caryophyllene and α-humulene). In addition, increased α-terpinolene levels in yellow and orange carrot directly correlated with the expression of the α-terpinolene synthase *DcTPS03*. Although multiple enzymes can produce bisabolenes in carrot, we found that β-bisabolene and γ-bisabolene levels strongly correlated with the expression pattern of *DcTPS42* (Fig. S1). No additional correlations were found due to multiple enzymes involved in the formation of the compounds (e.g. α-pinene, β-pinene, β-farnesene) or unknown biochemical origin of the compound (bornyl acetate).

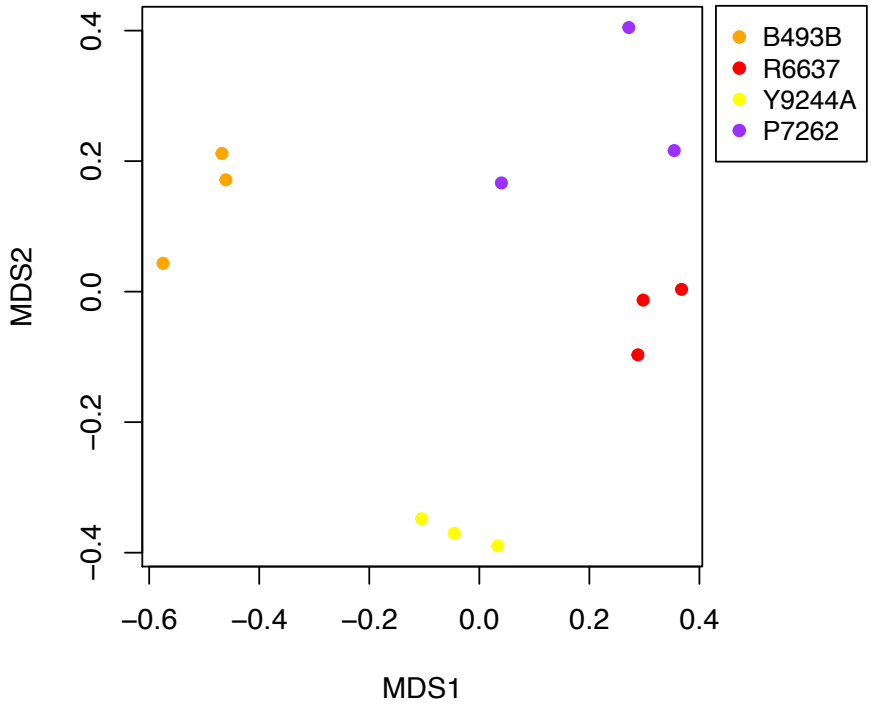


Figure 6. Multidimensional scaling plot (MDS) based on a random forest analysis of root terpene volatiles from field grown colored carrots. Yellow-Y9244A, Orange-B493B, Red-R6637 and Purple-Orange-P7262.

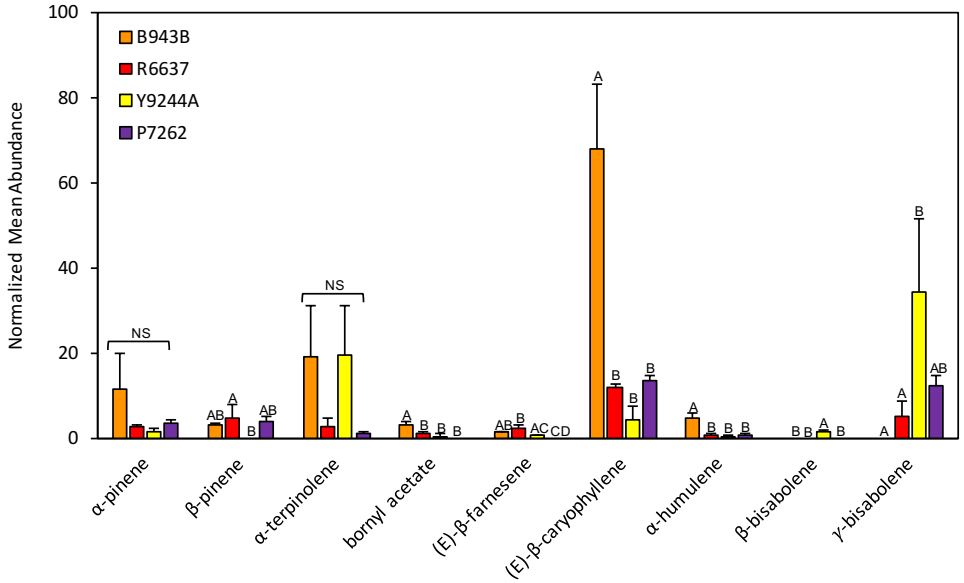


Figure 7. Comparison of relative terpene levels for compounds identified by random forest analysis. Letters indicate significant differences determined by multiple ANOVAs and pot-hoc Tukey-HSD comparisons. Yellow-Y9244A, Orange-B493B, Red-R6637 and Purple-Orange-P7262.

DISCUSSION

Carrot (*Daucus carota L.*) genotypes have been extensively studied for their commercial and nutritional value, essential oil content, and resistance against pathogens and herbivores [25], [26], [27], [28]. However, the genetic determinants of volatile terpene constituents of carrot essential oil, though first reported 50 years ago [29], had largely remained biochemically undescribed. Here we report the major terpene volatiles in the genomic model carrot (DH1), and correlate major compounds with the in vitro function of 19 recombinant TPS enzymes. In combination with the two *TPSs* previously reported (*DcTPS01* and *DcTPS02;* [19]), biochemical characterization of enzymes responsible for the production of the major carrot terpene volatiles has now been described.

Carrot roots and leaves produce a structurally diverse blend of mono- and sesquiterpene compounds that directly contribute to carrot aroma and flavor [6,26, Kreutzmann, 2008 #768]. Although carrot aroma and flavor are complex sensory traits, both attributes can be strongly influenced by specific terpene compounds [7]. We show that expression of *TPSs* can be highly tissue specific and, therefore, help shape the distinct aroma of above and belowground tissues (Fig. 3). For example, sweet, citrus and fruity aromatic notes in carrot are associated with the accumulation of the monoterpenes γ-terpinene, limonone, α-terpinolene, and the sesquiteprnes β-bisabolene and β-farnesene [7]. We found high levels of the monoterpenes γ-terpinene and α-terpinolene, which provide a sweet and fruity odor [7], in roots but not leaves or petioles of carrot (cv. DH1). Interestingly, although γ-terpinene was detected in all tested tissues (roots, petioles and leaves), biosynthesis of the compound likely occurs exclusively in roots by *Dc*TPS30 (Figs. 3 and 4). In vitro, *Dc*TPS11 can also catalyze the formation of γ-terpinene from GPP (Fig. 4); however, this reaction is unlikely to occur in planta as *Dc*TPS11 is not predicted to be plastid localized and therefore potentially restricted by substrate availability in the cytosol. As lipophilic interconnected

phloem oil ducts occur throughout roots, petioles and leaves [5], it is conceivable that terpene hydrocarbons are mobilized from roots to shoots through schizogenous spaces.

The bicyclic monoterpenes (α/β)-pinene and sabinene accumulate in carrot leaves and petioles and are responsible for the "pine" and "carrot top" aroma [6,7]. We found that the dense *TPS* gene cluster (*DcTPS*04, *DcTPS*26, *DcTPS*54 and *DcTPS*55; *DcTPS27* non-functional) on chromosome 4 was involved in formation of these compounds in vitro (Fig. 4). Based on high sequence homology and functional redundancy, it is likely that this gene cluster represents a recent gene duplication event in carrot.

Carrot roots maintain a distinctive spicy and woody aroma and flavor mainly attributed to accumulation of sesquiterpene hydrocarbons [6,7]. In addition to *Dc*TPS01, catalyzing the formation of (*E*)-β-caryophyllene and α-humulene [19], we characterized *Dc*TPS42 as a root phloem expressed multi product sesquiterpene synthase (Figs. 3 and 4). Verified major compounds produced by recombinant *Dc*TPS42 in vitro include germacrene-D and (α/β)-patchoulene, which likely contribute to the woody fragrance of carrot root [7]. Interestingly, *Dc*TPS42 also catalyzed the formation of more than 20 additional volatile sesquiterpene products, many of which can be observed across carrot tissues (Fig. 4), therefore it is likely that this enzyme is a major contributor to carrot flavor and aroma.

Raw colored carrot genotypes have unique flavors and aromas attributed to the accumulation of specific terpene compounds [6]. We found that volatile terpene profiles from four field grown colored carrot varieties (orange, red, yellow and purple) show significant differences in levels of compound accumulation, likely contributing to the variation in sensory attributes of colored carrots (Figs. 6 and 7). Correlation of *TPS* gene expression with terpene accumulation in colored carrots corresponded well to the major terpene compounds observed in carrots roots

(Supplemental Figs. S10 and S11). Results from this study provide substantial insight into *TPS* contributions to flavor and aroma in carrot and could be directly applied to breeding and improving the sensory quality of carrots.

MATERIALS AND METHODS

Plant Growth and Conditions

Seeds from the doubled haploid orange Nantes type carrot DH1 were kindly provided by Rijk Zwaan and directly seeded into 6-inch clay pots filled with 50% potting mix and 50% composted soil. Seedlings were grown at the University of Wisconsin, Madison, Walnut Street Greenhouse under a 12-h photoperiod with an average temperature cycle of 20^{o}C - 25^{o}C (night/day). Colored carrot cultivars (yellow-Y9244A, orange-B493B, red-R6637 and purple-7262) were field grown at the University of Wisconsin, Madison. Whole plants were harvested 100 days after planting and frozen immediately in liquid nitrogen for later isolation of RNA and metabolite extraction.

Identification of *TPS* Genes in the Carrot Genome

Publically available RNA-seq data for above- and belowground tissues were retrieved from the NCBI Short Read Archive [21], biosample SAMN03216637 and quality assessed using FastQC. Reads were truncated nine bp using Trimmomatic [30] to remove low quality sequences and assembled de novo using Trinity [31]. Assembled transcriptomes (20 total) were individually queried with a representative TPS sequence (*Dc*TPS01) using TBLASTX. The resulting "hits" were manually curated for putative functionality based on length and presence of aspartate rich conserved motifs (DDxxD, DxDD). Gene models were refined further by comparing transcripts to genome sequences available in Phytozome (*Daucus carota v2.0*). Exon/intron structure was predicted by alignment of coding sequences to genomic sequences using the Gene Structure

Display Server [32]. Putative N-terminal plastidic transit peptides were predicted using multiple sequence alignments and analysis of each sequence using the transit peptide prediction software ChloroP [33]. Phylogenetic analysis was conducted in Geneious (v8.0.2) using default settings (bootstrap = 1000) based on multiple sequence alignments generated with Clustal Omega [34].

Gene Expression Analysis

Qualitative Analysis of 43 Carrot TPSs by RT-PCR

Total RNA was extracted from each tissue type in biological triplicate (young leaves, fully expanded leaves, petiole, root xylem, root phloem and whole root) using the TRIzol® Plus RNA Purification Kit (Life Technologies, Carlsbad, CA) in accordance with the manufacturer's protocol. RNA was treated for DNA contamination with the TurboDNA-free kit (Life Technologies, Carlsbad, CA) and used for first strand cDNA synthesis with SuperScriptII reverse transcriptase and oligo$(dT)_{18}$ primers (Invitrogen) according to the manufacturer's instructions. PCR amplification of 43 *TPS* genes was performed with each cDNA, gene specific primers (Supplemental Table S3), and Taq DNA Polymerase (Promega) with an initial denaturing step of 95^{o}C for 5 min, followed by denaturation for 30 s at 95^{o}C, annealing for 30 s at 50^{o}C, extension for 1 min at 72^{o}C and a final extension for 7 min for 30 cycles. Actin and PP2A were used as internal controls.

Quantitative Analysis of 43 full-length Carrot TPSs by RT-qPCR

Total RNA extraction and first-strand cDNA synthesis were performed as described above, however RNA was first normalized between samples and replicates to 2.5 µg based on denaturing gel electrophoresis and spectrophotometer measurements at 260 nm. The resulting cDNA was diluted to 100 ng/µl. Reactions were performed with 1 µl cDNA in a 20 µl reaction using Power SYBR Green PCR master mix (Applied Biosystems) and gene specific primers (Supplemental

Table S3). PCR amplifications were done with a CFX96 Touch real-time PCR detection system (Bio-Rad) with the following cycles: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, 50°C 30 s and 60°C for 1 min. Melt curve analysis was performed at the end of amplification to ensure specificity of each primer pair. Relative expression levels across tissues for each *TPS* gene were calculated using the relative quantification method and normalized to actin.

Amplification of Full-Length TPS cDNAs and Plasmid Construction

Full-length cDNAs for *DcTPS07*, *DcTPS11*, *DcTPS15*, *DcTPS19*, *DcTPS23*, *DcTPS26*, *DcTPS42*, *DcTPS53* and those truncated based on predicted transit peptide coding regions *DcTPS03*, *DcTPS04*, *DcTPS10*, *DcTPS14*, *DcTPS25*, *DcTPS27*, *DcTPS28*, *DcTPS30*, *DcTPS48*, *DcTPS54* and *DcTPS55* were obtained by PCR-amplification with gene-specific primers carrying restriction sites (Supplemental Table S4). Template cDNA was derived from root and stem RNA as described above. Amplification was performed with Q5 High-Fidelity DNA polymerase in a 25 µl reaction volume with the following PCR conditions: 98°C for 30 s, followed by 30 cycles of 98°C for 30 s, 55°C for 30 s, 72°C for 1 min 45 s and a final extension at 72°C for 2 min. Amplified fragments were gel purified using a NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel, MN) and concentrated to ~10 µl. A 10 µl A-tailing reaction was prepared with 3 µl of purified PCR product incubated in the presence of 10 mM dATPs and Taq polymerase at 72°C for 30 min. The resulting product was ligated overnight into the pGEM-T Easy vector (Promega) and Sanger sequenced to verify the insert. Open reading frames were then digested with the appropriate restriction enzymes (typically *BamHI, XhoI*) and ligated overnight into the corresponding restriction sites of the bacterial expression vector pET28a (Novagen). Constructs were Sanger sequenced again prior to expression in *Escherichia coli*.

Recombinant Protein Expression in *Escherichia coli* and TPS Assays

All verified plasmids were transformed into *E. coli* BL21-CodonPLus(DE3) cells (Stratagene) and grown at 37°C in 100 ml Luria-Bertani (LB) media supplemented with 50 µM kanamycin until the optical density at 600 nm (OD_{600}) reached 0.5 to 0.7. Protein production was then induced with 0.5 mM isopropyl 1-thio-ß-D-galactopyranoside (IPTG) and incubated with shaking at 18°C for 16 h. Cell pellets were washed with 10 mM Tris base and 50 mM potassium chloride, resuspended in 4 ml phosphate buffered saline (PBS, 50 mM sodium phosphate, 100 mM sodium chloride, 10% glycerol) supplemented with 1 mM dithiothreitol (DTT) and 0.5 mM phenylmethylsulfonyl fluoride (PMSF), and lysed by sonication. Clarified extracts were mixed with equal parts PBS and recombinant His(6x)-tagged proteins were partially purified by Ni^{2+} affinity chromatography according to the manufacturer's instructions (Qiagen). Partially purified proteins were then desalted on PD-10 desalting columns (GE) equilibrated with assay buffer (10 mM MOPSO, 10% glycerol [v/v] and 1 mM DTT, pH 7.0) and visualized by SDS-PAGE (10%, GenScript). Enzyme reactions (125 µl total volume) were prepared in a 10-ml screw cap vial (Supelco) by combing partially purified protein with 20 mM $MgCl_2$ and 60 µM commercially available prenyl diphosphate substrates including GPP, NPP, (*E,E*)-FPP, (*Z,Z*)-FPP and GGPP (Echelon Biosciences). Assay mixtures were analyzed using automated solid phase microextraction (SPME, AOC-5000 Shimadzu) following a 5 min incubation at 30°C in the presence of a 100-µM polydimethylsiloxane fiber (Supelco). Collected volatiles where thermally desorbed for 4 min and analyzed using a gas chromatograph (240°C injector port) coupled with a quadrupole mass spectrometer (GC-MS-QP2010S, Shimadzu). Extracts were separated with a 5:1 split on a 30 m x 0.25 mm i.d. x 0.25 µm Zebron capillary column (Phenomenex) using Helium as the carrier gas (1.4 ml min^{-1} flow rate) and a temperature gradient of 5°C min^{-1} from 40°C (hold 2

min) to 220°C. Identification of major volatile compounds was confirmed by comparisons of retention times and mass spectra to authentic standards when available (Sigma) and mass spectral libraries (Wiley and NIST).

***Dc*TPS25**

Diterpene cyclase activity of *Dc*TPS25 was tested by incubating partially purified protein as described above with 60 µM GGPP and 10 mM $MgCl_2$ for 1 h at 30°C with the addition of a 1 ml hexane overlay. Following incubation, 80 µl of 5 M HCl or water was added and mixed by vortex to facilitate acid hydrolysis of terpene products. Hexane fractions were dried over magnesium sulfate ($MgSO_4$), concentrated to ~40 µl, and 1 µl was injected into the GC-MS as described above. Identification of major compounds was confirmed by comparisons of retention times and to authentic standards when available (Sigma) and mass spectral libraries (Wiley and NIST).

***Dc*TPS28**

Diterpene synthase activity of TPS28 was tested as described above, either alone or co-incubated with partially purified TPS25, and by co-expression of pET28a-*Dc*TPS28 with a pGGeC plasmid (provided by Dr. Reuben Peters) which carries a *GGPPS* gene from *Abies* grandis and a *CPS* gene from *Arabidopsis thaliana* [24]. Constructs, including a known *ent*-kaurene synthase gene from *Bradyrhizobium japonicum* as a control (pDEST15-BjKS courtesy of Dr. Reuben Peters), were co-transformed into *E. coli* C41 (DE3) OverExpress cells (Lucigen) and a single bacterial colony was selected to inoculate 5 ml LB media and incubated for 16 h at 37°C. The saturated culture was used to inoculate a 50 ml LB culture, which was incubated at 37°C until the OD_{600} reached 0.5 to 0.7. Protein expression was induced by the addition of 0.5 mM IPTG and incubated with shaking at 18°C for 72 h. Cultures were extracted with equal parts hexane, dried over $MgSO_4$,

concentrated to ~40 µl, and 1 µl was injected for GC-MS analysis as described above. Identification of *ent*-kaurene was achieved as described above, and by comparisons to the *ent*-kaurene product produced by the BjKS.

GC-MS and GC-FID Analysis of Terpenes

Volatile terpenes were extracted from 1 g of leaf, petiole, root phloem, root xylem and whole root samples in biological triplicate from culture conditions described above (cv. DH1). Samples were rinsed with deionized water, dried with tissue paper and immediately frozen in liquid nitrogen for processing. Samples were then ground to a fine powder for 2 min in the presence of liquid nitrogen, weighed, transferred to 5 ml hexanes and mixed by vortex for 20 s. The ground material was placed in an ultrasonic bath (Fisher Scientific) for 10 min and then pelleted by centrifugation. Following the collection of two fractions, 1-bromodecane was added for a final concentration of 20 ng/µl as an internal standard and extracts were dried over a $MgSO_4$ column and concentrated on ice to ~40 µl. Extracts were separated as above with a 2:1 split using the same column and conditions above, and by GC-FID (ThermoFinnigan) using Helium as the carrier gas (1.4 ml min^{-1} flow rate) and Nitrogen, Hydrogen and Air (25, 35, 350 ml min^{-1}, respectively) as makeup and combustion gasses. Annotation of major terpene compounds was achieved as described above. Chromatograms were compared between GC-MS and GC-FID results for compound identification and quantification using the multipoint internal standard method (Alltech). Standard curves for monoterpene and sesquiterpene compounds were constructed with authentic α-pinene and α-humulene (Sigma), respectively, and obtained values were normalized to gram fresh weight. Analysis of volatile compounds for colored carrot cultivars (B493B, R6637, Y9244A, P7262) followed identical methodology with the exception that compounds were only

analyzed in roots by GC-MS and reported as normalized relative abundance ((peak area analyte/peak area internal standard)/gram fresh weight).

Random Forest Analysis and Boruta Factor Selection

Abundance of major terpene compounds from roots of colored carrot cultivars (see above) were analyzed by random forest and Boruta using R (v3.5.0). Random forest analysis was set to 5000 bootstrap iterations for compound selection. Compounds selected in bootstrapped models were retained for use in MANOVAs comparing normalized relative abundance of compounds across cultivars. Significant differences among groups were followed with ANOVAs and Tukey-HSD post-hoc comparisons of compound levels among groups.

RNA-seq Analysis of *TPS* Gene Expression in Colored Carrots

Total RNA was extracted from 14 week old whole roots, of colored carrot cultivars (B493B, R6637, Y9244A, P7262), with three roots (i.e., three biological replicates) per sample set. Total RNA was extracted using the TRIzol® Plus RNA Purification Kit (Life Technologies, Carlsbad, CA) following the manufacturer's protocol. DNA was removed with the 'DNA free-kit' provided with the RNA purification kit. RNA quantification was measured on a Nanodrop One Spectrophotometer and quality control was done on an Agilent 2100 Bioanalyzer RNA NanoChip. For each RNA sample, libraries were prepared at the University of Wisconsin-Madison Gene Expression Center and sequenced on an Illumina HiSeq 2000 using 1x100 nt reads. After quality control with FastQC (http://www.bioinformatics.babraham.ac.uk /projects/fastqc/), reads were filtered with Trimmomatic version 0.32 (Bolger et al., 2014) with adapter trimming and using a sliding window of length ≥50 and Phred quality score ≥28. Reads were mapped against the carrot genome sequence (GenBank accession LNRQ01000000.1) using Bowtie2 (Langmead and Salzberg 2012). Illumina reads were mapped against the carrot genome sequence (GenBank

accession LNRQ01000000.1) using Rsubread version 1.24.2. (Liao *et al* 2013). Transcript expression was analyzed using the Bioconductor package limma (Ritchie *et al*. 2015).

REFERENCES

1 Simon, P. W. *Domestication, historical development, and modern breeding of carrot*. Vol. 19 (2000).

2 Simon, P. W., Peterson, C. E. & Lindsay, R. C. Correlations between sensory and objective parameters of carrot flavor. *J Agr Food Chem* **28**, 559-562, doi:DOI 10.1021/jf60229a041 (1980).

3 Alasalvar, C., Grigor, J. M., Zhang, D. L., Quantick, P. C. & Shahidi, F. Comparison of volatiles, phenolics, sugars, antioxidant vitamins, and sensory quality of different colored carrot varieties. *J Agr Food Chem* **49**, 1410-1416, doi:DOI 10.1021/jf000595h (2001).

4 Rosenfeld, H. J., Vogt, G., Aaby, K. & Olsen, E. Interaction of terpenes with sweet taste in carrots (*Daucus carota L.*). *Acta Hortic*, 377-386, doi: 10.17660/ActaHortic.2004.637.47 (2004).

5 Senalik, D. & Simon, P. W. Relationship between oil ducts and volatile terpenoid content in carrot roots. *Am J Bot* **73**, 60-63, doi:Doi 10.2307/2444277 (1986).

6 Kreutzmann, S., Thybo, A. K., Edelenbos, M. & Christensen, L. P. The role of volatile compounds on aroma and flavour perception in coloured raw carrot genotypes. *Int J Food Sci Tech* **43**, 1619-1627, doi:10.1111/j.1365-2621.2007.01662.x (2008).

7 Fukuda, T., Okazaki, K. & Shinano, T. Aroma characteristic and volatile profiling of carrot varieties and quantitative role of terpenoid compounds for carrot sensory attributes. *Journal of Food Science* **78**, S1800-S1806, doi:10.1111/1750-3841.12292 (2013).

8 *The carrot genome*. 1 edn, (Springer International Publishing, 2019).

9 Iorizzo, M. *et al.* A cluster of myb transcription factors regulates anthocyanin biosynthesis in carrot (daucus carota l.) root and petiole. *Front. Plant Sci.* **9**, doi:10.3389/fpls.2018.01927 (2019).

10 Rosenfeld, H. J., Aaby, K. & Lea, P. Influence of temperature and plant density on sensory quality and volatile terpenoids of carrot (*Daucus carota L.*) root. *J Sci Food Agr* **82**, 1384-1390, doi:10.1002/jsfa.1200 (2002).

11 Dudareva, N., Pichersky, E. & Gershenzon, J. Biochemistry of plant volatiles. *Plant Phys* **135**, 1893-1902, doi:10.1104/pp.104.049981 (2004).

12 Goyal, S., Lambert, C., Cluzet, S., Mérillon, J. M. & Ramawat, K. G. in *Prog Bio Control* 109-138 (Springer, 2012).

13 Bartwal, A., Mall, R., Lohani, P., Guru, S. K. & Arora, S. Role of secondary metabolites and brassinosteroids in plant defense against environmental stresses. *J Plant Growth Regul* **32**, 216-232, doi:10.1007/s00344-012-9272-x (2013).

14 Theis, N. & Lerdau, M. The evolution of function in plant secondary metabolites. *Int J Plant Sci* **164**, S93-S102 (2003).

15 Lichtenthaler, H. K., Rohmer, M. & Schwender, J. Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. *Physiol Plantarum* **101**, 643-652, doi:10.1111/j.1399-3054.1997.tb01049.x (1997).

16 Tholl, D. & Lee, S. Terpene specialized metabolism in arabidopsis thaliana. *The Arabidopsis Book / ASPB* **9**, e0143, doi:10.1199/tab.0143 (2011).

17 Chen, F., Tholl, D., Bohlmann, J. & Pichersky, E. 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, doi:10.1111/j.1365-313X.2011.04520.x (2011).
18 Pichersky, E. & Gang, D. R. Genetics and biochemistry of secondary metabolites in plants: An evolutionary perspective. *Trends Plant Sci* **5**, 439-445 (2000).
19 Yahyaa, M. *et al.* Identification and characterization of terpene synthases potentially involved in the formation of volatile terpenes in carrot (*Daucus carota L.*) roots. *J Agr Food Chem* **63**, 4870-4878, doi:10.1021/acs.jafc.5b00546 (2015).
20 Keilwagen, J. *et al.* The terpene synthase gene family of carrot (*Daucus carota L*): Identification of qtls and candidate genes associated with terpenoid volatile compounds. *Front Plant Sci* **8**, doi:193010.3389/fpls.2017.01930 (2017).
21 Iorizzo, M. *et al.* A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nat Genet* **48**, doi:10.1038/ng.3565 (2016).
22 Zerbe, P. & Bohlmann, J. Plant diterpene synthases: Exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol* **33**, 419-428, doi:10.1016/j.tibtech.2015.04.006 (2015).
23 Nagegowda, D. A., Gutensohn, M., Wilkerson, C. G. & Dudareva, N. Two nearly identical terpene synthases catalyze the formation of nerolidol and linalool in snapdragon flowers. *Plant J* **55**, 224-239, doi:10.1111/j.1365-313X.2008.03496.x (2008).
24 Cyr, A., Wilderman, P. R., Determan, M. & Peters, R. J. A modular approach for facile biosynthesis of labdane-related diterpenes. *J Am Chem Soc* **129**, 6684. doi:10.1021/ja071158n (2007).
25 Sharma, K. D., Karki, S., Thakur, N. S. & Attri, S. Chemical composition, functional properties and processing of carrot-a review. *J Food Sci Tech Mys* **49**, 22-32, doi:10.1007/s13197-011-0310-7 (2012).
26 Habegger, R. & Schnitzler, W. H. Aroma compounds in the essential oil of carrots (D*aucus carota L. ssp sativus*). 1. Leaves in comparison with roots. *J Appl Bot-Angew Bot* **74**, 220-223 (2000).
27 Jayaraj, J., Rahman, M., Wan, A. & Punja, Z. K. Enhanced resistance to foliar fungal pathogens in carrot by application of elicitors. *Ann Appl Biol* **155**, 71-80, doi:10.1111/j.1744-7348.2009.00321.x (2009).
28 Guerin, P. M. & Ryan, M. F. Relationship between root volatiles of some carrot cultivars and their resistance to the carrot fly, *Psila rosae*. *Entomol Exp Appl* **36**, 217-224, doi:DOI 10.1111/j.1570-7458.1984.tb03431.x (1984).
29 Buttery, R. G., Seifert, R. M., Guadagni, D. G., Black, D. R. & Ling, L. C. Characterization of some volatile constituents of carrots. *J Agr Food Chem* **16**, 1009-&, doi:DOI 10.1021/jf60160a012 (1968).
30 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).
31 Grabherr, M. G. *et al.* Full-length transcriptome assembly from rna-seq data without a reference genome. *Nat Biotechnol* **29**, 644-U130, doi:10.1038/nbt.1883 (2011).
32 Hu, B. *et al.* Gsds 2.0: An upgraded gene feature visualization server. *Bioinformatics* **31**, 1296-1297, doi:10.1093/bioinformatics/btu817 (2015).

33 Emanuelsson, O., Nielsen, H. & Von Heijne, G. Chlorop, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* **8**, 978-984, doi:DOI 10.1110/ps.8.5.978 (1999).
34 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol Syst Biol* **7**, doi: 10.1038/msb.2011.75 (2011).

SUPPLEMENTAL DATA

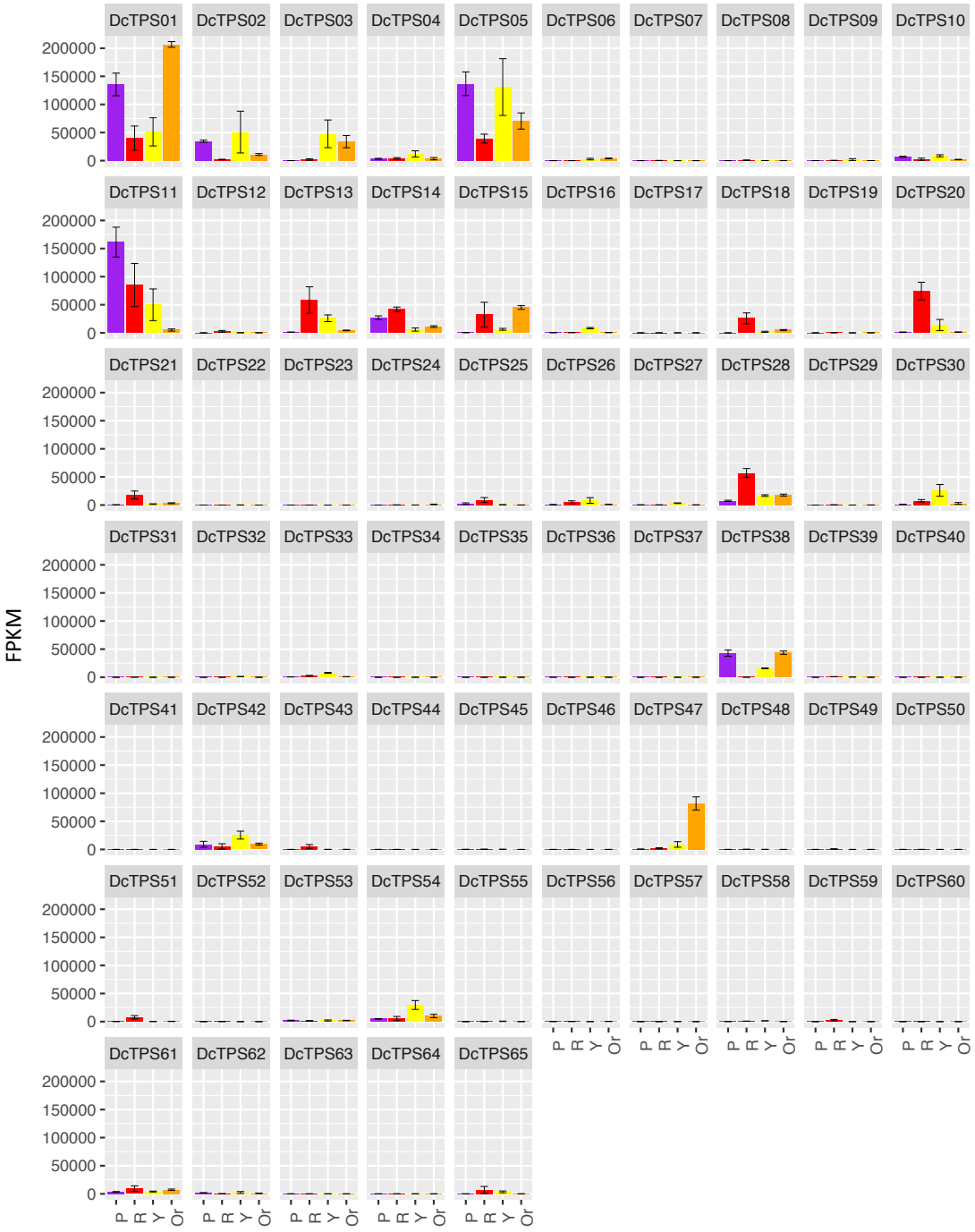


Figure S1. RNA-seq analysis of the 65 TPS gene models predicted by Keilwagen et al., (2017) across four colored carrot varieties. Gene expression is presented as reads per kilobase of transcript per million mapped reads (RPKM). P: purple (cv. P7262), R: red (cv. R6637), Y: yellow (cv. Y9244A) and O: orange (cv. B493B).

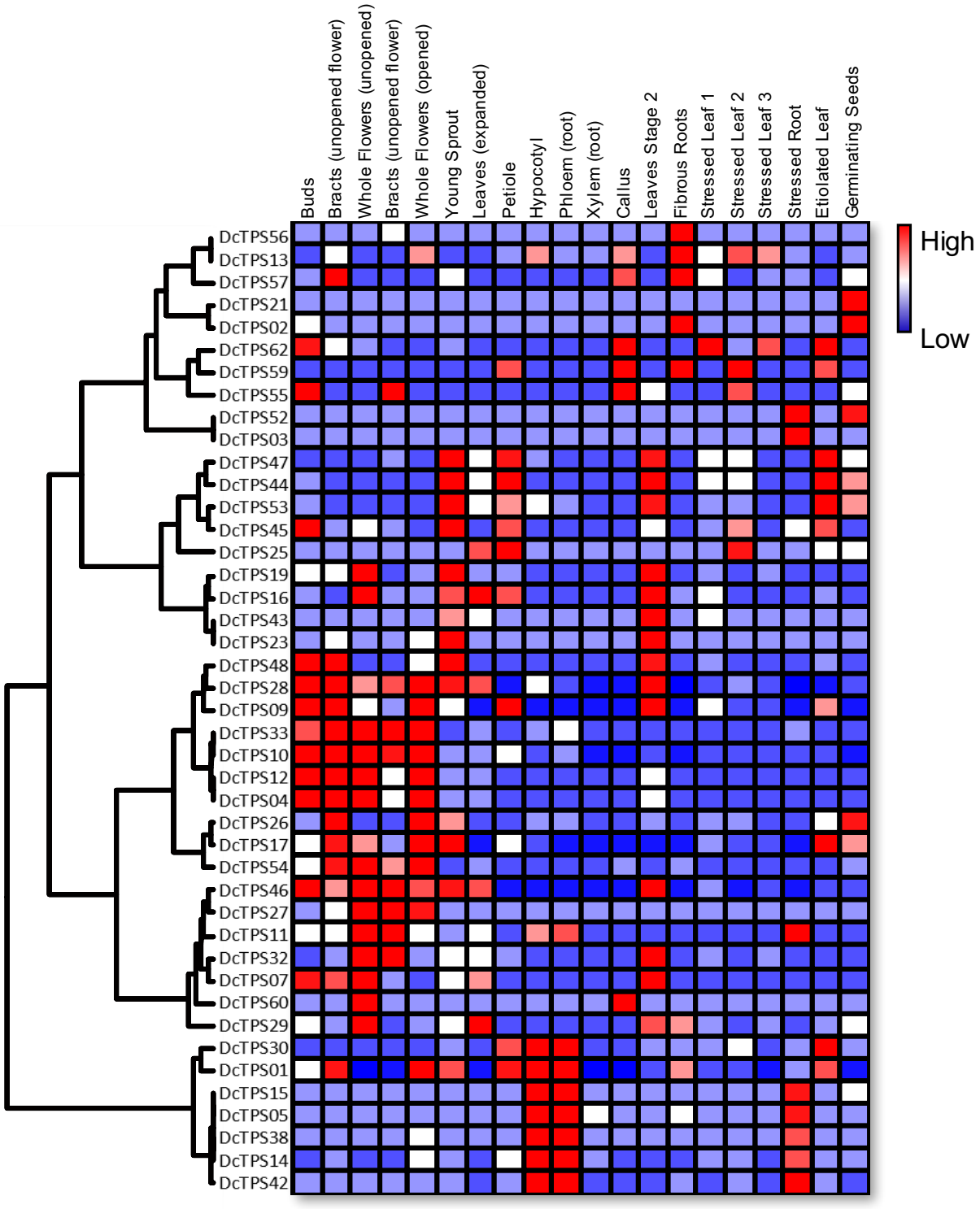


Figure S2. Hierarchical cluster analysis of *TPS* expression profiles across carrot tissues. Heat map commparing relative transcript abundance for all 43 *TPS* genes candidate in FPKM (Fragments Per Kilobase of transcript per Million mapped reads) across 20 tissue specific gene expression data sets from Iorizzzo et al. (2016).

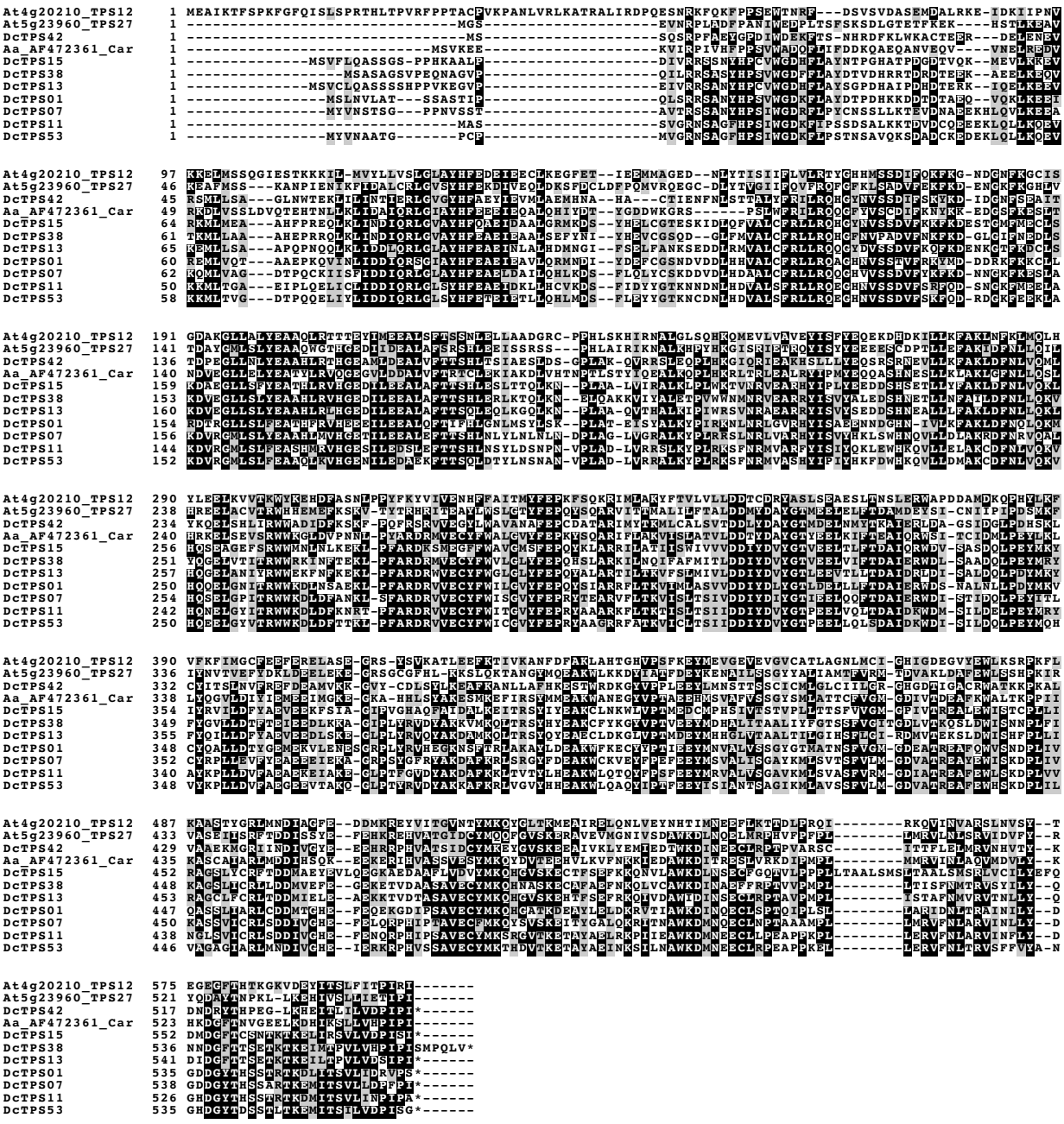


Figure S3. Carrot TPS-a sub-family amino acid alignment using Clustal W (Larkin et al. 2007, *Bioinformatics*) and visualized BoxShade (v3.21).

Figure S4. Carrot TPS-b sub-family amino acid alignment using Clustal W (Larkin et al. 2007, *Bioinformatics*) and visualized BoxShade (v3.21).

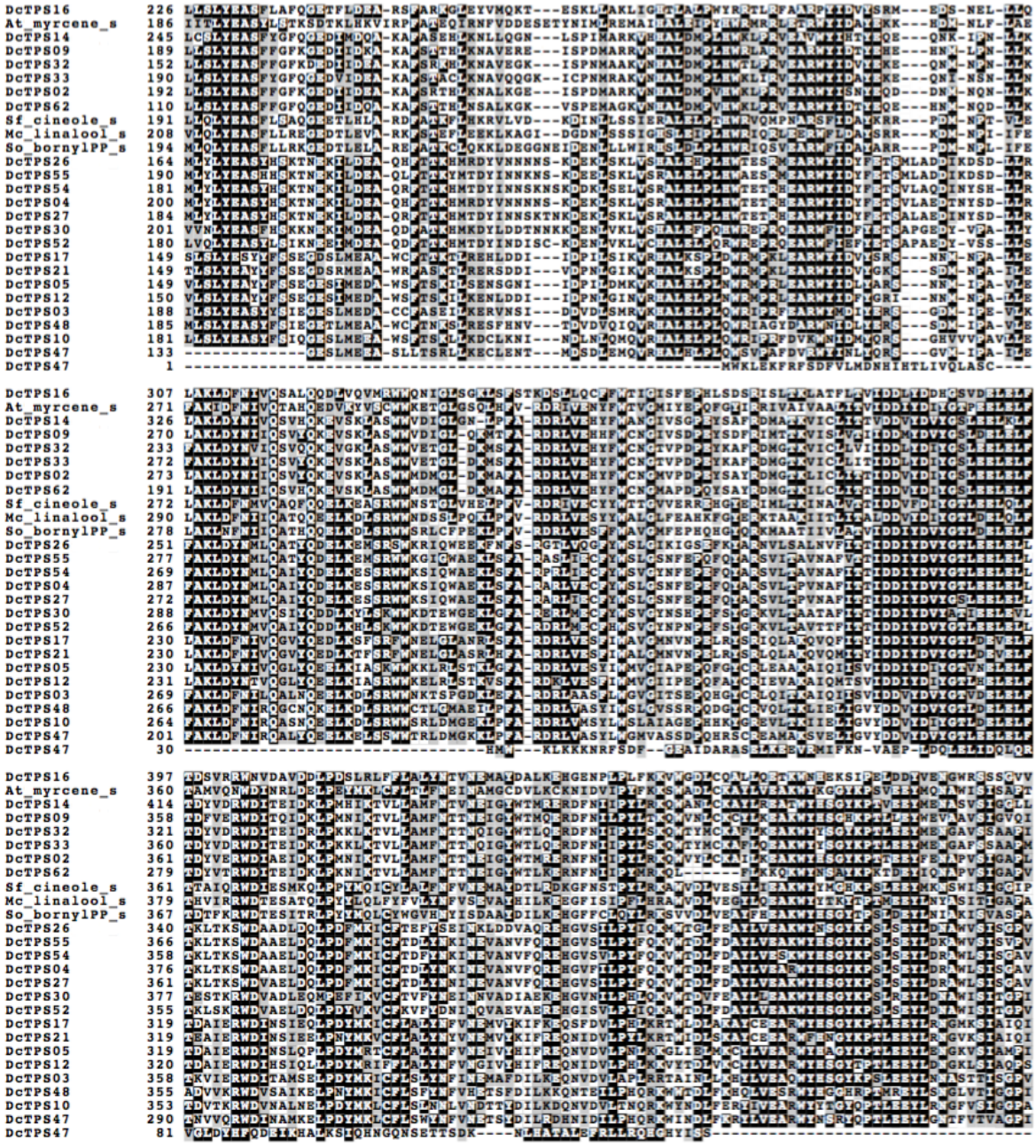


Figure S4. Continued

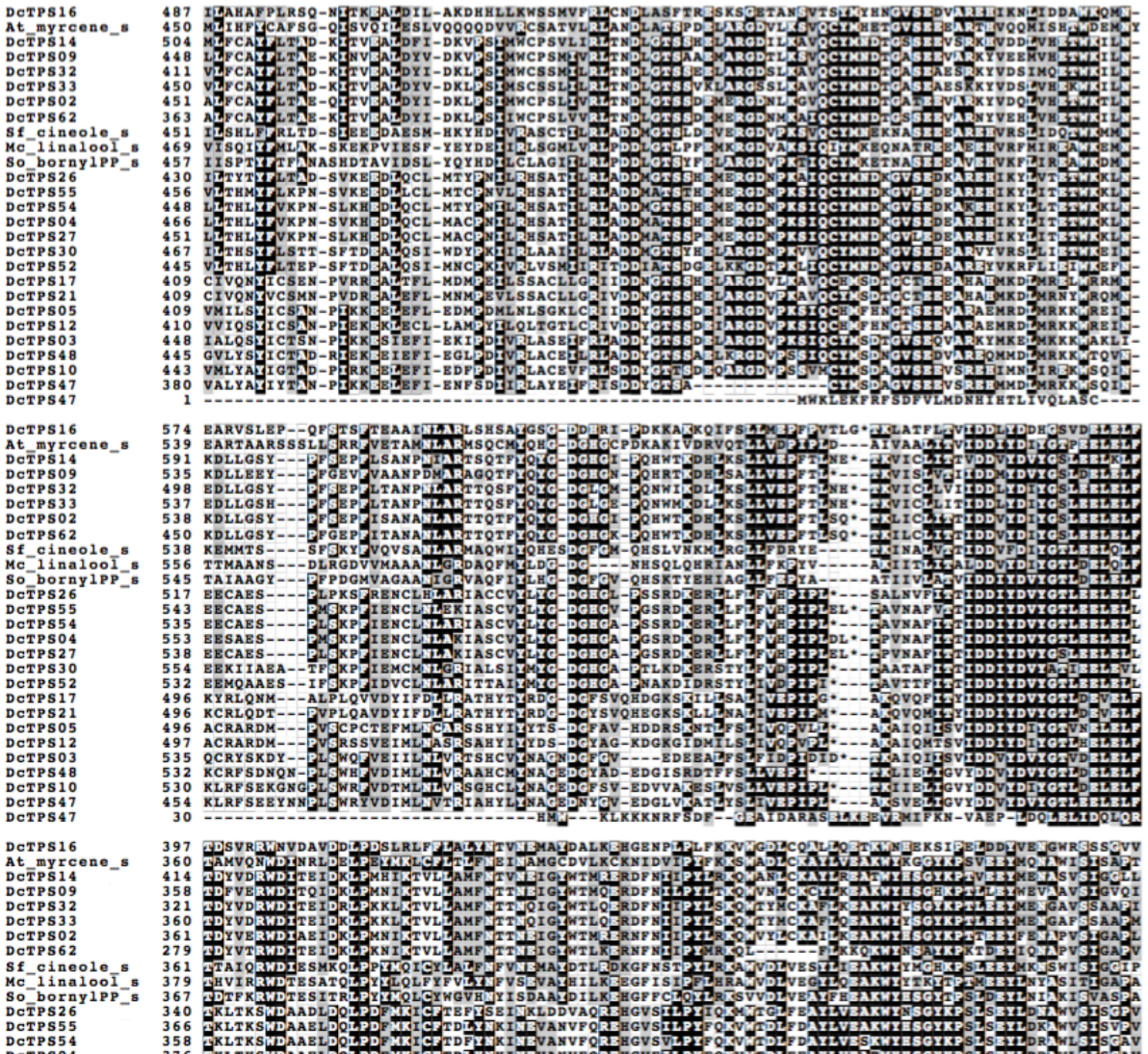


Figure S4. Continued

Figure S5. Carrot TPS-g sub-family amino acid alignment using Clustal W (Larkin et al. 2007, *Bioinformatics*) and visualized BoxShade (v3.21).

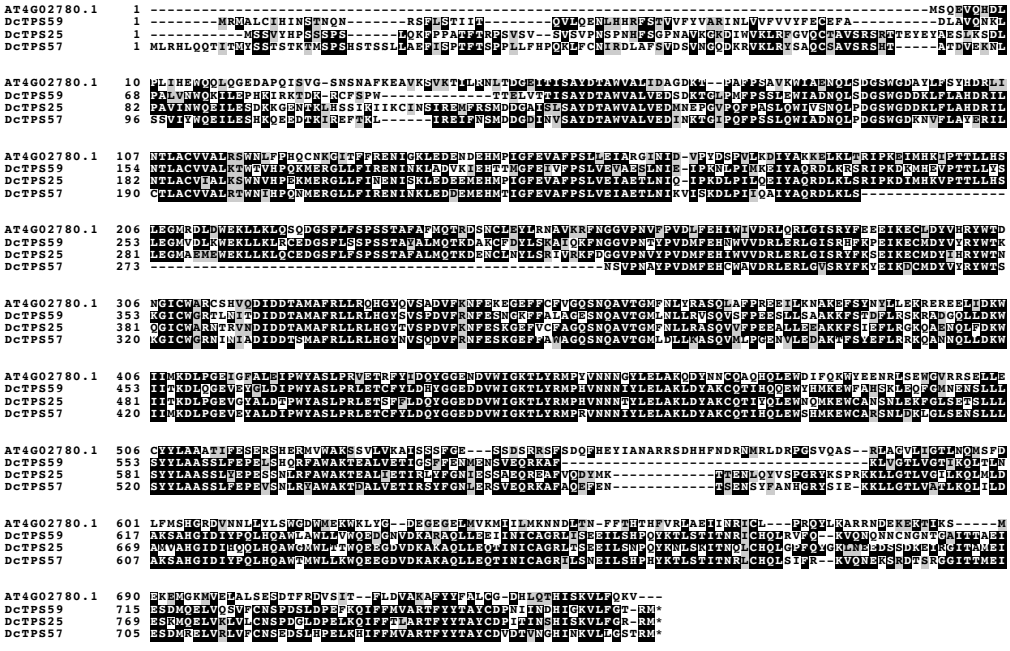


Figure S6. Carrot TPS-c sub-family amino acid alignment using Clustal W (Larkin et al. 2007, *Bioinformatics*) and visualized BoxShade (v3.21).

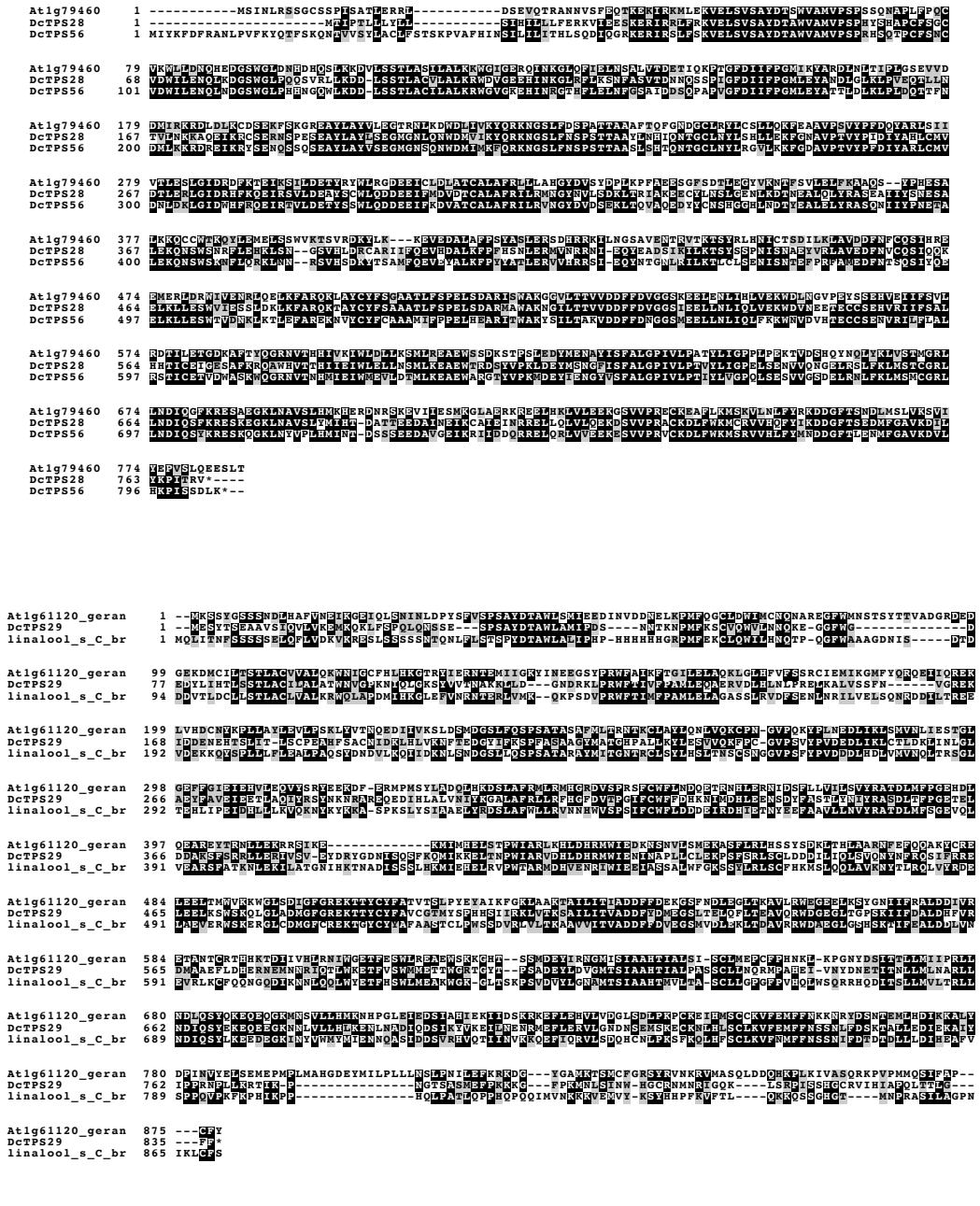


Figure S7. Carrot TPS-e/f sub-family amino acid alignment using Clustal W (Larkin et al. 2007, *Bioinformatics*) and visualized BoxShade (v3.21).

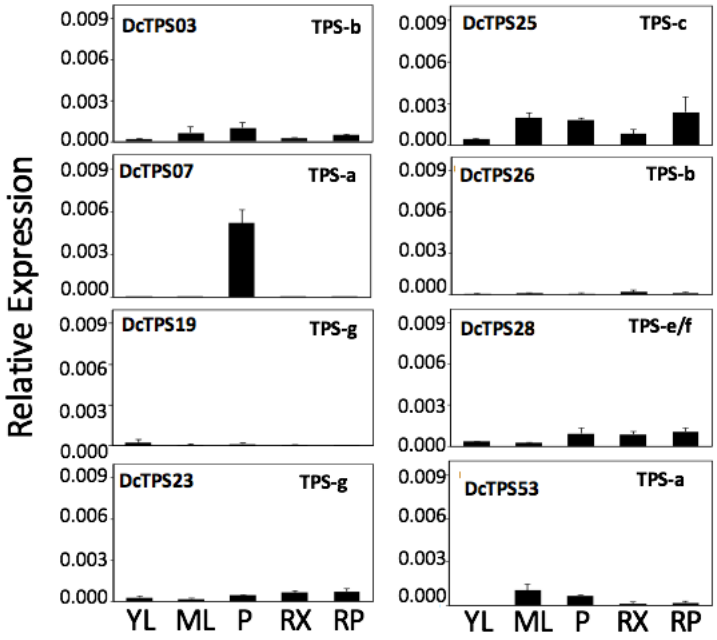


Figure S8. Analysis of gene expression by RT-qPCR for additional gene candidates. Relative expression levels across tissues for each TPS was calculated using the relative quantification method and normalized to actin. Samples were run in biological and technical triplicate and error bars indicate standard deviation from the mean. YL: young leaf, ML: mature leaf, P: petiole, RX: root xylem and RP: root phloem.

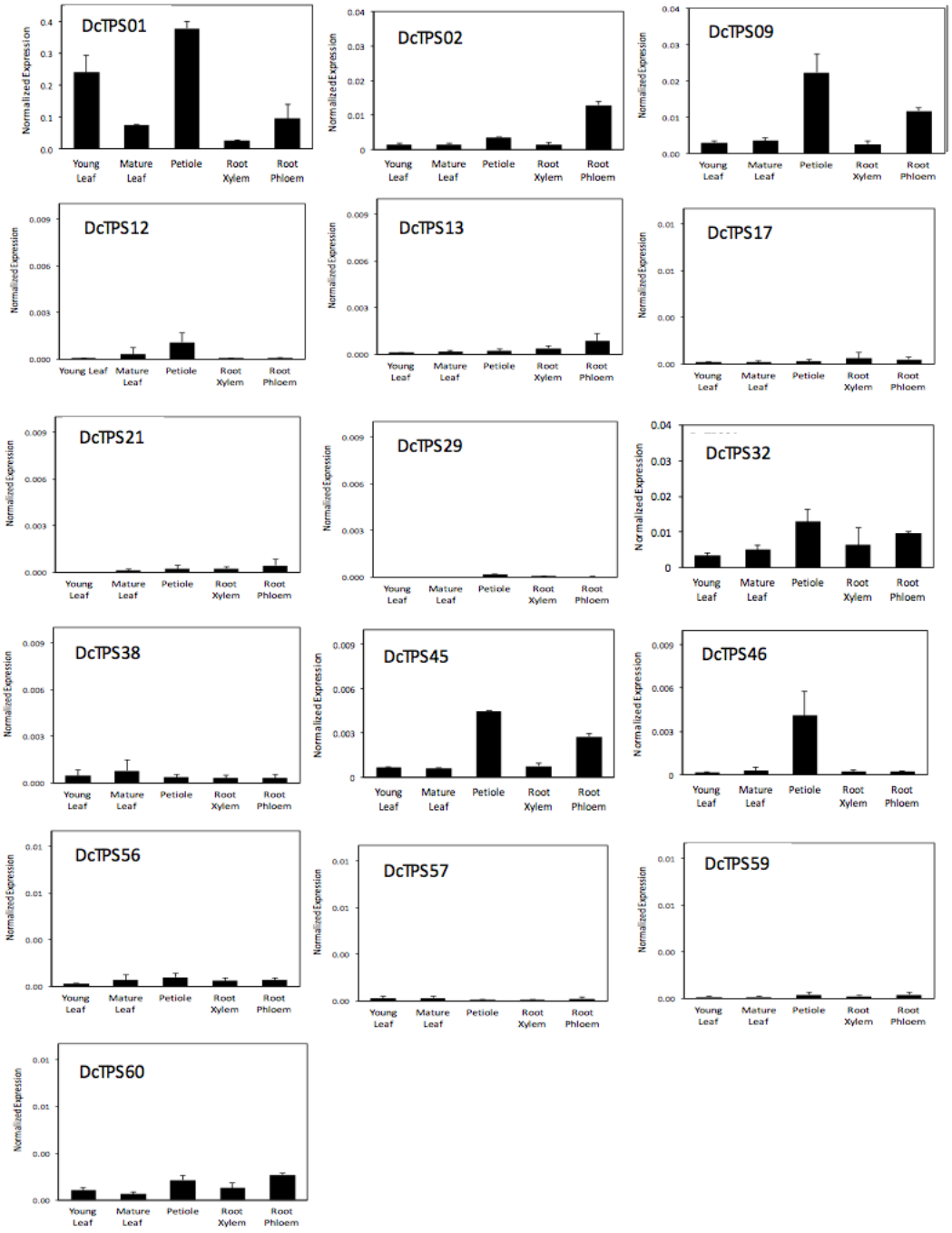


Figure S9. Analysis of gene expression of by RT-qPCR for additional genes of interest. Relative expression levels across tissues for each TPS was calculated using the relative quantification method and normalized to actin. Samples were run in biological and technical triplicate and error bars indicate standard deviation from the mean.

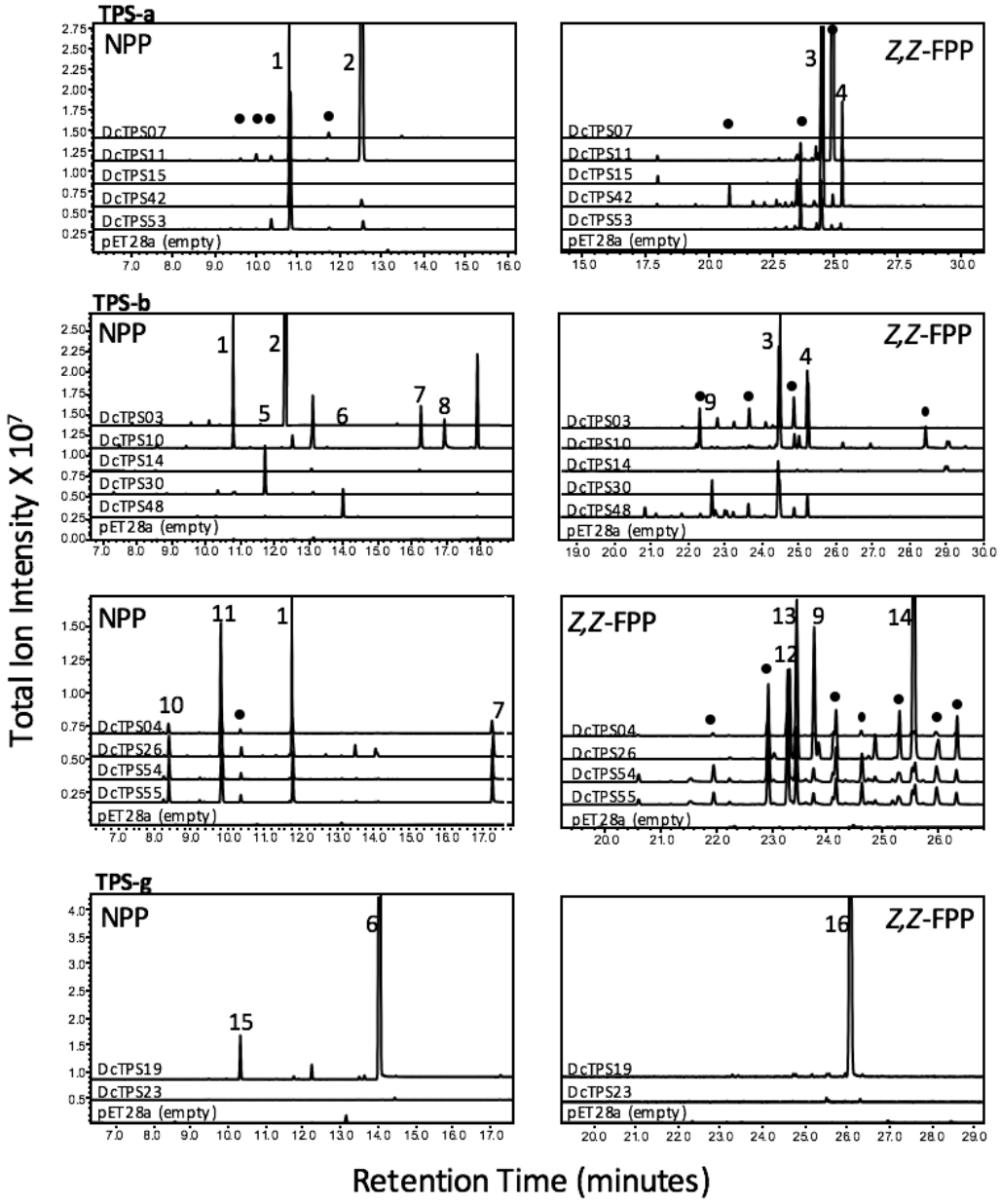


Figure S10. SPME-GC-MS analysis of enzyme assays with partially purified proteins. Assays were incubated for 5 minutes in the presence of a SPME fiber prior to thermal desorption. 1: limonene, 2: α-terpinolene, 3: β-bisabolene, 4: α-bisabolene, 5: γ-terpinene, 6: linalool, 7: α-terpineol, 8: nerol, 9: sesquisabinene, 10: α-pinene, 11: sabinene, 12: β-farnesene, 13: bergamotene, 14: bisabolene isomer, 15: β-myrcene, 16: nerolidol. Compound identification is based on comparisons to authentic standards and essential oils provided by Dr. Tobias Köllner (MPI-CE). NPP: neryl diphosphate, *Z,Z*-FPP: all-*cis*-farnesyl diphosphate.

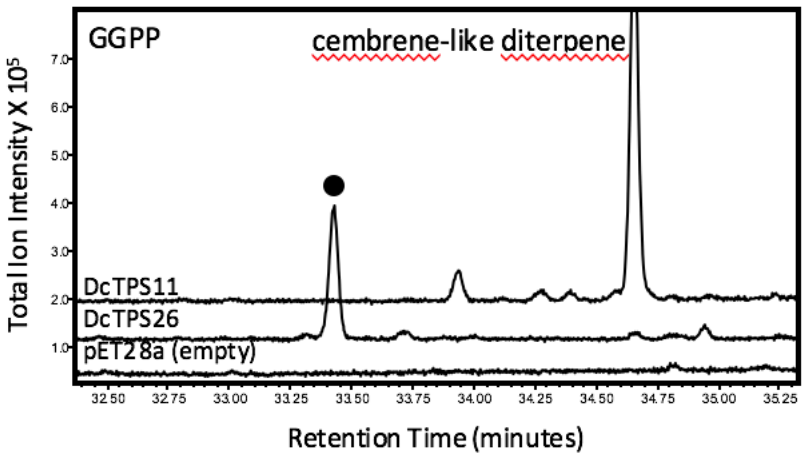


Figure S11. SPME-GC-MS analysis of enzyme assays with GGPP and partially purified DcTPS11, DcTPS26 or empty vector control. Assays were incubated for 5 minutes in the presence of a SPME fiber prior to thermal desorption. Putative compound identification is based on comparisons to reference libraries (NIST and Wiley).

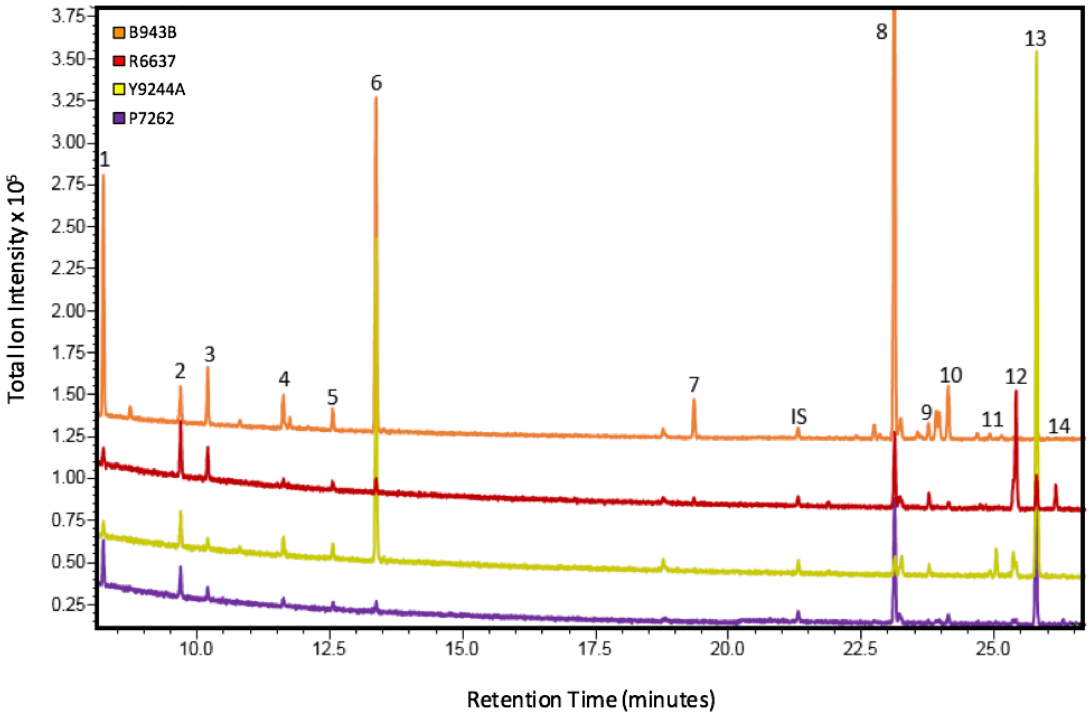


Figure S12. GC-MS analysis of hexane extracts from 11-week old field grown carrot roots. 1: α-pinene, 2: β-pinene, 3: β-myrcene, 4: D-limonene, 5: γ-terpinene, 6: α-terpinolene, 7: bornyl acetate, 8: (*E*)-β-caryophyllene, 9: β-farnesene, 10: α-humulene, 11: β-bisabolene, 12: (*Z*)-γ-bisabolene, 13: (*E*)-γ-bisabolene, 14: α-bisabolene, IS: internal standard 1-bromodecane. Compound identification is based on comparisons to authentic standards and essential oils provided by Dr. Tobias Köllner (MPI-CE). Purple (cv. P7262), Red (cv. R6637), Yellow (cv. Y9244A) and Orange (cv. B493B).

SUPPLEMENTAL TABLES

Table S1. Volatile terpenes from triplicate carrot tissues quantified by GC-FID using multi point internal standard calibration curves to authentic α-pinene and α-humulene. Identification was based on comparison to mass spectral libraries (NIST, Wiley) and comparison to authentic standards and oils when available. Values are reported as micrograms per gram fresh weight.

Compound	Leaf	Petiole	Root	Phloem	Xylem
α-pinene	6.00 ± 0.39	3.71 ± 0.12	1.19 ± 0.01	1.01 ± 0.01	1.06 ± 0.00
camphene	0.26 ± 0.07	0.17 ± 0.09	trace	trace	trace
β-pinene	0.59 ± 0.00	0.30 ± 0.00	0.21 ± 0.02	0.07 ± 0.00	0.07 ± 0.02
β-myrcene	3.54 ± 0.44	0.40 ± 0.35	0.07 ± 0.07	0.08 ± 0.04	0.10 ± 0.04
p-cymene	0.06 ± 0.20	0.05 ± 0.02	0.05 ± 0.03	trace	trace
limonene	0.46 ± 0.00	0.29 ± 0.05	0.09 ± 0.00	0.05 ± 0.02	0.08 ± 0.03
E-β-ocimene	1.87 ± 0.4	0.43 ± 0.1	trace	trace	trace
terpinolene	0.1 ± 0.02	0.06 ± 0.03	0.85 ± 0.10	0.03 ± 0.00	0.02 ± 0.02
γ-terpinene	0.88 ± 0.30	0.76 ± 0.05	0.41 ± 0.20	0.39 ± 0.09	0.39 ± 0.20
E-β-caryophyllene	12.73 ± 0.10	5.61 ± 0.20	2.13 ± 1.17	1.22 ± 0.40	1.70 ± 1.40
Δ-elemene	4.57 ± 0.60	1.41 ± 0.06	0.05 ± 0.00	0.07 ± 0.02	0.20 ± 0.01
α-humulene	0.92 ± 0.02	0.36 ± 0.03	0.12 ± 0.08	0.12 ± 0.06	0.17 ± 0.13
germacrene-D	4.26 ± 0.93	0.68 ± 0.15	0.02 ± 0.00	0.03 ± 0.01	0.04 ± 0.03
β-bisabolene	1.34 ± 0.02	0.45 ± 0.06	0.08 ± 0.00	0.13 ± 0.03	0.35 ± 0.02

Table S2. Predicted transit peptides based on ChloroP (v1.1) analysis of full-length proteins.

Name	Length	Score	cTP	CS-score	cTP-length
*Dc*TPS01	560	0.452	-	3.619	6
*Dc*TPS02	594	0.555	Y	0.149	36
*Dc*TPS03	589	0.548	Y	5.606	40
*Dc*TPS04	608	0.565	Y	1.456	43
*Dc*TPS05	550	0.436	-	1.234	14
*Dc*TPS07	563	0.498	-	3.592	13
*Dc*TPS09	589	0.526	Y	3.824	33
*Dc*TPS10	588	0.519	Y	4.694	34
*Dc*TPS11	551	0.442	-	4.265	35
*Dc*TPS12	551	0.455	-	0.006	15
*Dc*TPS13	566	0.442	-	3.438	4
*Dc*TPS14	647	0.431	-	4.367	88
*Dc*TPS15	577	0.457	-	4.698	6
*Dc*TPS16	632	0.454	-	9.959	69
*Dc*TPS17	551	0.443	-	-0.508	30
*Dc*TPS19	562	0.505	Y	-2.852	37
*Dc*TPS21	551	0.471	-	-0.470	30
*Dc*TPS23	553	0.469	-	7.505	54
*Dc*TPS25	826	0.588	Y	6.700	59
*Dc*TPS26	570	0.447	-	-1.257	30
*Dc*TPS27	593	0.522	Y	1.923	44
*Dc*TPS28	769	0.431	-	1.214	50
*Dc*TPS29	836	0.462	-	-1.711	47
*Dc*TPS30	609	0.571	Y	7.929	43
*Dc*TPS32	554	0.455	-	-0.817	13
*Dc*TPS33	593	0.53	Y	4.210	17
*Dc*TPS38	567	0.448	-	7.202	2

*Dc*TPS42	541	0.43	-	0.692	8
*Dc*TPS43	562	0.428	-	4.212	23
*Dc*TPS44	582	0.455	-	1.078	30
*Dc*TPS45	517	0.432	-	5.587	25
*Dc*TPS46	534	0.436	-	2.549	16
*Dc*TPS47	512	0.435	-	9.424	50
*Dc*TPS48	587	0.534	Y	1.490	39
*Dc*TPS52	587	0.538	Y	5.584	26
*Dc*TPS53	560	0.461	-	6.295	4
*Dc*TPS54	588	0.565	Y	3.527	43
*Dc*TPS55	598	0.535	Y	0.822	34
*Dc*TPS56	804	0.492	-	3.792	39
*Dc*TPS57	763	0.554	Y	6.064	78
*Dc*TPS59	772	0.479	-	-2.482	5
*Dc*TPS60	554	0.482	-	7.111	56
*Dc*TPS62	506	0.431	-	-1.898	2

Table S3. RT-qPCR primers used in this study.

Primer_Name	**Primer_Seq**	**Tm**	**Amplicon (bp)**
*Dc*TPS01F_int	CAACATCTTCAGCTTCTACC	53.7	281
*Dc*TPS01R_int	CATACGCTGCAACACAGC	57.2	281
*Dc*TPS02F_int	CTAACACTTGTGCTCTAAAACC	55.3	248
*Dc*TPS02R_int	CAGCAGCTCCAGCTTGTC	58.4	248
*Dc*TPS32F_int	GTTTGGGAGGTTCAAAGATG	54.5	391
*Dc*TPS32R_int	CCATGCTCCCTGAGC	52.9	391
*Dc*TPS45F_int	CTTTGAATTTGCGGAAAACTC	54.6	276
*Dc*TPS45R_int	CTTGATCCATTGCCTTATACTGGC	59.8	276
*Dc*TPS46F_int	GGTGAAGAACGTGGTACG	55.5	218
*Dc*TPS46R_int	CGAAGTTGTCAAAAACATCTGC	57.1	218
*Dc*TPS19F_int	GATATGGCAAATATTAAGCGTG	53.8	384
*Dc*TPS19R_int	GCGGCTCAAGTATTCAG	52.5	384
*Dc*TPS25F_int	CCATCCCTCCTCCTCC	53.9	301
*Dc*TPS25R_int	CGACGAATGAAGCTTAGTGTTTTC	58.9	301
*Dc*TPS30F_int	GCATGAGGTTCTCTTCATTG	54.4	292
*Dc*TPS30R_int	CTCACAATTTCCTTCAATCC	52.1	292
*Dc*TPS09F_int	GCTGCACCTCCCCG	55.8	251
*Dc*TPS09R_int	GCATGGATCAGAAGCTTCG	56.5	251
*Dc*TPS04F_int	GTGAGTGTTGGTATGAGTTTTTCTG	58.6	286
*Dc*TPS04R_int	CCTTCAAATCCCCTGC	51.2	286
*Dc*TPS26F_int	CTGCACTAGGAATGATACTGC	56.2	325
*Dc*TPS26R_int	CCACTTGTCATTCTTGTTCAC	55.4	325
*Dc*TPS56F_int	CAGGGGAGAAAAGAAAGAATC	54.2	297
*Dc*TPS56R_int	CTCAAGAAAATGTGTACCCC	53.9	297
*Dc*TPS28F_int	CTTTCACCCAAATTCTTACATAC	53.7	281
*Dc*TPS28R_int	CATCCTTAAGTTGATTTTCCAATATC	55	281
*Dc*TPS14F_int	CACGGTCTTTTTTCCCC	52.8	334
*Dc*TPS14R_int	GCCTATGAACAGAGTCGAG	54.8	334
*Dc*TPS17F_int	CACACACTGCCAGTTC	52.1	291
*Dc*TPS17R_int	GATTTCCCAGTGATTATTATTTTG	52.5	291
*Dc*TPS57F_int	CATTCCACGTCTTCTTTACTTGC	58.3	345
*Dc*TPS57R_int	CAGAAACATTTATATCTCCATCGTC	55.9	345
*Dc*TPS59F_int	GACAATAATTACACAGGATTCG	53.1	326
*Dc*TPS59R_int	CTTGAAGGGAACATAGGTAGTCC	58	326
*Dc*TPS23F_int	GTTAGGCAAGAACTCAAGG	53.3	355
*Dc*TPS23R_int	CTTCATCTTCTATCCTCAGC	52.6	355
*Dc*TPS60F_int	GCTAATAACAAGGTAGGAATCG	54.5	331
*Dc*TPS60R_int	GTTTTCTCACTAGCTTCTTCTCCTTG	59.9	331

*Dc*TPS44F_int	GCCTTGGATCATTTCAGTTTG	55.7	303
*Dc*TPS44R_int	CAATAAGCTCCAACTTTTCCC	55.1	303
*Dc*TPS29F_int	CTCTCCACAGCTACAAAATTC	54.6	335
*Dc*TPS29R_int	CTGGAAAAACGATGGTAAAC	52.8	335
*Dc*TPS16F_int	GCACTTTCCTCAGTTGGG	55.6	289
*Dc*TPS16R_int	CTAAATCATTGATACTCCACGC	55.1	289
*Dc*TPS15F_int	GTTCTCCACCACATAAGGCTG	58.6	307
*Dc*TPS15R_int	CAAAGCTCATGATACGAATC	52.2	307
*Dc*TPS38F_int	GGAACAAAATGCTGGGG	53.2	329
*Dc*TPS38R_int	CGACCATGAAAAGGCC	52.2	329
*Dc*TPS42F_int	CAGTCTCGGCCATTCGC	58.2	319
*Dc*TPS42R_int	CTTAAGATTCGGAAGTAAAGTGC	55.7	319
*Dc*TPS10F_int	CATCTATAATTTTCCCAGTTTCAAC	54.7	330
*Dc*TPS10R_int	CAACTCTCTGCAACTGATC	53.4	330
*Dc*TPS11F_int	GCTAAAACAGGAAGTGAAG	51.4	395
*Dc*TPS11R_int	GAGAGGTGGTAAATTCAAGAG	53.7	395
*Dc*TPS03F_int	CTATAGTCCACATTTCTCCCC	54.9	377
*Dc*TPS03R_int	CTTCAAGGCACGTTCTATTTC	55.3	377
*Dc*TPS07F_int	CATACTGTAACTCGTCGC	52.8	338
*Dc*TPS07R_int	GTTGTTGTCCTTGAACTTGTAG	55.5	338
*Dc*TPS53F_int	GTTAACGCTGCAACTGG	53.8	353
*Dc*TPS53R_int	CTTAGTGCAACATCATGTAAATTATC	55.3	353
*Dc*TPS12F_int	GCTGCCTGTTATTAGACG	52.9	339
*Dc*TPS12R_int	CCATGTTGTCTGAGAAGTC	53	339
*Dc*TPS48F_int	CTCCATTAGTCACCTGCTGC	58.3	282
*Dc*TPS48R_int	GATCTAGCGGCTCAACAG	54.9	282
*Dc*TPS13F_int	GTGTGCCTGAGATTGTTCG	56.7	339
*Dc*TPS13R_int	GTAACCGGAAGCAAAGAG	53.1	339
*Dc*TPS21F_int	CAGGCGTTCGGGGAATTAC	58.6	361
*Dc*TPS21R_int	CATAAAACTCTCGAATATATCTTC	51.6	361

Table S4. Primers for gene cloning.

Primer_Name	**Primer_Seq**	**Purpose**
*Dc*Actin_F	CGGTATTGTGTTGGACTCTGGTGAT	Housekeeping gene
*Dc*Actin_R	CAGCAAGGTCAAGACGGAGTATGG	Housekeeping gene
*Dc*PP2A_F	GTGTATCAATGTACCACCAGCAACT	Housekeeping gene
*Dc*PP2A_R	GCTCACCAAGGAACATGACTTCTT	Housekeeping gene
Tubulin_F	TCTTGGAGGTGGCACAGGAT	Housekeeping gene
Tubulin_R	ACCTTAGGAGACGGGAACACAGA	Housekeeping gene
*Dc*TPS1_F	ATGTCTCTCAATGTTCTGGC	Control TPS
*Dc*TPS1_R	TGATGGAACCCGATCAATGA	Control TPS
*Dc*TPS19_F	GGATCCATGTTTAAACAGGCACAAG	Cloning
*Dc*TPS19_R	CTCGAGTCAACAATTTAAAGGTAAAAC	Cloning
*Dc*TPS25_F	GGATCCATGACTGCAGTTTCAAG	Cloning
*Dc*TPS25_R	CTCGAGTTACATCCTACGGCC	Cloning
*Dc*TPS23_F	GGATCCATGGCGATTGTTAGG	Cloning
*Dc*TPS23_R	CTCGAGCTAGAACAATTGTTCCC	Cloning
*Dc*TPS16_F	GGATCCATGGCTGCAGTACTG	Cloning
*Dc*TPS16_R	CTCGAGTTAACCAAGAGTTAC	Cloning
*Dc*TPS04_TCF	GGATCCATGATGGAAAGAGTTTCTTG	Cloning
*Dc*TPS04_TCR	CTCGAGTCAGGTGATCGAAAGAGGTTC	Cloning
*Dc*TPS05F_TC	GAATTCATGAGTACTCTTATTGTGAATC	Cloning
*Dc*TPS05R_TC	GTCGACTTACAGAAGAACAGG	Cloning
*Dc*TPS03F_TC	GGATCCATGGGTTCTCAATCC	Cloning
*Dc*TPS03R_TC	GTCGACTCATATATCGATGGGGTC	Cloning
*Dc*TPS07R_TC	GGATCCATGAATTCTACTTCTGG	Cloning

*Dc*TPS07F_TC	CTCGAGTTATATAGGAAAAGGGTC	Cloning
*Dc*TPS53R_TC	GGATCCATGGCTATGTATGTTAACG	Cloning
*Dc*TPS53F_TC	CTCGAGTTATGCTGGAATCGGATC	Cloning
*Dc*TPS03B_TCF	GGATCCATGTCCATGGGAATTTCTG	Cloning
*Dc*TPS03B_TCR	GTCGACTCAGTCAATATCGATGGGG	Cloning
*Dc*TPS53B_TCR	CTCGAGTTACCCTGAAATCGGATC	Cloning
*Dc*TPS48_F	GGATCCATGGAGGGATCTGTGTCACCAG	Cloning
*Dc*TPS48_R	GTCGACTCATATAGGCTCAACAAGAAGGG	Cloning
*Dc*TPS14_R	CTCGAGTCACTCATTAAGAGTGAAGG	Cloning
*Dc*TPS42_F	GAGCTCATGAGCAGCCAGTCTCG	Cloning
*Dc*TPS42_R	GTCGACTCATATCGGTATCGGATCCACC	Cloning
*Dc*TPS11_F	GGATCCATGGCTTCAAGTGTGGG	Cloning
*Dc*TPS42_R	CTCGAGTTATGCTGGAATGGGATTTATG	Cloning
*Dc*TPS10_F	CTCGAGATGGCCGCTGAAACAAC	Cloning
*Dc*TPS10_R	GTCGACTCAGAGAGGGATGGGCTC	Cloning
*Dc*TPS30_F	GGATCCATGTGCAGTAGCAGTAGTCATG	Cloning
*Dc*TPS30_R	CTCGAGCTATAGAGGTATGGGGTCAAC	Cloning
*Dc*TPS15_F	GAGCTCATGTCTGTGTTTCTTCAGGC	Cloning
*Dc*TPS15_R	GTCGACTCATATTGAGATGGGATCCACTAG	Cloning
*Dc*TPS38_F	CTCGAGATGTCGGCTTCAGCAGG	Cloning
*Dc*TPS38_R	GTCGACTTAGACAAGTTGTGGCATTG	Cloning
*Dc*TPS28_F	GGATCCATGGCCTCACTGGAGTC	Cloning
*Dc*TPS28_R	CTCGAGCTATACTCGGGTAATAGGTTTG	Cloning
*Dc*ZZFPPSF_TC	GGATCCATGAGTAGTCGTCGAACAG	Cloning
*Dc*ZZFPPSR_TC	GTCGACTTATTGCTGCTTTAGTTTACCC	Cloning

Table S5. Terpene compounds identified from random forest analysis and boruta variable selection of colored carrot root tissue. MANOVAs comparing the concentrations of the selected compounds among groups are reported (F-statistic and p value).

Compound	Boruta IMP Factor	F Value	p Value
α-pinene	4.7437306	3.6368	6.40E-02
β-pinene	4.9868365	4.7613	3.448e-02*
α-terpinolene	7.2058394	4.1936	4.659e-02*
bornyl acetate	8.1511563	20.344	4.228e-04***
(*E*)-β-farnesene	8.003681	21.099	3.721e-04***
(*E*)-β-caryophyllene	7.2949134	42.36	2.955e-05***
α-humulene	5.5921902	30.176	1.034e-04***
β-bisabolene	4.4540192	12.795	2.022e-03**
(*E*)-γ-bisabolene	6.7570118	8.3796	7.506e-03**

CHAPTER V. FINAL DISCUSSION AND FUTURE PERSPECTIVES

OVERVIEW

Plants produce a large number of specialized compounds as a form of chemical communication, and for protection against biotic and abiotic stress. Plants are also known to sustain intimate relationships with diverse microbiota thereby enhancing plant productivity and resilience to stress. The plant traits that influence and structure microbial communities by "habitat" filtering remain poorly understood. Plant derived volatile organic compounds (VOCs), which are known to affect the microbial colonization of aboveground tissues, have been lesser explored for their influence on plant-microbe interactions belowground. We investigated the function of volatile terpenoids as chemo-selective factors in roots using the model bioenergy plant switchgrass. We found that switchgrass accumulates the volatile monoterpenoid borneol in roots, which was of particular interest based on literature evidence for its antimicrobial activity and potential for acting as unique bacterial carbon source (Tabanca et al., 2001; Tsang et al., 2016). To determine if borneol is a plant trait necessary for structuring the root microbiome, we first biochemically characterized the terpene synthase (*TPS*) gene family of switchgrass to identify all enzymes that could produce borneol in vitro (Chapter II). We then associated in vitro enzymatic function with constitutive and induced terpenoids in roots and leaves to better understand terpenoid metabolism and function in switchgrass (Chapter II). We have directly, and through collaboration, demonstrated that switchgrass has an extensive volatile and non-volatile terpene metabolism (Muchlinski et al., submitted; Pelot et al., 2018), much more expanded than most other model plants (Tholl et al., 2004; Falara et al., 2011; Keilwagen et al., 2017; Block et al., 2019). Within the TPS-a subfamily of switchgrass, we identified *PvTPS04* as the only terpene synthase that catalyzes the formation of borneol as a major product in vitro. We show that this gene is highly

expressed in leaf tissues and not in roots where the compound accumulates (Chapter III). This suggests a yet undiscovered chemical transport mechanism between leaves and roots, perhaps similar to mechanisms identified in the Lamiaceae (Croteau et al., 1987) or observed for specific maize sesquiterpenes (Köllner et al., 2008). In addition, we demonstrated that switchgrass has a defined root microbiome, and that this microbial community can influence the levels of terpenoids accumulated in switchgrass roots. We further developed RNAi based knock-down lines of *PvTPS04* in attempts to modify the root chemical environment, however we experienced challenges in reducing borneol production likely due to the variation and complexity of the genetic and biochemical background (Chapter III). We, therefore, continued our efforts in a related model grass *Setaria*, and could successfully deplete roots from borneol (Appendix B). However, analysis of microbial communities between wild-type and transgenic lines indicated no significant influence of terpenoids on the root microbiome. In this final Chapter, I highlight the importance of these discoveries and present potential future research directions.

In an applied project, we described terpenoid biosynthesis in the major agricultural crop carrot (*Daucus carota*). Here, we determined central enzymes that contribute to the terpene component of carrot volatile blends, which accounts for the primary aroma and flavor of carrot (Chapter IV). We demonstrated tissue specific expression patterns of carrot terpene synthase genes and correlated expression with compound accumulation in leaves, petioles and roots. We revealed key biosynthetic enzymes which directly correlated with major compounds of carrot flavor and aroma including germacrene-D (DcTPS11), gamma-terpinene (DcTPS30) and alpha-terpinolene (DcTPS03). In addition, we analyzed root volatiles of four aromatically unique colored carrot genotypes (orange-4943B, red-R6637, yellow-Y9244A and purple-P7262). We further showed that colored carrot aroma and flavor can be distinguished by analysis of levels of terpene VOCs.

Interestingly, accumulation of specific terpene compounds rather than chemical diversity was responsible for differences in colored carrot genotypes (Muchlinski et al., submitted). As accumulations of specific terpene compounds can contribute to the undesired flavor in carrot, our work provides a detailed roadmap for future breeding efforts to enhance or manipulate carrot flavor and aroma. I will discuss the importance and potential future research directions of this study in this Chapter.

Terpenoid mediated stress protection in switchgrass

Switchgrass cultivars have a broad spectrum of resistance and susceptibility to biotic and abiotic stress (Parrish and Fike, 2005; Casler et al., 2007; Schuh et al., 2019). Research into underlying mechanisms of protection in specific cultivars has been somewhat limited (Ayoub et al., 2013; Meyer et al., 2014; Ye et al., 2016; Demers et al., 2017; Donze-Reiner et al., 2017). In general, switchgrass is considered to be more tolerant to stress than other model grasses (Casler et al., 2007). For instance, switchgrass cultivars spanning upland and lowland ecotypes have been shown to be more tolerant to herbivory from fall armyworm than maize (Prasifka et al., 2011; Schuh et al., 2019). Interestingly, armyworm larval survival declines rapidly when fed a diet of switchgrass leaves, especially those rich in non-volatile steroidal triterpene saponins (Prasifka et al., 2011). It is therefore conceivable that saponins are involved in direct defense of switchgrass against herbivory. Besides the documented presence of triterpenoids in switchgrass, we have shown that the switchgrass genome contains 70 *TPS* genes that encode enzymes with functions in monoterpenoid, sesquiterpenoid, and diterpenoid biosynthesis (Muchlinski et al., submitted; Pelot et al., 2018). We have further demonstrated that switchgrass (cv. Alamo) mounts a diverse terpenoid-based response when plants are treated with SA, JA, drought, UV, copper sulfate and

herbivory from fall armyworm (Muchlinski et al., submitted; Appendix A; Pelot et al., 2018). This response occurs in both roots and leaves, and follows an orchestrated pattern with tissue specific terpene synthase gene expression (Muchlinski et al., submitted; Appendix A; Pelot et al., 2018).

Although some terpenoids are constitutively produced in switchgrass (i.e. borneol), the majority of volatile and non-volatile terpenoid products were only detected following stress treatments. This is consistent with observations in other model grasses (e.g. sorghum and maize), where terpenoid biosynthetic genes are upregulated when plants are exposed to herbivory and abiotic stress (Rasmann et al., 2005; Zhuang et al., 2012; Vaughan et al., 2015). Terpenoid based stress responses can be linked to direct defense against herbivory based on specialized terpenoids being cytotoxic at elevated levels (Doll-Boscardin et al., 2012). Therefore, induction of specific terpenoids in switchgrass could serve as a direct feeding deterrent. We found that (*E*)-beta-caryophyllene is emitted as part of a sesquiterpenoid volatile blend in response to herbivory or treatment with jasmonic acid in roots and leaves. The release of (*E*)-beta-caryophyllene from maize roots, following herbivory from Western corn rootworm, has been shown to recruit entomopathogenic nematodes which parasitize and kill the herbivore (Rasmann et al., 2005). A similar mechanism occurs in the maize phyllosphere where (*E*)-beta-caryophyllene is released from leaves during attack from lepidopteran larvae, which in turn attracts parasitic wasps that oviposit on the surface of the larvae (Turlings et al., 1990). Based on these studies, and the emission of (*E*)-beta-caryophyllene from roots and leaves, it is plausible such tritrophic interactions occur in switchgrass. A mutant based approach could be used to further test if (*E*)-beta-caryophyllene is an attractant of parasites by silencing the major induced (*E*)-beta-caryophyllene synthase(s) in switchgrass and evaluating for decreased fitness as a result of decreased (*E*)-beta-caryophyllene production.

The sesquiterpene (*E*)-beta-farnesene is released from switchgrass leaves following treatment with SA and this terpene is likely produced by PvTPS16 based on in vitro enzyme assays and correlated gene expression patterns. Although SA regulates pathogen defense responses, there is no evidence that (*E*)-beta-farnesene acts directly in pathogen defense. On the other hand, (*E*)-beta-farnesene serves as an aphid alarm pheromone and may be acting as a repellent to aphids that potentially vector plant pathogens although the role of plant-derived (*E*)-beta-farnesene on aphid behavior has been discussed controversially (Kunert et al., 2010; Jaouannet et al., 2014). Interestingly, pea aphids (*Acyrthosiphon pisum*) infected with fungal pathogens are less sensitive to (*E*)-beta-farnesene, which likely enhances fungal transmission (Roy et al., 1999). The most direct approach would be to generate *PvTPS16* mutants and comparing leaf pathogen infection (load/rates) and aphid colonization rates.

Infection of leaves by rust fungal pathogens has been shown to cause major yield losses in switchgrass (Sykes et al., 2016). However, specific cultivars of switchgrass show enhanced resistance to fungal infection (Sykes et al., 2016). It would, therefore, be interesting to investigate if such resistant cultivars have elevated levels of particular volatile terpenoid compounds that serve as natural fungicides. However, as previously mentioned, the presence of non-volatile triterpene saponins in switchgrass, which are known to have antifungal properties, are also likely involved in warding off fungal pathogens (Lee et al., 2009; Thimmappa et al., 2014). In addition, switchgrass could produce non-volatile phytoalexins from volatile precursors. This could be especially true for in vitro enzymatic products that could not be detected under tested conditions but may exist in a modified non-volatile form in planta. In fact, such non-volatile derivatives have been identified as sesquiterpene acids in maize (Huffaker et al., 2011). Therefore, it is possible

that an expansion of the terpene synthase gene family in switchgrass occurred in conjunction with emerging pathways leading from volatile precursors to non-volatile defensive metabolites.

Besides their roles as defensive compounds, it has recently been shown that diterpenoid phytoalexins are important for drought stress tolerance in maize (Vaughan et al., 2015), and possibly in switchgrass as well (Pelot et al., 2018). However, it is still unclear what roles terpenoids have in abiotic stress protection, other than volatile terpenes such as isoprene and monoterpenes which accumulate during abiotic stress (Loreto and Schnitzler, 2010). We found that switchgrass accumulated many oxygenated terpenoids in roots following droughts stress treatment (Appendix A). Interestingly, this accumulation was not consistent across ecotypes where Dacotah accumulated more putative diterpenoids than Alamo. Based on our transcriptomic analyses of drought stress tissues, it could not be immediately determined if this response was due to de novo formation of terpenoid compounds, or release from non-volatile storage pools (e.g. terpenoid glycosides). It has been suggested that terpenoids could stabilize membranes and quench reactive oxygen species to protect against ozone and high temperature stress (Loreto and Schnitzler, 2010). Therefore, the observed induction of terpenes in switchgrass roots exposed to drought, could indicate terpene involvement in osmoprotective roles in switchgrass. In addition, the root microbiome could be directly influenced by terpenoid accumulation under drought stress conditions. In sorghum, drought stress caused shifts in the root microbial community which was associated with increased metabolite production (Xu et al., 2018). Furthermore, Vaughan et al. (2015) argued that accumulations of phytoalexins in maize during drought stress could serve a defensive function against drought tolerant pathogens. On the other hand, specific diterpenoids could be used to recruit beneficial microbes to help tolerate drought stress. Switchgrass plants cultivated with the ectomycorrhizal fungus *Sebacina vermifera*, for example, show higher

resilience to drought stress than uninoculated plants (Ghimire and Craven, 2011). Terpenoids might further promote root mycorrhizal fungal associations or even recruit drought tolerant microbes that may help tolerate drought stress. Mutant based studies are needed to provide stronger evidence for terpenoid mediated abiotic stress resilience. Future work could aim to engineer plants with modifications in terpene chemical composition and measure effects on plant growth during abiotic stress conditions. Enhanced terpene levels during stress could affect the root microbial community which may also influence plant health. Therefore, mutant based stress studies could also incorporate analysis of microbial communities to better understand the complex interactions which may contribute to plant stress resilience.

The results from this work highlight the importance of switchgrass as a model system for exploring biological functions of terpenoid compounds. Though we have made advances in switchgrass functional genomics, more work is needed to decipher the biological functions of the expanded terpenoid metabolism in switchgrass.

Borneol function and the microbial community of switchgrass

Monoterpenoids have multiple roles in plants including defense against pathogens, plant-plant communications, and serving as unique carbon sources to beneficial bacteria (Kleinheinz et al., 1999; Tabanca et al., 2001; Chen et al., 2004; Riedlmeier et al., 2017; Chen et al., 2018). In switchgrass roots, we found that the monoterpenoid borneol is accumulated at substantial levels (Chapter III). Production of specialized metabolites at high concentrations has a considerable metabolic cost (Gershenzon, 1994) and is likely to be maintained only under particular selective pressures. Borneol is a known antimicrobial, but can also be metabolized by common soil bacteria (Stojanovic et al., 2006; Tsang et al., 2016). We proposed, therefore, that borneol was required

for maintaining a healthy microbiome in switchgrass roots. Although we were unable to support this hypothesis with current results, follow-up investigations in switchgrass are needed to draw further conclusions. In our backup model system *S. viridis*, we significantly lowered borneol levels but did not see any significant effect on the microbiome (Appendix B). This could be due to the adaptability of the microbial community in response to changes in borneol levels, or possible localization differences between compound accumulation and microbial distribution in the root. In vetiver grass it has been shown that terpenes accumulate in cells of the outer cortex (Guidice et al, 2008). The study further suggested that bacteria located in the vicinity of these cells might be metabolizing terpenes. It is currently uncertain whether similar distributions of borneol and possible borneol metabolizing bacteria occur in switchgrass roots and if terpene metabolizing species might represent only a small portion of the microbiome and, therefore, remained unidentified in our analysis. There is growing evidence from studies on triterpenoids and other secondary metabolites with clear effects on root microbial community composition (Huang et al., 2019); however, these non-volatile compounds are exuded from roots and likely have different and more direct effects on root colonizing microbes.

It is also conceivable that borneol could serve another biological function in *Setaria* plants. Future work could determine if borneol has toxic or growth inhibiting effects on root rot pathogens, and whether a reduction in borneol levels makes *Setaria* more susceptible to herbivory or pathogen infection. In rice, for example, the monoterpene limonene was shown to confer resistance to the blast fungus (*Magnaporthe oryzae*). Further, overexpression lines accumulating more limonene, were more resistant to infection than wild-type suggesting limonene directly contributes to the enhanced resistance, presumably by inhibiting spore germination (Chen et al., 2018). Recent work from poplar has also demonstrated that monoterpenes are released from roots in response to

herbivore damage and can inhibit the growth of the oomycete plant pathogen *Phytophtora cactorum* (Lackus et al., 2018).

In addition to putative defensive functions, borneol has been shown to influence root morphology in multiple model systems (Horiuchi et al., 2007; Tian et al., 2014). *Arabidopsis* seedlings exposed to borneol, for example, produced thickened root tips and showed a reduced root growth phenotype (Horiuchi et al., 2007). Further, this response was dependent on the stereochemistry of the compound where (+)-borneol application produced plants with shorter roots than (-)-borneol, although both plants had significantly shorter roots than methanol treated controls (Horiuchi et al., 2007). In the wormwood *Artemisia annua*, exogenous borneol application was found to promote adventitious root formation (Tian et al., 2014). Unlike *Arabidopsis,* borneol (and camphor) accumulate in *A. annua* roots. Therefore, borneol may serve as a chemical cue to promote adventitious root formation, but only in plants that can produce borneol. Similar exogenous borneol application experiments could be done to see if *Setaria* (or switchgrass) root morphology is influenced by the presence of borneol.

Unexpectedly, we found that borneol biosynthesis is occurring in leaves and not in roots where the compound is accumulating. This could be due to higher pools of the substrate GPP in leaf tissues based on a higher number of plastids in leaves of this C4 grass than in roots. In addition, a transport mechanism must exist to mobilize borneol to the roots. Due to the hydrophobic nature of terpene compounds, it is likely that borneol would first need to be converted to a more polar intermediate (e.g. via addition of an acetate group or glycosylation) prior to its transport through the phloem. Once the compound arrives in the root, the chemical modification would be reversed releasing borneol, which can be further converted to camphor. To further investigate borneol biosynthesis and function, pulse-labeling $^{14}CO_2$ studies could be used to track the flow of borneol

throughout the plant and potentially into rhizosphere microorganisms (Vidal et al., 2018). For these experiments, switchgrass leaves would be exposed to $^{14}CO_2$ and labeled borneol isolated by liquid extraction. Labeled borneol would then be applied to switchgrass leaves and tissues/microbes would be collected for analysis of radioactivity by radio GC-MS. In addition, any non-volatile intermediates or conversion products could be identified similar to observations in sage (Croteau et al., 1987).

Genetic determinants of carrot aroma and flavor

Carrot sensory quality is known to be negatively affected by abiotic and biotic stressors (Rosenfeld et al., 2002; Alegria et al., 2012; Seljasen et al., 2013). Carrot palatability is substantially reduced due to accumulation of terpenoid compounds resulting in a harsh and bitter taste (Rosenfeld et al., 2004). Accumulation of specific terpenoids in carrot roots e.g. (*E*)-beta-caryophyllene, as a result of temperature stress, are known to reduce palatability and potentially marketability (Rosenfeld et al., 2002). This observed increase could be due to terpenes providing a protective function for the plant experiencing elevated temperature stress by enhancing membrane stability. Therefore, future studies should address whether TPS gene expression is induced in response to temperature stress and if specific genes are more responsive to the stress than others and, hence, have a stronger effect on changes in flavor. Additionally, an RNAi based gene silencing approach could be used to knock down particular TPS genes and potentially improve carrot taste at elevated temperature. In general, our work has provided a resource for engineering flavor and aroma in carrot. Similar to other domesticated crops e.g. tomato, carrot flavor and aroma have been subjected to targeted breeding for enhanced yield, size, shape and stress resilience (Aharoni et al., 2005; Jayaraj et al., 2009; Simon, 2010; Gascuel et al., 2017).

However, such breeding efforts can reduce flavor and aroma traits (Aharoni et al., 2005). Therefore, work should focus on preserving strong flavor and aroma while maintaining other domesticated traits (e.g. yield). Moreover, the contribution of the terpene rich carrot essential oil to protection against pests and pathogens is largely unexplored. Alternaria leaf blight caused by the fungus *Alternaria dauci* causes major foliar damage to carrot. Interestingly, analysis of terpene compounds in resistant and susceptible carrot genotypes suggested that terpene compounds are potentially involved in fungal defense (Koutouan et al., 2018). In addition, carrots are susceptible to several herbivore pests and therefore could be engineered for enhanced terpene metabolism and stronger pest protection (Guerin and Ryan, 1984; Seljasen et al., 2001; Seljasen et al., 2013). Overall this work provides a strong resource for the research community and for future breeding and metabolic engineering efforts.

REFERENCES

Aharoni, A., Jongsma, M.A., and Bouwmeester, H.J. (2005). Volatile science? Metabolic engineering of terpenoids in plants. *Trends Plant Sci.* 10**,** 594-602. doi: 10.1016/j.tplants.2005.10.005.

Alasalvar, C., Grigor, J.M., Zhang, D.L., Quantick, P.C., and Shahidi, F. (2001). Comparison of volatiles, phenolics, sugars, antioxidant vitamins, and sensory quality of different colored carrot varieties. *J. Agric. Food Chem.* 49**,** 1410-1416. doi: 10.1021/jf000595h.

Alegria, C., Pinheiro, J., Duthoit, M., Goncalves, E.M., Moldao-Martins, M., and Abreu, M. (2012). Fresh-cut carrot (cv. Nantes) quality as affected by abiotic stress (heat shock and UV-C irradiation) pre-treatments. *Food Sci. Technol. Int.* 48**,** 197-203. doi: 10.1016/j.lwt.2012.03.013.

Ayoub, A., Venditti, R.A., Pawlak, J.J., Sadeghifar, H., and Salam, A. (2013). Development of an acetylation reaction of switchgrass hemicellulose in ionic liquid without catalyst. *Crop Improv.* 44**,** 306-314. doi: 10.1016/j.indcrop.2012.10.036.

Badri, D.V., and Vivanco, J.M. (2009). Regulation and function of root exudates. *Plant Cell* 32**,** 666-681. doi: 10.1111/j.1365-3040.2009.01926.x.

Block, A.K., Vaughan, M.M., Schmelz, E.A., and Christensen, S.A. (2019). Biosynthesis and function of terpenoid defense compounds in maize (*Zea mays*). *Planta* 249**,** 21-30. doi: 10.1007/s00425-018-2999-2.

Buttery, R.G., Seifert, R.M., Guadagni, D.G., Black, D.R., and Ling, L.C. (1968). Characterization of some volatile constituents of carrots. *J. Agric. Food Chem.* 16**,** 1009-&. doi: 10.1021/jf60160a012.

Casler, M.D., Vogel, K.P., Taliaferro, C.M., Ehlke, N.J., Berdahl, J.D., Brummer, E.C., Kallenbach, R.L., West, C.P., and Mitchell, R.B. (2007). Latitudinal and longitudinal adaptation of switchgrass populations. *Crop Sci.* 47**,** 2249-2260. doi: 10.2135/cropsci2006.12.0780.

Chen, F., Ro, D.K., Petri, J., Gershenzon, J., Bohlmann, J., Pichersky, E., and Tholl, D. (2004). Characterization of a root-specific *Arabidopsis* terpene synthase responsible for the formation of the volatile monoterpene 1,8-cineole. *Plant Physiol.* 135**,** 1956-1966. doi: 10.1104/pp.104.044388.

Chen, X.J., Chen, H., Yuan, J.S., Köllner, T.G., Chen, Y.Y., Guo, Y.F., Zhuang, X.F., Chen, X.L., Zhang, Y.J., Fu, J.Y., Nebenfuhr, A., Guo, Z.J., and Chen, F. (2018). The rice terpene synthase gene OsTPS19 functions as an (*S*)-limonene synthase in planta, and its overexpression leads to enhanced resistance to the blast fungus *Magnaporthe oryzae*. *Plant Biotechnol. J.* 16**,** 1778-1787. doi: 10.1111/pbi.12914.

Croteau, R., Elbialy, H., and Dehal, S.S. (1987). Metabolic fate of (+)-camphor in sage *Plant Physiol.* 84**,** 643-648. doi: 10.1104/pp.84.3.643.

Del Giudice, L., Massardo, D.R., Pontieri, P., Bertea, C.M., Mombello, D., Carata, E., Tredici, S.M., Tala, A., Mucciarelli, M., Groudeva, V.I., De Stefano, M., Vigliotta, G., Maffei, M.E., and Alifano, P. (2008). The microbial community of *Vetiver* root and its involvement into essential oil biogenesis. *Environ. Microbiol.* 10**,** 2824-2841. doi: 10.1111/j.1462-2920.2008.01703.x.

Demers, J.E., Liu, M., Hambleton, S., and Castlebury, L.A. (2017). Rust fungi on *Panicum*. *Mycologia* 109**,** 1-17. doi: 10.1080/00275514.2016.1262656.

Doll-Boscardin, P.M., Sartoratto, A., Maia, B.H.L.D.S., de Paula, J.P., Nakashima, T., Farago, P.V., and Kanunfre, C.C. (2012). In vitro cytotoxic potential of essential oils of Eucalyptus benthamii and its related terpenes on tumorcell lines. *Evid. Bas. Men. Hea.* doi: 10.1155/2012/342652.

Donze-Reiner, T., Palmer, N.A., Scully, E.D., Prochaska, T.J., Koch, K.G., Heng-Moss, T., Bradshaw, J.D., Twigg, P., Amundsen, K., Sattler, S.E., and Sarath, G. (2017).

Transcriptional analysis of defense mechanisms in upland tetraploid switchgrass to greenbugs. *BMC Plant Biol.* 17. doi: 10.1186/s12870-017-0998-2.

Falara, V., Akhtar, T.A., Nguyen, T.T.H., Spyropoulou, E.A., Bleeker, P.M., Schauvinhold, I., Matsuba, Y., Bonini, M.E., Schilmiller, A.L., Last, R.L., Schuurink, R.C., and Pichersky, E. (2011). The tomato terpene synthase gene family. *Plant Physiol.* 157**,** 770-789. doi: 10.1104/pp.111.179648.

Gascuel, Q., Diretto, G., Monforte, A.J., Fortes, A.M., and Granell, A. (2017). Use of natural diversity and biotechnology to increase the quality and nutritional content of tomato and grape. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00652.

Gershenzon, J. (1994). Metabolic costs of terpenoid accumulation in higher plants. *J. Chem. Ecol.* 20**,** 1281-1328. doi: 10.1007/Bf02059810.

Ghimire, S.R., and Craven, K.D. (2011). Enhancement of switchgrass (Panicum virgatum L.) biomass production under drought conditions by the ectomycorrhizal fungus Sebacina vermifera. *Appl. Environ. Microbiol.* 77**,** 7063-7067. doi: 10.1128/Aem.05225-11.

Guerin, P.M., and Ryan, M.F. (1984). Relationship between root volatiles of some carrot cultivars and their resistance to the carrot fly *Psila rosae*. *Entomol. Exp. Appl.* 36**,** 217-224. doi: 10.1111/j.1570-7458.1984.tb03431.x.

Hejl, A.M., and Koster, K.L. (2004). Juglone disrupts root plasma membrane ATPase activity and impairs water uptake, root respiration, and growth in soybean (*Glycine max*) and corn (*Zea mays*). *J. Chem. Ecol.* 30**,** 453-471. doi: 10.1023/B:JOEC.0000017988.20530.d5.

Horiuchi, J.I., Muroi, A., Takabayashi, J., and Nishioka, T. (2007). Exposing Arabidopsis seedlings to borneol and bornyl acetate affects root growth: Specificity due to the chemical and optical structures of the compounds. *J. Plant Inter.* 2**,** 101-104. doi: 10.1080/17429140701575624.

Huffaker, A., Kaplan, F., Vaughan, M.M., Dafoe, N.J., Ni, X.Z., Rocca, J.R., Alborn, H.T., Teal, P.E.A., and Schmelz, E.A. (2011). Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiol.* 156**,** 2082-2097. doi: 10.1104/pp.111.179457.

Iorizzo, M., Ellison, S., Senalik, D., Zeng, P., Satapoomin, P., Huang, J.Y., Bowman, M., Iovene, M., Sanseverino, W., Cavagnaro, P., Yildiz, M., Macko-Podgorni, A., Moranska, E., Grzebelus, E., Grzebelus, D., Ashrafi, H., Zheng, Z.J., Cheng, S.F., Spooner, D., Van Deynze, A., and Simon, P. (2016). A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nat. Genet.* 48**,** doi: 10.1038/ng.3565.

Jaeger, C.H., Lindow, S.E., Miller, S., Clark, E., and Firestone, M.K. (1999). Mapping of sugar and amino acid availability in soil around roots with bacterial sensors of sucrose and Tryptophan. *Appl. Environ. Microbiol.* 65**,** 2685-2690.

Jaouannet, M., Rodriguez, P.A., Thorpe, P., Lenoir, C.J.G., MacLeod, R., Escudero-Martinez, C., and Bos, J.I.B. (2014). Plant immunity in plant-aphid interactions. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00663.

Jayaraj, J., Rahman, M., Wan, A., and Punja, Z.K. (2009). Enhanced resistance to foliar fungal pathogens in carrot by application of elicitors. *Ann. Appl. Biol.* 155**,** 71-80. doi: 10.1111/j.1744-7348.2009.00321.x.

Keilwagen, J., Lehnert, H., Berner, T., Budahn, H., Nothnagel, T., Ulrich, D., and Dunemann, F. (2017). The terpene synthase gene family of carrot (*Daucus carota L.*). *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01930.

Kjeldsen, F., Christensen, L.P., and Edelenbos, M. (2001). Quantitative analysis of aroma compounds in carrot (*Daucus carota L.*) cultivars by capillary gas chromatography using large-volume injection technique. *J. Agric. Food Chem.* 49**,** 4342-4348. doi: 10.1021/jf010213n.

Kleinheinz, G.T., Bagley, S.T., St John, W.P., Rughani, J.R., and McGinnis, G.D. (1999). Characterization of alpha-pinene-degrading microorganisms and application to a bench-scale biofiltration system for VOC degradation. *Arch. Environ. Contam. Toxicol.* 37**,** 151-157. doi: 10.1007/s002449900500.

Köllner, T.G., Schnee, C., Li, S., Svatos, A., Schneider, B., Gershenzon, J., and Degenhardt, J. (2008). Protonation of a neutral (*S*)-beta-bisabolene intermediate is involved in (*S*)-beta-macrocarpene formation by the maize sesquiterpene synthases TPS6 and TPS11. *J. Biol. Chem.* 283**,** 20779-20788. doi: 10.1074/jbc.M802682200.

Koutouan, C., Le Clerc, V., Baltenweck, R., Claudel, P., Halter, D., Hugueney, P., Hamama, L., Suel, A., Huet, S., Merlet, M.H.B., and Briard, M. (2018). Link between carrot leaf secondary metabolites and resistance to *Alternaria dauci*. *Sci. Rep.* 8. doi: 10.1038/s41598-018-31700-2.

Kreutzmann, S., Thybo, A.K., Edelenbos, M., and Christensen, L.P. (2008). The role of volatile compounds on aroma and flavour perception in coloured raw carrot genotypes. *Int. J. Food Sci. Technol.* 43**,** 1619-1627. doi: 10.1111/j.1365-2621.2007.01662.x.

Kunert, G., Reinhold, C., and Gershenzon, J. (2010). Constitutive emission of the aphid alarm pheromone, (*E*)-β-farnesene, from plants does not serve as a direct defense against aphids. *BMC Ecol.* 10**,** 23. doi: 10.1186/1472-6785-10-23.

Lackus, N.D., Lackner, S., Gershenzon, J., Unsicker, S.B., and Köllner, T.G. (2018). The occurrence and formation of monoterpenes in herbivore-damaged poplar roots. *Sci. Rep.* 8. doi: 10.1038/s41598-018-36302-6.

Lee, S.T., Mitchell, R.B., Wang, Z., Heiss, C., Gardner, D.R., and Azadi, P. (2009). Isolation, characterization, and quantification of steroidal saponins in switchgrass (*Panicum virgatum L.*). *J. Agric. Food Chem.* 57**,** 2599-2604. doi: 10.1021/jf803907y.

Loreto, F., and Schnitzler, J.P. (2010). Abiotic stresses and induced BVOCs. *Trends Plant Sci.* 15**,** 154-166. doi: 10.1016/j.tplants.2009.12.006.

Meyer, E., Aspinwall, M.J., Lowry, D.B., Palacio-Mejia, J.D., Logan, T.L., Fay, P.A., and Juenger, T.E. (2014). Integrating transcriptional, metabolomic, and physiological responses to drought stress and recovery in switchgrass (*Panicum virgatum L.*). *BMC Genomics* 15**,** 527. doi: 10.1186/1471-2164-15-527.

Papadopoulou, K., Melton, R.E., Leggett, M., Daniels, M.J., and Osbourn, A.E. (1999). Compromised disease resistance in saponin-deficient plants. *Proc. Natl. Acad. Sci. U.S.A.* 96**,** 12923-12928. doi:10.1073/pnas.96.22.12923.

Parrish, D.J., and Fike, J.H. (2005). The biology and agronomy of switchgrass for biofuels. *Crit. Rev. Plant Sci.* 24**,** 423-459. doi: 10.1080/07352680500316433.

Pelot, K.A., Chen, R., Hagelthorn, D.M., Young, C.A., Addison, J.B., Muchlinski, A., Tholl, D., and Zerbe, P. (2018). Functional diversity of diterpene synthases in the biofuel crop switchgrass. *Plant Physiol.* 178**,** 54-71. doi: 10.1104/pp.18.00590.

Prasifka, J.R., Bradshaw, J.D., Lee, S.T., and Gray, M.E. (2011). Relative feeding and development of armyworm on switchgrass and Corn. *J. Econ. Entomol.* 104**,** 1561-1567. doi: 10.1603/Ec10304.

Rasmann, S., Köllner, T.G., Degenhardt, J., Hiltpold, I., Toepfer, S., Kuhlmann, U., Gershenzon, J., and Turlings, T.C.J. (2005). Recruitment of entomopathogenic nematodes by insect-damaged maize roots. *Nature* 434**,** 732-737. doi: 10.1038/nature03451.

Riedlmeier, M., Ghirardo, A., Wenig, M., Knappe, C., Koch, K., Georgii, E., Dey, S., Parker, J.E., Schnitzler, J.P., and Vlot, A.C. (2017). Monoterpenes support systemic acquired

resistance within and between plants. *Plant Cell* 29**,** 1440-1459. doi: 10.1105/tpc.16.00898.
Rosenfeld, H.J., Aaby, K., and Lea, P. (2002). Influence of temperature and plant density on sensory quality and volatile terpenoids of carrot (*Daucus carota L.*) root. *J. Sci. Food Agric.* 82**,** 1384-1390. doi: 10.1002/jsfa.1200.

Rosenfeld, H.J., Vogt, G., Aaby, K., and Olsen, E. (2004). Interaction of terpenes with sweet taste in carrots (*Daucus carota L.*). *Ad. Veg. Breed.* 377-386. doi: 10.17660/ActaHortic.2004.637.47.

Roy, H.E., Pell, J.K., and Alderson, P.G. (1999). Effects of fungal infection on the alarm response of pea aphids. *J. Invertebr. Pathol.* 74**,** 69-75. doi: 10.1006/jipa.1999.4856.

Rudrappa, T., Czymmek, K.J., Pare, P.W., and Bais, H.P. (2008). Root secreted malic acid recruits beneficial soil bacteria. *Plant Physiol.* 148**,** 1547-1556. doi: 10.1104/pp.108.127613.
Scervino, J.M., Ponce, M.A., Erra-Bassells, R., Vierheilig, H., Ocampo, J.A., and Godeas, A. (2005). Flavonoids exhibit fungal species and genus specific effects on the presymbiotic growth of *Gigaspora* and *Glomus*. *Mycol. Res.* 109**,** 789-794. doi: 10.1017/S0953756205002881.

Schuh, M.K., Bahlai, C.A., Malmstrom, C.M., and Landis, D.A. (2019). Effect of switchgrass ecotype and cultivar on establishment, feeding, and development of fall armyworm *J. Econ. Entomol.* 112**,** 440-449. doi: 10.1093/jee/toy292.

Seljasen, R., Hoftun, H., and Bengtsson, G.B. (2001). Sensory quality of ethylene-exposed carrots (*Daucus carota L*, cv 'Yukon') related to the contents of 6-methoxymellein, terpenes and sugars. *J. Sci. Food Agric.* 81**,** 54-61. doi: 10.1002/1097-0010(20010101)81.

Seljasen, R., Vogt, G., Olsen, E., Lea, P., Hogetveit, L.A., Taje, T., Meadow, R., and Bengtsson, G.B. (2013). Influence of field attack by carrot psyllid on sensory quality, antioxidancapacity and content of terpenes of carrots (*Daucus carota L.*). *J. Agric. Food Chem.* 61**,** 2831-2838. doi: 10.1021/jf303979y.

Sharma, K.D., Karki, S., Thakur, N.S., and Attri, S. (2012). Chemical composition, functional properties and processing of carrot-a review. *J. Food Sci. Tech.* 49**,** 22-32. doi: 10.1007/s13197-011-0310-7.

Simon, P.W. (2010). Domestication, historical development, and modern breeding of carrot. *Plant Breed. Rev.* 19**,** 157 - 190. doi: 10.1002/9780470650172.ch5.

Simon, P.W., Peterson, C.E., and Lindsay, R.C. (1980). Correlations between sensory and objective parameters of carrot flavor. *J. Agric. Food Chem.* 28**,** 559-562. doi: 10.1021/jf60229a041.

Stojanovic, G., Palic, I., and Ursic-Jankovic, J. (2006). Composition and antimicrobial activity of the essential oil of Micromeria cristata and Micromeria juliana. *Flavour Fragrance J.* 21**,** 77-79. doi: 10.1002/ffju.1507.

Sykes, V.R., Allen, F.L., Mielenz, J.R., Stewart, C.N., Windham, M.T., Hamilton, C.Y., Rodriguez, M., and Yee, K.L. (2016). Reduction of ethanol yield from switchgrass infected with rust caused by Puccinia emaculata. *Bioenergy Research* 9**,** 239-247. doi: 10.1007/s12155-015-9680-4.

Tabanca, N., Kirimer, N., Demirci, B., Demirci, F., and Baser, K.H.C. (2001). Composition and antimicrobial activity of the essential oils of *Micromeria cristata subsp phrygia* and the

enantiomeric distribution of borneol. *J. Agric. Food Chem.* 49**,** 4300-4303. doi: 10.1021/jf0105034.

Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P., and Osbourn, A. (2014). Triterpene biosynthesis in plants. *Annu. Rev. Plant Biol.* 65**,** 225-257. doi: 10.1146/annurev-arplant-050312-120229.

Tholl, D., Chen, F., Gershenzon, J., and Pichersky, F. (2004). Arabidopsis thaliana, a model system for investigating volatile terpene biosynthesis, regulation, and function. *Rec. Adv. Phyto.* 38**,** 1-18.

Tian, N., Liu, S., Li, J., Xu, W., Yuan, L., Huang, J., and Liu, Z. (2014). Metabolic analysis of the increased adventitious rooting mutant of Artemisia annua reveals a role for the plant monoterpene borneol in adventitious root formation. *Physiol. Plant.* 151**,** 522-532. doi: 10.1111/ppl.12139.

Tsang, H.L., Huang, J.L., Lin, Y.H., Huang, K.F., Lu, P.L., Lin, G.H., Khine, A.A., Hu, A., and Chen, H.P. (2016). Borneol dehydrogenase from Pseudomonas catalyzes the oxidation of (+)-borneol *Appl. Environ. Microbiol.* 82**,** 6378-6385. doi: 10.1128/Aem.01789-16.

Turlings, T.C.J., Tumlinson, J.H., and Lewis, W.J. (1990). Exploitation of herbivore-induced plant odors by host-seeking parasitic wasps. *Science* 250**,** 1251-1253. doi: 10.1126/science.250.4985.1251.

Umehara, M., Cao, M.M., Akiyama, K., Akatsu, T., Seto, Y., Hanada, A., Li, W.Q., Takeda-Kamiya, N., Morimoto, Y., and Yamaguchi, S. (2015). Structural requirements of strigolactones for shoot branching inhibition in rice and *Arabidopsis*. *Plant and Cell Physiology* 56**,** 1059-1072. doi: 10.1093/pcp/pcv028.

Vaughan, M.M., Christensen, S., Schmelz, E.A., Huffaker, A., Mcauslane, H.J., Alborn, H.T., Romero, M., Allen, L.H., and Teal, P.E.A. (2015). Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance. *Plant Cell and Environment* 38**,** 2195-2207. doi: 10.1111/pce.12482.

Vidal, A., Hirte, J., Bender, S.F., Mayer, J., Gattinger, A., Hoschen, C., Schadler, S., Iqbal, T.M., and Mueller, C.W. (2018). Linking 3D soil structure and plant-microbe-soil carbon transfer in the rhizosphere. *Frontiers in Environmental Science* 6. doi: 10.3389/fenvs.2018.00009.

Xu, L., Naylor, D., Dong, Z.B., Simmons, T., Pierroz, G., Hixson, K.K., Kim, Y.M., Zink, E.M., Engbrecht, K.M., Wang, Y., Gao, C., DeGraaf, S., Madera, M.A., Sievert, J.A., Hollingsworth, J., Birdseye, D., Scheller, H.V., Hutmacher, R., Dahlberg, J., Jansson, C., Taylor, J.W., Lemaux, P.G., and Coleman-Derr, D. (2018). Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria *Proc. Natl. Acad. Sci. U.S.A.* 115**,** E4952-E4952. doi: 10.1073/pnas.1807275115.

Yahyaa, M., Tholl, D., Cormier, G., Jensen, R., Simon, P.W., and Ibdah, M. (2015). Identification and characterization of terpene synthases potentially involved in the formation of volatile terpenes in carrot (*Daucus carota L.*) roots. *J. Agric. Food Chem.* 63**,** 4870-4878. doi: 10.1021/acs.jafc.5b00546.

Yang, J., Kloepper, J.W., and Ryu, C.M. (2009). Rhizosphere bacteria help plants tolerate abiotic stress. *Trends Plant Sci.* 14**,** 1-4. doi: 10.1016/j.tplants.2008.10.004.

Ye, Z.J., Sangireddy, S., Okekeogbu, I., Zhou, S.P., Yu, C.L., Hui, D.F., Howe, K.J., Fish, T., and Thannhauser, T.W. (2016). Drought-induced leaf proteome changes in switchgrass seedlings. *Interl J Molec Scs* 17. doi: 10.3390/ijms17081251.

Zhuang, X.F., Köllner, T.G., Zhao, N., Li, G.L., Jiang, Y.F., Zhu, L.C., Ma, J.X., Degenhardt, J., and Chen, F. (2012). Dynamic evolution of herbivore-induced sesquiterpene biosynthesis in *Sorghum* and related grass crops. *Plant J.* 69**,** 70-80. doi: 10.1111/j.1365-313X.2011.04771.x.

APPENDIX A: EFFECTS OF DROUGHT STRESS ON VOLATILE TERPENOID PRODUCTION IN SWITCHGRASS ROOTS

Contributions:
Andrew Muchlinski (VT): Designed and conducted drought stress experiment, data analysis, hexane and methanol extractions of root and leaves, metabolite profiling and analysis by LC- and GC/MS, RNA extraction from roots and leaves; Yiming Liu (Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, China): conducted experimental design, physiological measurements and implementation of drought stress treatments; Sherry Hildreth, Rich Helm (VT): provided training for non-volatile metabolite analysis; Dorothea Tholl (VT): performed experimental design and metabolite analysis; Erik Nilsen, Bingyu Zhao (VT): conducted experimental design; Adrian Peña: performed leaf volatile metabolite extractions and analysis; Eva Collakova (VT): provided training for principle component analysis; Katherine Berg (VT- undergraduate researcher): performed RNA and hexane extractions; Mary Ellen Gill, David Brodsky, Susan Boohaker, Hanna Holt (VT undergraduate researchers): performed physiological measurements and sample collection; Avinash Sreedasyam, Jeremy Schmutz, Yasuo Yoshikuni, Kerrie Barry (JGI): conducted transcriptome assembly and analysis.

OVERVIEW

In conjunction with Chapter II, we sought to examine ecotype-specific differences of volatile and non-volatile specialized metabolites in switchgrass roots under no stress and drought stress conditions with a focus on terpenoids. There is limited knowledge for the role of terpenoids in drought stress response; but, there is evidence for the function of diterpenes in maize in drought stress tolerance. The goal was to better understand the contribution of terpenoids to abiotic stress responses in roots of switchgrass. This evaluation has been done by investigating changes in terpene abundance and biosynthesis in leaves and roots by a targeted analysis of terpene

specialized metabolites in conjunction with a broader metabolomics analysis and by obtaining transcriptome data under control and drought stress conditions.

RESULTS AND DISCUSSION

Effects of drought stress on switchgrass physiology

To determine changes of terpenoid metabolite profiles under drought stress conditions, we first compared physiological responses of two ecotypes of switchgrass with documented differences in drought stress tolerance (Liu et al., 2015). To assess drought tolerance, a drought tolerance index (DTI) was used for evaluating physiological traits. DTI was calculated using the formula: DTI = (value of trait under stress condition)/(value of trait under controlled condition) × 100 (Liu et al., 2015). We found that switchgrass cv. Alamo (lowland) was more tolerant to drought stress based on maintaining a higher DTI for conductivity and transpiration rate at 20% soil water content (SWC) compared to cv. Dacotah (upland ecotype). However, the DTI for photosynthetic rate was higher for Dacotah suggesting both cultivars maybe have some resilience to drought based similar to findings reported by Liu et al. (2015).

Compared to well-watered plants used as controls, drought-stress plants of both ecotypes showed significantly decreased photosynthetic rate (Fig. 2; Table 1). By day 13, drought stressed Alamo plants had a significantly lower photosynthetic rate compared to well-watered Alamo and Dacotah control plants. Interestingly, drought stressed Dacotah plants did not have significantly lower photosynthetic rate at day 13 compared to Dacotah control plants, further suggesting Dacotah may exhibit drought tolerance at that stage of drought stress (Table 1). Re-watered plants (Day 22) showed no significant differences in photosynthetic rate compared to controls (Table 1).

Soil water content and leaf water content were also found to decrease as a result of withholding water (Figs. 3 and 4).

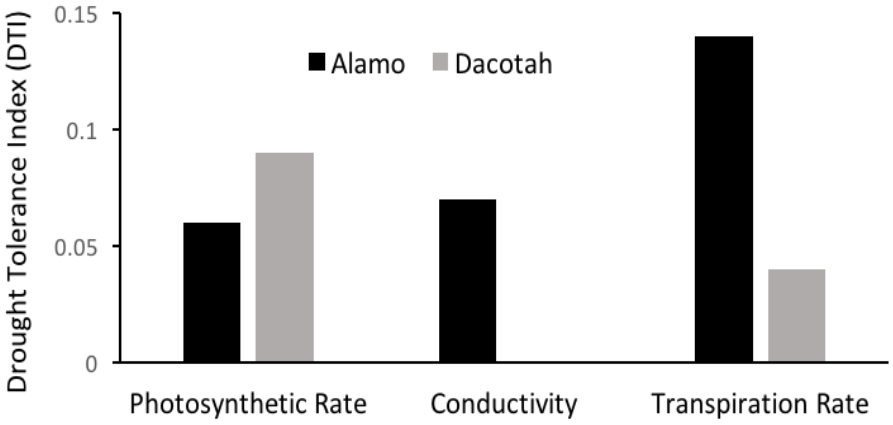


Figure 1. Drought tolerance index (DTI) of photosynthetic rate, conductivity, and transpiration rate in Alamo and Dacotah at relative soil water content = 20%. DTI was calculated by dividing the average ratio of the physiological traits from drought stressed plants to those of control samples. n=4.

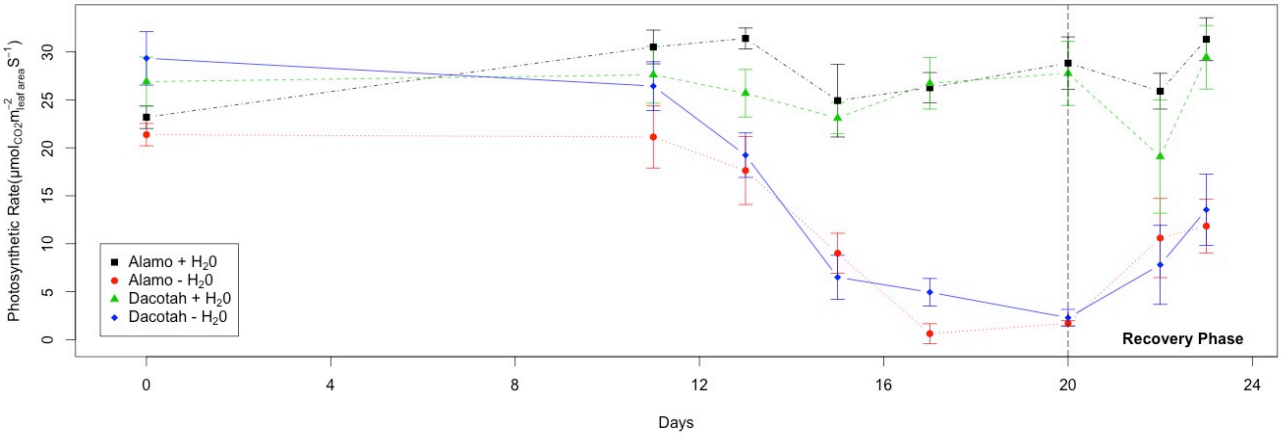


Figure 2. Analysis of photosynthetic rate in leaves of switchgrass cultivars Alamo and Dacotah measured using a LICOR-6400 (LICOR Inc, Lincoln, NE, USA). Plants were grown from tillers and allowed to establish in a soil/sand (50/50) substrate for 11 days prior to application of drought treatment. Beginning on day 11, soil water content was gradually reduced to 20%, held for a period of 3 days, and then re-watered on day 20. Statistical analysis and figure generation were conducted in R (The R Foundation for Statistical Computing) where significant differences in drought versus well-watered treatments were considered if $\alpha \leq 0.05$.

Table 1. Statistical analysis of drought treatment effects on photosynthetic carbon assimilation across all time points. Relative differences were calculated using analysis of variance (ANOVA) and exact differences were calculated using post-hoc Tukey-Kramer HSD tests in R (The R Foundation for Statistical Computing). Results from the ANOVA and post-hoc comparisons are reported as adjusted p-values where asterisks indicate significant values ≤ 0.05.

Pairwise comparisons	Day 0 (ANOVA)	Day 11 (ANOVA)	Day 13 (adj. p-value)	Day 15 (adj. p-value)	Day 17 (adj. p-value)	Day 20 (adj. p-value)	Day 22 (ANOVA)
Alamo + H_20 : Alamo - H_20	0.74434	0.0524	0.0053*	0.0022*	< 0.001*	< 0.001*	0.2212
Alamo + H_20 : Dacotah + H_20	0.74434	0.0524	0.5639	0.9958	0.9999	0.9993	0.2212
Alamo + H_20 : Dacotah - H_20	0.74434	0.0524	0.0104*	0.0003*	< 0.001*	< 0.001*	0.2212
Alamo - H_20 : Dacotah - H_20	0.74434	0.0524	0.5963	0.9958	0.9882	0.9997	0.2212
Dacotah + H_20 : Dacotah - H_20	0.74434	0.0524	0.2713	0.0010*	< 0.001*	< 0.001*	0.2212

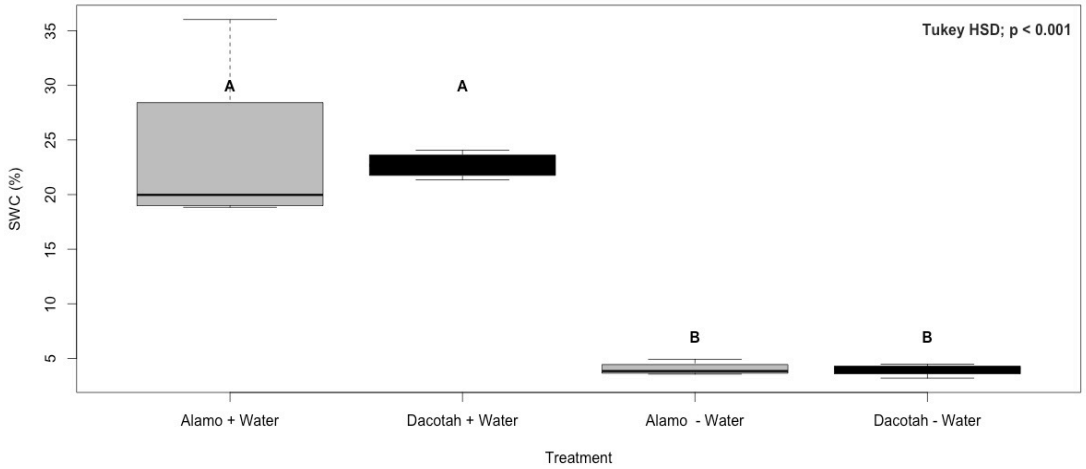


Figure 3. Absolute soil water content (SWC). Values were calculated using the following formula: $SWC(\%) = \left[\frac{W-DW}{DW}\right] x100$ where W= fresh weight and DW= dry weight. Absolute SWC for well-watered plants represents full water capacity.

Volatile terpenoids accumulate in switchgrass roots during drought stress

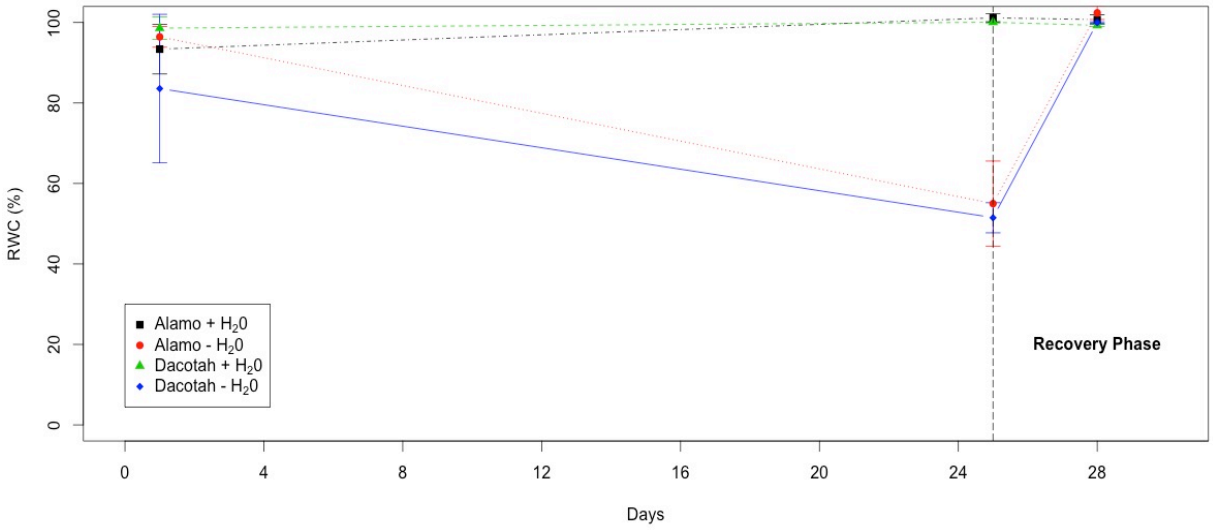


Figure 4. Leaf relative water content (RWC). Values were calculated using the following formula: $RWC(\%) = \left[\frac{W-DW}{TW-DW}\right] x100$ where W= fresh weight, T= turgor weight and DW = dry weight.

To determine changes in levels of volatiles as a result of drought stress, we generated chemical profiles of replicate control and drought stressed root samples for both ecotypes (see Methods). We found several putative oxygenated sesquiterpenoids in roots of drought-treated Alamo and Dacotah plants as well as a large number of putative oxygenated diterpenoid compounds in drought-treated roots of Dacotah (Fig. 5).

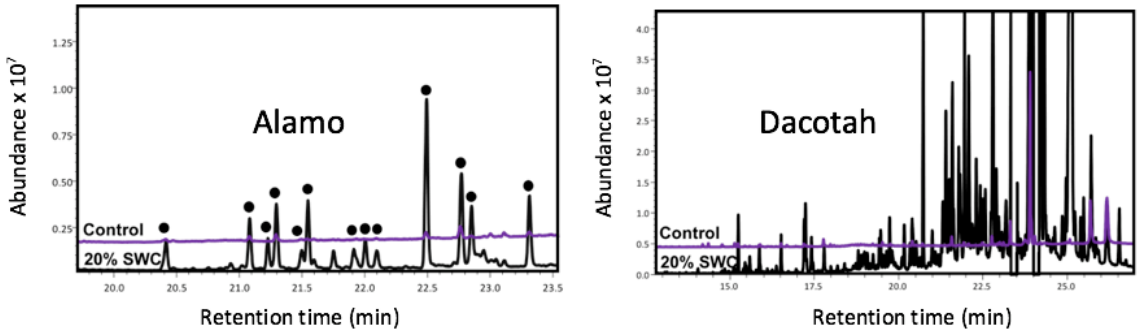


Figure 5. Representative GC/MS profiles of volatile compounds in roots of Alamo and Dacotah plants under drought and control conditions. Shaded circles indicate putative oxygenated sesquiterpene compounds. Peaks were not identified for Dacotah due to the large number of induced putative oxygenated sesquiterpenoid and diterpenoid compounds.

Volatile metabolites show differential responses in roots as compared to leaves during drought stress

Alamo and Dacotah showed an increased concentration of root volatiles during drought stress. To better understand whether these changes are specific to the belowground tissues, we compared volatile profiles of leaves and roots of both ecotypes under drought conditions using principle component analysis (PCA). Hexane extracts from four biological replicates from each tissue (root/shoot) and ecotype (Alamo/Dacotah) were analyzed by GC/MS (see Methods). The resulting peaks were integrated, normalized to gram fresh weight, and putatively annotated based on major compound category (terpenoid, phenolic, fatty acid derived etc.). Data were then filtered based on consistency among replicates as well as tissue and ecotype. The resulting 39 metabolites that were supported consistently among at least three replicates were used for PCA (Fig. 6).

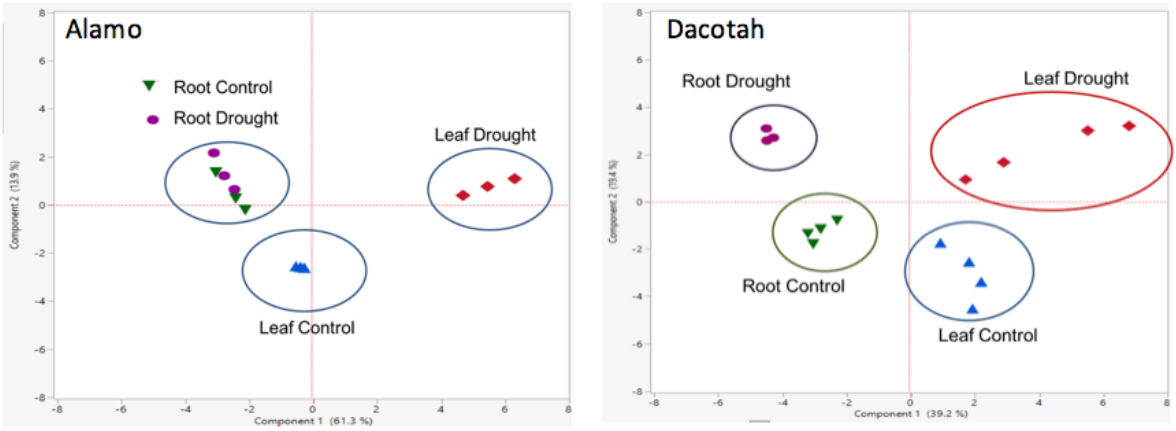


Figure 6. Principle component analysis of volatile metabolites in roots and leaves of Alamo and Dacotah under controlled and drought stress conditions. Metabolites were extracted from four replicate samples following standards protocols. Data analysis was conducted using the JMP statistical package.

In analyzing tissue specific volatile data under well-watered and drought stress conditions, using PCA, we found that leaves of both cultivars responded similarly to drought stress by reducing the production of C6-green leaf volatiles. We also observed an increase in several short chain (6-10 carbon) organic acids in leaves of both cultivars during drought stress potentially suggesting a reallocation of resources to better cope with the stress via production of osmoprotectants. Interestingly, the monoterpene alcohol linalool appeared to be drought-inducible in Alamo leaves, and possibly in Dacotah as well; however, variation in compound levels across samples and treatments did not completely support this. In general, we obtained clearly clustered replicates of leaf samples in the PCA suggesting that both cultivars do respond to drought stress by modifying leaf volatile production. In Alamo roots, we did not observe major shifts in volatile profiles as a result of drought stress using PCA. However, upon comparing GC/MS profiles between control and drought stressed samples, elevated levels of putative oxygenated

sesquiterpenoids could be detected in the roots (Fig. 5). In Dacotah roots, several putative terpenoids including camphor, cycloisosativene and germacrene appeared to accumulate upon drought treatment. In addition, the magnitude of the response (both in number of compounds and relative amounts) was much higher than that observed in Alamo roots (Fig. 5).

Non-volatile primary and secondary metabolites show significant changes in roots during drought stress

To determine more global metabolic changes in drought stress responses of the switchgrass cultivars, we investigated methanolic extracts from root tissue and analyzed samples by UPLC/MS (see Methods). We found that Alamo and Dacotah share 129 features during drought stress, but also have unique ecotype specific features as well (Fig. 7A). Statistical analysis of these features identified 242 features that changed as a results of droughts tress. Specifically, 24 features were shared between ecotypes whereas 97 were specific to Alamo and 121 were specific to Dacotah (Fig. 7B).

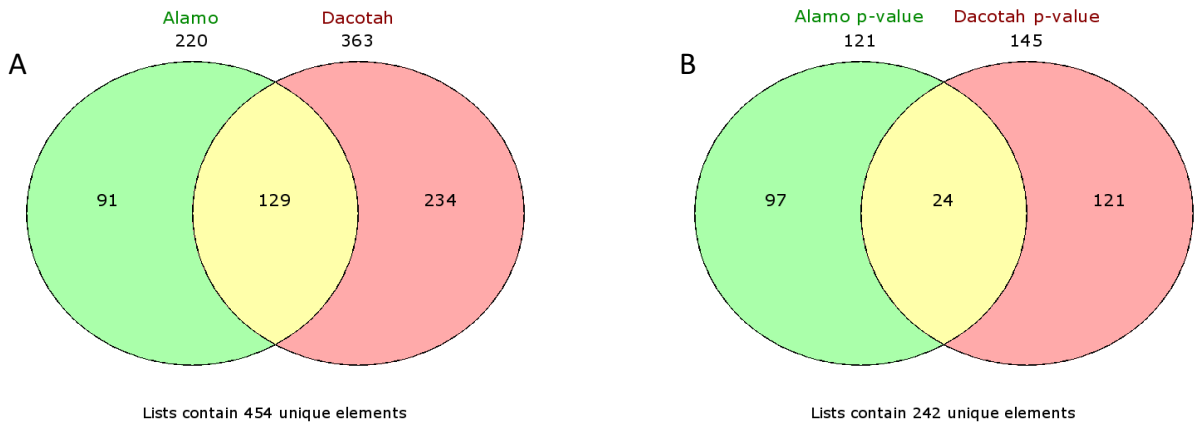


Figure 7. Venn diagrams generated from a list of UPLC/MS features analyzed in positive ion mode (VennPlex). A) Comparison of total features across Alamo and Dacotah during drought stress. B) Comparison of filtered features based on significant deviation from control plants.

TPS gene expression analysis

To identify which terpene synthases (TPSs) contribute to the observed increase in terpene metabolites in roots during drought stress, we isolated RNA from samples described above and performed RNA-seq analysis in collaboration with the DOE Joint Genome Institute. Results comparing expression patterns of TPS genes across samples are reported in Fig. 8 as transcript per million (TPM). We found five genes in roots that had differences in transcript levels between Alamo drought and Alamo control (at 20% SWC); some of these differences remain in the recovery phase and some of these genes also show differential transcript abundance between Dacotah drought and Dacotah control pre-treatment as well as in the recovery phase. Interestingly, these genes were previously characterized to function in diterpenoid biosynthesis. We also found an upregulated diterpene synthase in leaves of Alamo in control versus drought conditions, however this diTPS remains uncharacterized. Although, we did not detect any induced volatile diterpenoids in Dacotah leaves, it is possible that the final product is not volatile and therefore not readily detected by GC/MS. Therefore, further targeted analysis of LC/MS data obtained from Dacotah leaves could be used to detect non-volatile compounds. RNA-seq results should also be further verified by RT-qPCR. Additionally, challenges with root RNA extraction and poor sequencing quality resulted in missing replicate data and unexpected divergence of transcript patterns in Alamo control - drought compared to Alamo control – pre-treatment. Nevertheless, the results suggest drought induced terpenoid biosynthesis in Alamo and Dacotah roots and leaves.

Figure 8. Heat map expression analysis of switchgrass TPS gene expression in Alamo and Dacotah root and leaves. Expression values are reported as transcript per million (TPM).

OUTLOOK

Based on metabolome and transcriptome profiling, we have observed several changes in specialized metabolism, especially terpenoids, in response to drought. We were able to identify induced transcript levels of some TPS genes, which might have a role in abiotic stress resilience in switchgrass roots. Our extensive investigation has resulted in the collection of several data sets that could be of use to the research community. We will therefore evaluate the possibility of releasing the obtained sets as part of a resource article.

METHODS

Plant materials and culture

Switchgrass ecotypes Alamo and Dacotah were kindly provided by Dr. Bingyu Zhao and were propagated from tillers and established under greenhouse conditions at Virginia Tech (Blacksburg, VA) during the summer of 2015. Tillers were planted in plastic pots (17 cm diam., 20 cm high, with four holes at the bottom for drainage) and filled with 3.7 kg of soil collected from an established switchgrass field (Moore Farm, Blacksburg, VA). To improve drainage, sand was mixed into the field soil at the following ratio: (soil: sand = 2:1 v/v, sand: 0.1-1.0 mm diam.). Plants were grown at temperatures of ~30^{o}C/~25^{o}C (day/night) with ~14 h photoperiod, and with photosynthetically active radiation ~500 μmol $m^{-2}s^{-1}$. Plants were irrigated every two days and acclimated for a period of one month. Experimentation began once all plants had reached the E5 developmental stage (Hardin et al., 2013).

Drought stress treatment

To determine the relative soil water content of each pot (SWC), an equation of linear regression between SWC and volumetric soil moisture content (VWC) was first determined. Specifically, soil from 4 randomly selected pots was oven dried at 105°C for 48 h to obtain the dry weight (DW). The soil was then saturated and allowed to sit until no water leached from the pot at which time fresh weight (FW) of each pot was recorded. VWC was obtained using a soil moisture meter (model HH2, Delta-T Devices, Cambridge, England) and averaging triplicate measurements from three random locations in the pot. SWC was calculated according to the following formula: SWC (%) = (FW−DW)/DW×100. Absolute SWC was also determined during plant harvest.

Plants from each ecotype were randomly assigned to a control group (n=4), which was kept well-watered to maintain the soil moisture content at field capacity, or drought treatment group (n=4), in which the SWC was allowed to gradually decline to 20% of field capacity over a 23-day period. Once 20% SWC was reached, the treatment was maintained for a one-week period before re-watering beginning the recovery phase. To ensure equal decline in SWC for individual pots, each pot was weighted every two days and VWC was measured. Using the standard curve and obtained measurements, a specific amount of water was added to equalize all pots to the same percent SWC.

Physiological measurements

Leaf relative water content (RWC) was determined according to the following formula: RWC= (FW-DW)/ (TW-DW) ×100, where FW is leaf fresh weight, DW (dry weight) is the weight of the leaves after drying at 85 °C for 3 d, and TW (turgid weight) is the weight of the leaves after soaking them in distilled water for 24 h at 20°C (Barrs and Weatherley, 1962).

The photosynthetic rate (Pn), stomatal conductance (g_s), and transpiration rate (Tr) were measured using a portable photosynthesis system (Li-6400XT, LI-COR, Inc., Lincoln, NE, USA) under a controlled atmosphere (385 μmol·mol^{-1} CO_2, 500 μmol·s^{-1} flow rate, 26 ^{o}C) and a LI-COR 6400 LED external light source that provided a photosynthetic photon flux density (PPFD) of 2000 μmol·m^{-2}·s^{-1}. The uppermost fully expanded leaf on the main tiller in each pot was selected for these measurements. Three readings were collected for each sample, and the average was used for statistical analysis.

Tissue collection

Samples from each ecotype (Alamo, Dacotah), and from each treatment (control, drought), were collected at three time points: Pre-treatment, Stressed (20% SWC) and Recovery. Root tissue was collected at each time point (n=4) by thoroughly rinsing root tissue with deionized water, drying, and harvesting the last 15 cm of root material as measured from the root tip. Collected tissue was immediately frozen in liquid nitrogen and stored at -80^{o}C for later processing. Leaf tissue was also collected at the "Stressed" time point and reserved for later analysis.

Metabolite extraction and analysis

Total collected tissue was pulverized under liquid nitrogen and mixed with extraction solvent in glass tubes or Eppendorf tubes (5 ml hexane/1 g or 1 ml methanol/150 mg) using a vortex. Samples were then placed in an ultrasonic bath (FisherSci) for 10 min at room temperature then centrifuged at 4^{o}C/12,000xg/15 min to remove insoluble material. The supernatant was transferred to a new vial and extraction was repeated with volumes described above. Pooled

fractions were analyzed by UPLC/MS as described by Jervis et al., 2015. Pooled hexane fractions were analyzed by GC/MS as described by Muchlinski et al., *in review*.

RNA extraction and transcriptome analysis

Total root RNA was extracted from frozen plant material (2 g) as described by Kumar et al. (2007). Following RNA isolation, samples were further removed of contaminants using a Qiagen RNeasy Mini Kit according to the manufactures protocol. Leaf RNA samples were extracted using a Qiagen RNeasy Mini Kit according to the manufactures protocol. RNA-seq was performed at the JGI using Illumina sequencing. Raw fastq file reads were filtered and trimmed using the JGI QC pipeline resulting in the filtered fastq file (*.filter-RNA.gz files). Using BBDuk, raw reads were evaluated for artifact sequence by kmer matching (kmer=25), allowing 1 mismatch and detected artifacts were trimmed from the 3' end of the reads. RNA spike-in reads, PhiX reads and reads containing any poor quality sequence were removed. Quality trimming was performed using the phred trimming method set at Q6. Finally, following trimming, reads under the length threshold were removed (minimum length 25 bases or 1/3 of the original read length - whichever is longer). Filtered reads from each library were aligned to the reference genome using HISAT2 version 2.1.0 (BAMs/ directory). featureCounts was used to generate the raw gene counts (counts.txt) file using gff3 annotations. Only primary hits assigned to the reverse strand were included in the raw gene counts (-s 2 -p --primary options). Raw gene counts were used to evaluate the level of correlation between biological replicates using Pearson's correlation and determine which replicates would be used in the differential gene expression analysis (replicate_analysis.txt, replicate_analysis_heatmap.pdf). In the heatmap view, the libraries were ordered as groups of replicates. DESeq2 (version 1.18.1) was subsequently used to determine which genes were

differentially expressed between pairs of conditions. The parameters used to call a gene DE between conditions were p-value < 0.05.

REFERENCES

Barrs, H.D., and Weatherley, P.E. (1962). A re-examination of relative turgidity technique for estimating water deficits in leaves. *J. Biol. Sci.* 15**,** 413-&. doi: 10.1071/Bi9620413.

Hardin, C.F., Fu, C.X., Hisano, H., Xiao, X.R., Shen, H., Stewart, C.N., Parrott, W., Dixon, R.A., and Wang, Z.Y. (2013). Standardization of switchgrass sample collection for cell wall and biomass trait analysis. *Bio. Res.* 6**,** 755-762. doi: 10.1007/s12155-012-9292-1.

Jervis, J., Hildreth, S.B., Sheng, X.Y., Beers, E.P., Brunner, A.M., and Helm, R.F. (2015). A metabolomic assessment of NAC154 transcription factor overexpression in field grown poplar stem wood. *Phytochemistry* 115**,** 112-120. doi: 10.1016/j.phytochem.2015.02.013.

Kumar, G.N., Iyer, S., and Knowles, N.R. (2007). Extraction of RNA from fresh, frozen, and lyophilized tuber and root tissues. *J. Agric. Food Chem.* 55**,** 1674-1678. doi: 10.1021/jf062941m.

Liu, Y.M., Zhang, X.Z., Tran, H., Shan, L., Kim, J., Childs, K., Ervin, E.H., Frazier, T., and Zhao, B.Y. (2015). Assessment of drought tolerance of 49 switchgrass (*Panicum virgatum*) genotypes using physiological and morphological parameters. *Biotech. Bio.* 8. doi: 10.1186/s13068-015-0342-8.

APPENDIX B: BIOSYNTHESIS OF BORNEOL AND EXAMINATION OF ITS EFFECT ON THE ROOT MICROBIAL COMMUNITY OF *SETARIA VIRIDIS*

Contributions:

Andrew Muchlinski (VT): designed and conducted experiments and performed data analysis. Specifically, AM conducted hexane extractions, SPME analysis of plant tissues, GC/MS, construct design, mutant screening, recombinant protein expression and biochemical characterization, stress treatments, microbial gDNA extraction, RNA extraction, RT-qPCR, sample collection, phylogeny construction, gene expression analysis; Feng Chen (UT Knoxville): performed gene identification; Joyce Van Eck, Kaitlin Pidgeon (Cornell): conducted plant transformation and regeneration; Frank Aylward, Alaina Weinheimer (VT): performed 16S rRNA data analysis; Dorothea Tholl (VT): was responsible for experimental design and manuscript preparation.

In conjunction with Chapter III, we extended our analysis of the role of borneol in root-microbe interactions to the model grass *Setaria viridis. Setaria* is a diploid close relative of switchgrass, which is more amenable to transformation. *S. viridis* also accumulates borneol in root tissues, although at much lower levels. The goal of this study was to identify a borneol synthase in *S. viridis* and to use an RNAi based approach to demonstrate that borneol is a host-specific factor affecting microbial community composition in the root endophytic compartment.

ABSTRACT

Plant roots host a diverse assemblage of microorganisms. This microbial community, or "microbiome", is considered beneficial in protecting plants against different environmental stressors. Specific plant traits that define the composition of microbial communities by "habitat" filtering remain poorly understood. As microbial community composition can be plant organ or tissue specific, mechanisms for selection of particular microbes to niches throughout the plant must

exist. Plant derived volatile organic compounds (VOCs), such as terpenoids, are known to affect colonization of aboveground tissues such as flowers by epiphytic bacteria. However, knowledge of the role of VOCs in plant-microbe interactions belowground is limited. We found that the model grass *Setaria viridis* constitutively produces the common volatile 10-carbon monoterpenoid (-)-borneol in roots. By comparing terpene synthase genes (*TPS*) from different grasses, we identified and biochemically characterized the root expressed (-)-borneol synthase *Sv*TPS04. Downregulation of *SvTPS04* in RNAi lines resulted in a depletion of borneol in roots compared to wild-type and empty vector lines. Analysis of endosphere and rhizosphere bacterial communities by 16S rRNA sequencing revealed distinct communities exist in these root zones consistent with findings from rice and maize. However, lack of (-)-borneol had little effect on the microbiome within the endosphere or rhizosphere of *S. viridis* roots suggesting alternative roles of this monoterpenoid in planta. Results from this study serve as a first major step in gaining a better understanding of the biological function of borneol in roots of *Setaria* and perhaps other plants.

RESULTS AND DISCUSSION

Borneol is constitutively released from detached root tissues of *S. viridis*

Based on previous findings that roots of different cultivars of switchgrass accumulate high levels of the monoterpenoid borneol (see Chapter III), we investigated whether this terpenoid also occurs in roots of the close switchgrass relative *S. viridis*. We found that borneol was constitutively released from detached *S. viridis* roots but could not be detected in any other tissue (panicle, culm, leaf; Fig. 1). In addition, borneol accumulation could not be stimulated or inhibited by treatment with the plant phytohormones methyl jasmonate (1 mM), salicylic acid (1 mM), and abscisic acid (1 mM) or by drought stress (14 days) and treatment with the oxidative stress inducer copper sulfate (10 mM) (Supplemental Fig. S1).

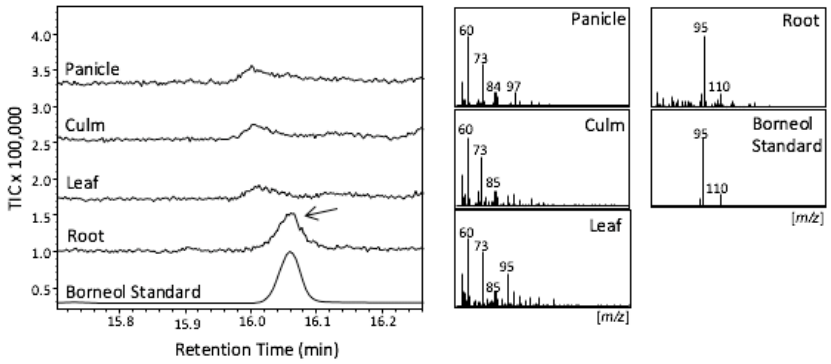


Figure 1. SPME-GC-MS analysis of borneol release from 1 g of whole detached *Setaria* tissues. Samples were analyzed in biological triplicate and compared to an authentic borneol standard.

Identification of borneol synthases genes in *Setaria*

To determine the biosynthetic origin of borneol production in *Setaria*, we first compiled a catalog of 225 putative terpene synthases from related Poaceae grasses using publically available genomic data (Phytozome; see Materials and Methods). Of these, only one is known to function as a borneol synthase; which is found in the closely related model bioenergy grass switchgrass (*PvTPS04*, Muchlinski *et al.*, *in review*). Multiple alignment using Fast Fourier Transform (MAAFT) and maximum likelihood phylogenic reconstruction revealed a small clade of putative borneol synthases in *Setaria viridis* and *Setaria italica* (Sevir.1G152600; Seita.1G151900) that closely align with the *Pv*TPS04 protein (Fig. 2). Detailed analysis of this clade indicated strong support (>80 %) over 500 bootstrap replications (Supplemental Fig. S2). Amino acid alignment of full-length *Pv*TPS04 with *Setaria* orthologs (see also synteny analysis, Muchlinski *et al.*, *in review*) indicated 88.9% sequence similarity (Supplemental Fig. S3). Both *Setaria* proteins designated *Si*- and *Sv*TPS04, share 99.8% sequence similarly and a putative 21 amino acid transit

peptide (Supplemental Fig. S3). We decided to focus our study on *TPS04* from *S. viridis* because of established transformation protocols and facilities for this species.

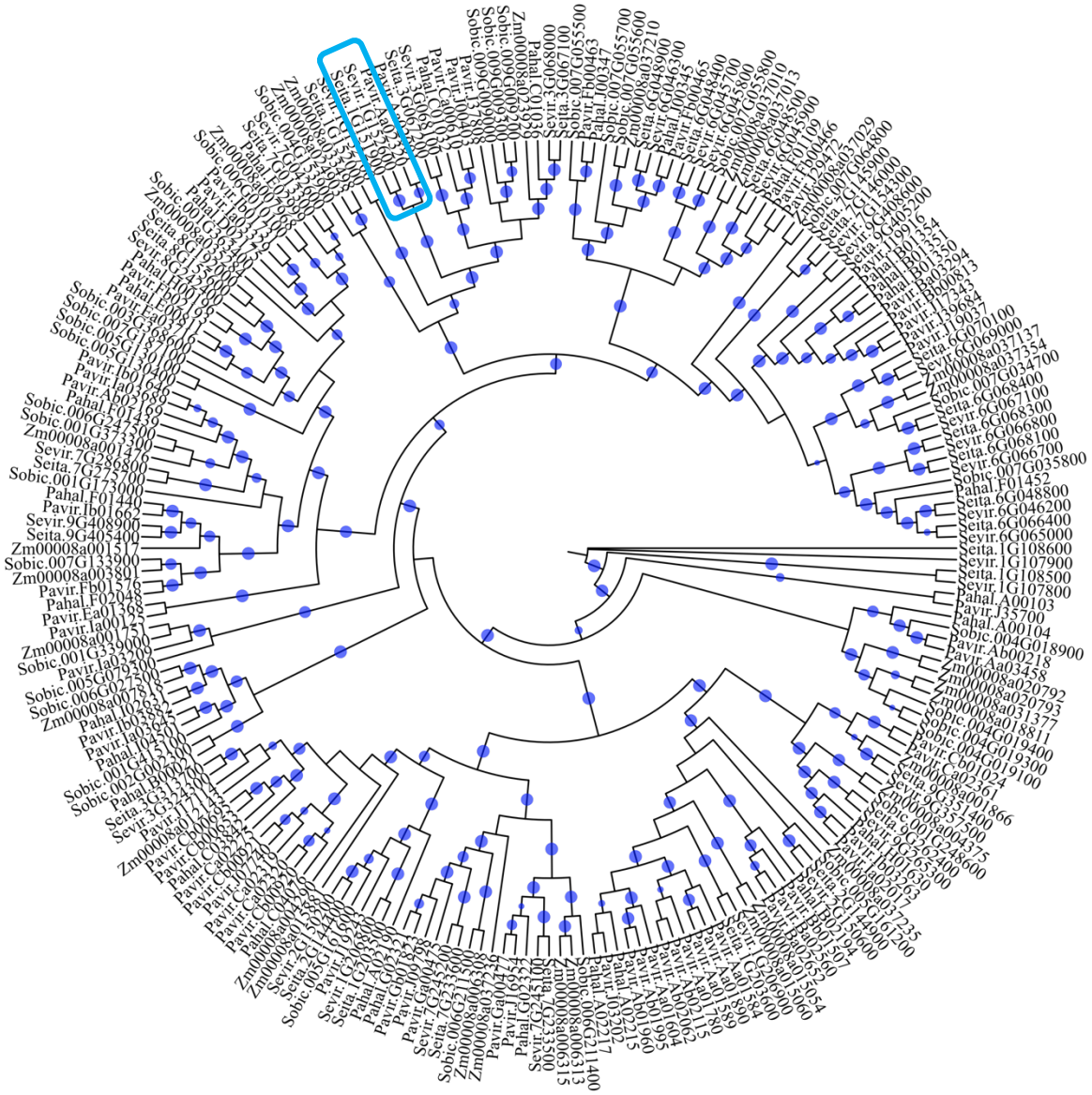


Figure 2. Maximum-likelihood phylogenetic analysis of putative terpene synthases across sequenced Panicoideae genomes (https://phytozome.jgi.doe.gov). Shaded blue circles represent ≥ 80% branch support over 500 bootstrap replications. Putative *Setaria* borneol synthase enzymes are highlighted in blue (Sevir.1G152600; Seita.1G151900, *viridis, itlaica*) with the previously characterized borneol synthase from switchgrass (Pavir.Aa02320, *Panicum virgatum*). Zm: *Zea mays*, Sobic: *Sorghum bicolor*.

***SvTPS04* is highly expressed in roots and acts as a (-)-borneol synthase in vitro**

To determine the tissue specificity of expression of the *SvTPS04* gene, we first performed hierarchal cluster analysis of gene expression in publically available transcriptomes of different *Setaria* tissues. We found highest expression of *SvTPS04* in belowground tissues consistent with borneol production (Supplemental Fig. S4; Fig. 3). RT-qPCR of *SvTPS04* transcript abundance, using gene specific primers (Supplemental Table S1), demonstrated significant differences in expression across panicle, culm, leaf, and root tissues (Fig. 3; Student's T-Test; $p < 0.001$). Highest expression was detected in root tissue indicating borneol biosynthesis occurs in roots of *S. viridis* potentially by the activity of *Sv*TPS04 (Fig. 3).

To determine if *Sv*TPS04 functions as a borneol synthase in vitro, we performed biochemical assays with purified recombinant protein (see Material and Methods). In the presence of geranyl diphosphate (GPP), *Sv*TPS04 catalyzed the formation of borneol as the major product (Fig. 4). As borneol occurs in the form of two enantiomers (+/-), we conducted chiral compound analysis by GC/MS (see Material and Methods). We found that *Sv*TPS04 produced (-)-borneol with GPP in vitro (Fig. 4), which is consistent with the chiral separation and detection of (-)-borneol extracted from *S. viridis* roots (Supplemental Fig. S5).

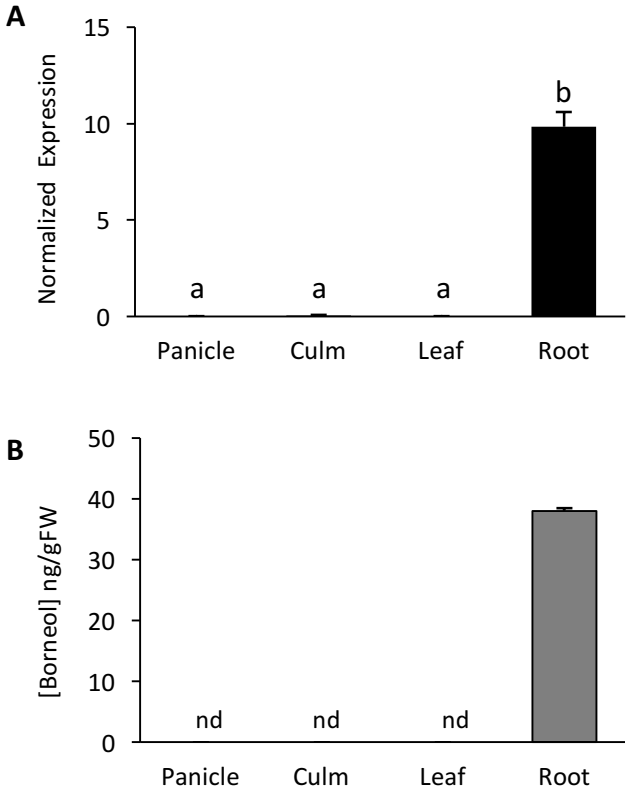


Figure 3. Analysis of *SvTPS*04 expression and borneol accumulation across *Setaria viridis* tissues. A) RT-qPCR from *Setaria* tissues. Samples were analyzed in biological and technical triplicate and normalized to the expression of actin using the ΔΔCt method. B) Quantitative analysis of borneol levels in *Setaria* tissues used in part A. Samples were analyzed by GC-MS in biological triplicate. Borneol levels were quantified using a standard curve with authentic borneol (Sigma-Aldrich) and normalized to gram fresh weight (gFW). Statistical analysis was performed with a Student's T-Test where $\alpha \leq 0.05$. Significant differences are indicated by different letters. nd = not detected

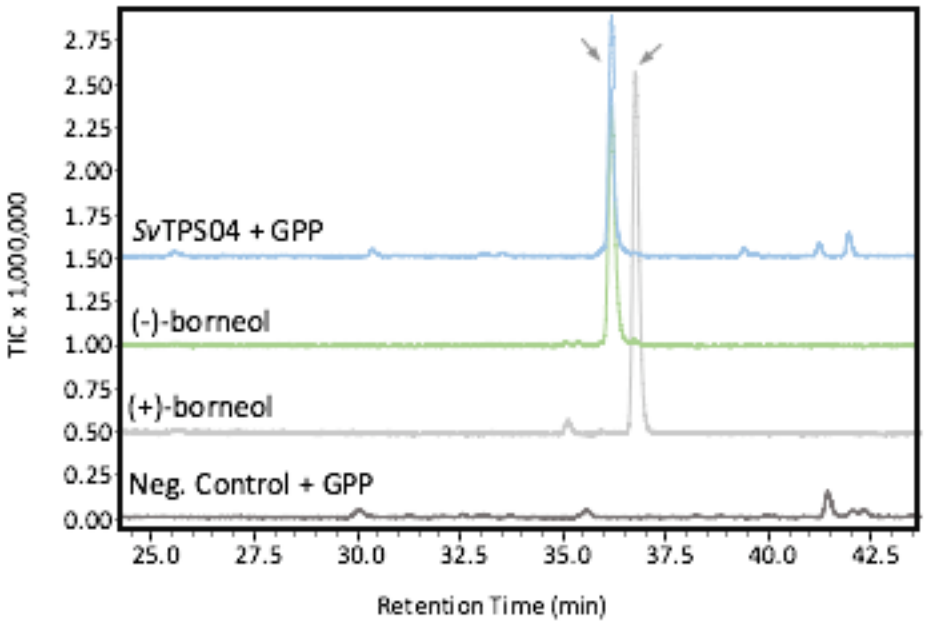


Figure 4. Chiral GC-MS analysis of hexane-extracted volatile products from an in vitro assay of purified *Sv*TPS04 with geranyl diphosphate (GPP) as a substrate. Mass spectra and retention times of enzyme-produced borneol were compared to authentic standards (Sigma-Aldrich) to determine compound identity and stereospecific configuration. Arrows indicate borneol peaks. Protein preparations from *E. coli* carrying the empty expression vector were assayed as a negative control.

Downregulation of *SvTPS04* leads to reduced borneol formation in planta

In attempts to determine the biological function of borneol production in *S. viridis* roots, we used an RNAi based approach to knock down borneol production and evaluate possible changes in plant-microbe associations. In collaboration with the Boyce Thompson Institute, we developed ten independent transgenic lines for evaluation of borneol reduction. Transgenic 0 plants were allowed to self-pollinate and produce T_1 lines which were further screened on selective media and grown until seed collection. T_2 plants were then germinated on selective media, transferred to potting substrate, and screened for reduced borneol formation (see Materials and Methods). We found that seven transgenic lines had no detectable borneol production compared

to wild-type (WT) and empty vector (EV) lines (Supplemental Fig. S6). Three lines (TPS04-1, TPS04-13 and TPS04-12) still produced borneol, although at lower levels compared to WT and EV plants, likely due to a partial knockdown of *SvTPS*04 expression.

We selected three borneol deficient lines for further quantitative analysis of borneol levels and expression of *SvTPS04* in roots. Borneol concentrations were found to be ~40 ng/gram fresh weight in WT and EV lines. By contrast, borneol was not observed in transgenic plants grown in potting substrate or in field soil under greenhouse conditions (Fig. 5, Supplemental Fig. S7). Soil grown plants were further analyzed for *SvTPS04* expression, and showed significant reduction of *SvTPS04* transcript abundance in mutant lines (TPS04-2, TPS04-4 and TPS04-13) compared to WT and EV lines (Fig. 5; Student's T-Test; $p < 0.001$). Although a trend for increased *SvTPS04* expression and borneol formation was observed in EV lines no significant differences were found between WT and EV plants (Fig. 5; Student's T-Test; $p = 0.348$). No phenotypic differences were observed between WT, EV and transgenic plants (Supplemental Fig. S8).

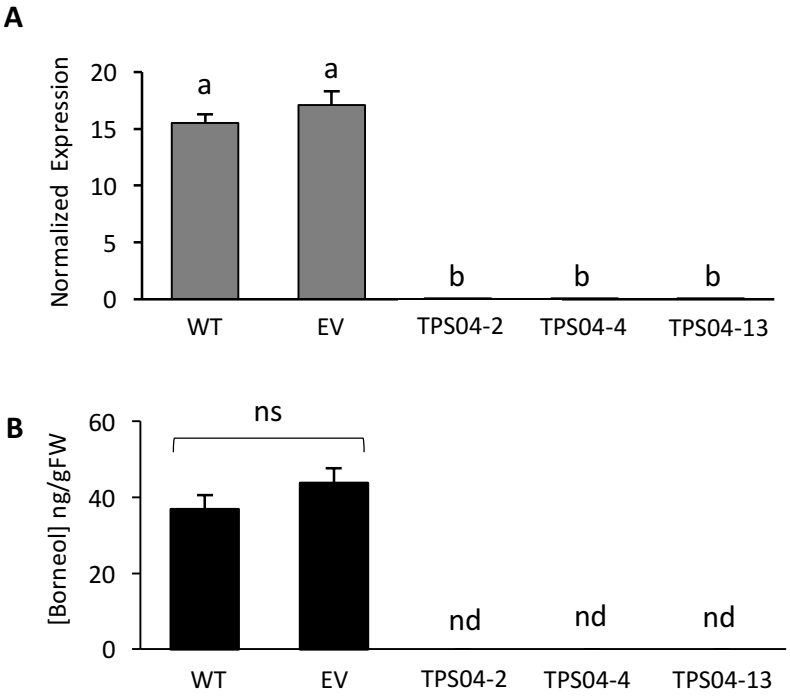


Figure 5. A) Expression of *SvTPS*04 in soil grown WT and transgenic *S. viridis* roots. Samples were analyzed in biological and technical triplicate and normalized to the expression of actin using the ΔΔCt method. Statistical analysis was performed with a Student's T-Test where $\alpha \leq 0.05$. Significant differences are indicated by different letters. B) Quantitative analysis of borneol levels in *S. viridis* tissues used in part A. Samples were analyzed by GC-MS in biological triplicate and borneol levels were quantified using a standard curve to authentic borneol (Sigma-Aldrich) and normalized to gram fresh weight (gFW). ns = non-significant difference, detected nd = not detected.

Borneol reduction in roots does not affect the root microbiome of *S. viridis*

To determine if reduction in borneol levels impacted the root microbiome of *S. viridis*, we extracted genomic DNA from bulk soil, rhizosphere, and endosphere samples of the plants described above. These samples were sent to the Vincent J. Coates Genomics Sequencing Laboratory (University of Berkeley) for 16S rRNA sequencing using an Illumina MiSeq platform. Generated reads (avg. 60,000-160,000 per sample; Supplemental Fig. S9) were then processed using QIIME2 (see Material and Methods). We found that rhizosphere and endosphere samples form distinct "groups" based on the UniFrac distance metric and visualization by non-metric multidimensional scaling (NMDS; Fig. 6 and Supplemental Fig. S10. Similar to switchgrass, Proteobacteria represent the dominant phyla in the root endosphere of *Setaria*. However, visual inspection of phyla level abundances across root zones and plant lines did not indicate any major patterns associated with reduction of borneol. Within endosphere samples, enrichment analysis by pair-wise comparisons between transgenic and WT samples using a Mann-Whitney U test did not result in any OTUs that were significantly enriched or depleted as a results of altered borneol levels ($p > 0.05$).

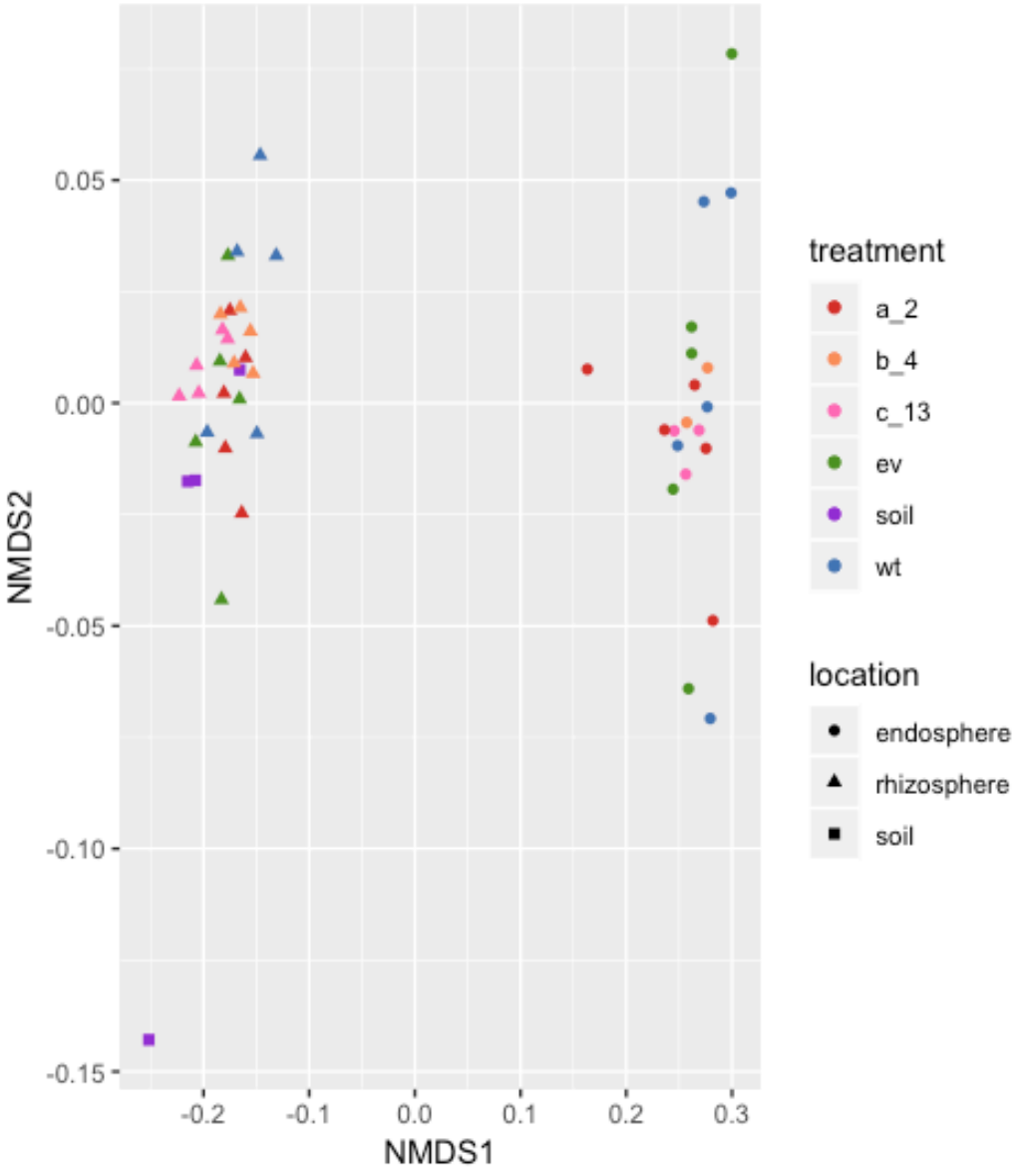


Figure 6. Non-metric multidimensional scaling (NMDS) plot based on Unifrac distance of microbial communities associated with roots of soil grown WT and transgenic *S. viridis* plants. Shapes and colors indicate sample type where the naming convention = wild-type (wt) and transgenic (empty vector - ev, mutant line 2, 4, and 13) plants. Samples with <2,000 reads were not included. Additional lettering (a-c) serves only as a naming convention. Unplanted bulk soil was included as negative controls.

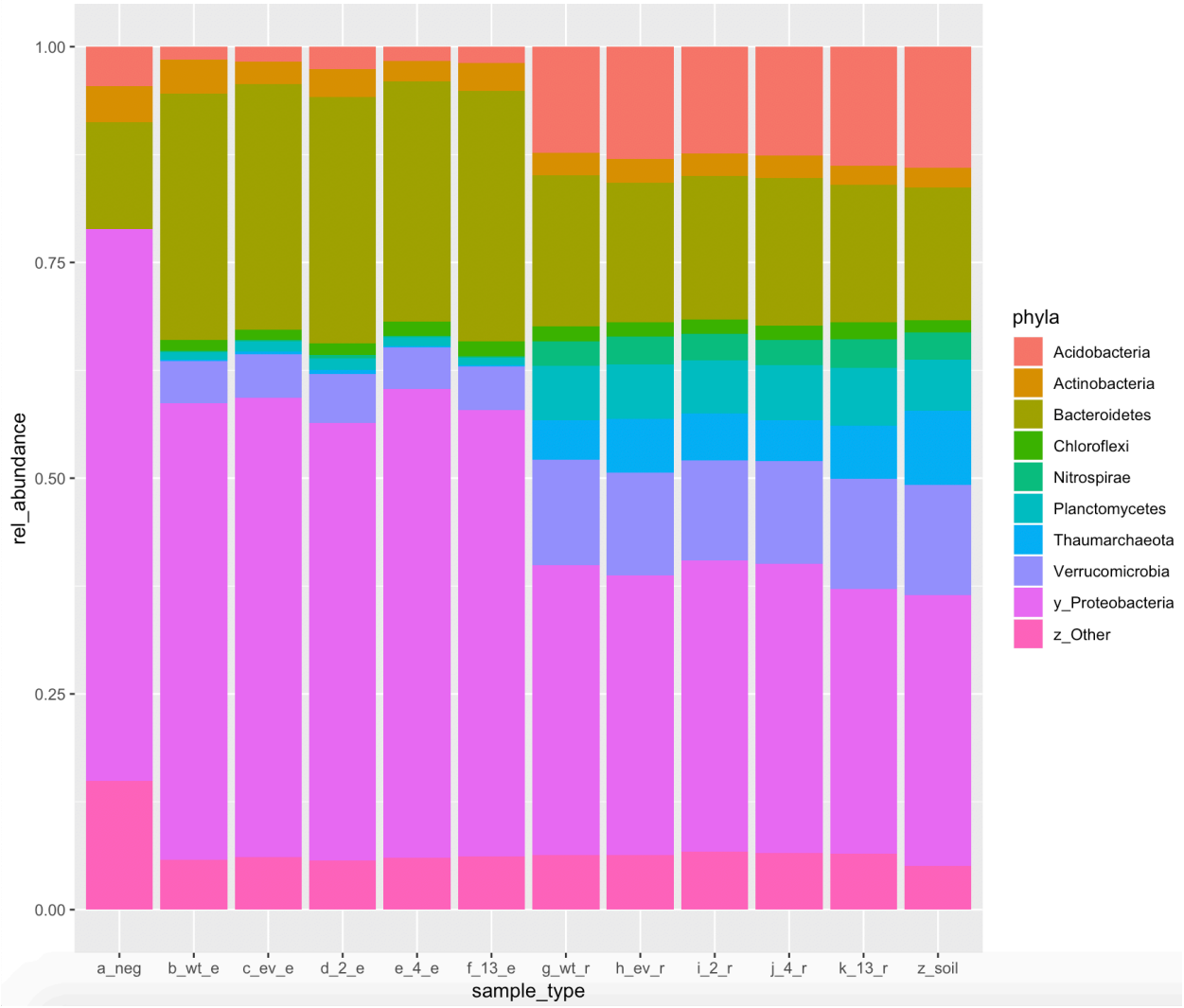


Figure 7. Stacked bar plot of average phyla level comparisons of 16S rRNA data from endosphere (e) and rhizosphere (r) of wild-type (wt) and transgenic (empty vector - ev, mutant line 2, 4, and 13) plants. Non-template sequencing (neg) and unplanted bulk soil (soil) were included as negative controls. Samples with <2,000 reads were not included. Additional lettering (a-k, z) serves only as a naming convention.

CONCLUSIONS AND FUTURE DIRECTIONS

Due to challenges with reducing borneol levels in roots of switchgrass, we used *Setaria* to continue pursuit of our hypothesis that borneol influences the microbial community inside roots of *S. viridis* at a significant scale. We were unable to support this hypothesis based on current analysis. Follow-up analyses will be performed to determine whether the abundance of few individual OTUs might be affected at the family or genus level. It is possible that borneol levels in *Setaria* roots are simply too low to have significant effects on root-associated bacteria. On the other hand, borneol may have different biological functions in constitutive defense since borneol concentrations are not affected by stress treatments. Future work could test if borneol deficient plants are more susceptible to monocot root pathogens e.g. *Pythium graminicola*. In addition, fungal communities associated with the root endosphere could be investigated to determine if borneol deficient plants have an altered fungal community composition including alterations in mycorrhiza fungal associations.

METHODS

Plant growth conditions

Seed and plant material provided by the Boyce Thompson Institute (Ithaca, NY) for *Setaria viridis* genotype A10.1 was used throughout this study. Wild-type and transgenic seedlings were grown in a Percival growth chamber at 23^{o}C with a 16 hr photoperiod at 210 µE m^{-2} s^{-1}. Seeds first were collected from T1 plants, dehulled and surface sterilized (10% bleach, 0.1% Tween, 3 min) and germinated on selective rooting media (per 1 L: 20 mg hygromycin, 7 g Phytoagar, 10 ml MS Vitamins (100X), 30 g sucrose, 2.2 g MS salts 1/2 strength) in Magenta boxes (Thomas Scientific).

Plants were transferred to SunGro potting substrate (8:1, soil:sand) in 4 inch pots once they reached a height of ~4 cm and grown until seeds could be collected (5-6 weeks). Collected seeds (T2) were sterilized and germinated as described above, grown in potting substrate as above for 4 weeks and used for initial screening for borneol reduction. Following screening, selected T2 RNAi lines (including WT and empty vector lines) were sterilized and germinated as above and transplanted to 4 inch pots containing freshly collected field soil from the Moore Farm (Blacksburg, Virginia). Plants were grown under greenhouse conditions for 4 weeks prior to tissue collection.

Qualitative tissue specific analysis of borneol production

Fresh WT (*Sv*A10.1) plant material (1 g) was detached and immediately placed in a 20 ml screw cap vial containing 1 ml of DI water and sealed with silicone septa caps (Supelco). Samples were incubated for 1 h at 30^{o}C in the presence of a 100-µM polydimethylsiloxane fiber (Supelco). Collected volatiles where thermally desorbed for 4 min and analyzed using a gas chromatograph (240^{o}C injector port) coupled with a quadrupole mass spectrometer (GC-MS-QP2010S, Shimadzu). Extracts were separated with a 2:1 split on a 30 m x 0.25 mm ID x 0.25 µm Zebron capillary column (Phenomenex) using Helium as the carrier gas (1.4 ml min^{-1} flow rate) and a temperature gradient of 5^{o}C min^{-1} from 40^{o}C (hold 2 min) to 220^{o}C. Identification of borneol was confirmed by comparisons of retention time and mass spectra to those of an authentic borneol standard.

Plant stress treatments

WT plants were germinated and grown under greenhouse conditions as described above for 4 weeks prior to treatments. Methyl jasmonate, salicylic acid and abscisic acid were first dissolved

in ethanol and applied in 1 mM concentrations in a 25 ml volume by watering plants. Copper sulfate was applied in a similar manner at a concentration of 10 mM. A mock treatment with ethanol was used as a control and root tissue was collected after 24 hours. For drought stress, water was omitted from plants for 14 days prior to collecting root tissue.

Gene identification and phylogenetic reconstruction

Protein sequence data were retrieved from Phytozome (http://www.phytozome.net/). Maximum likelihood (ML) trees were built from a MAFFT alignment using PhyML (Guindon, 2010 #2959) with 500 bootstrap replicates as previously described by Pelot et al. (2018). Final phylogeny annotation and design were performed in iTOL (Letunic and Bork, 2007).

Amplification of SvTPS04 and plasmid construction for bacterial expression

Using primers in Supplemental Table 1, a 63 bp truncated version of full-length *SvTPS04* was amplified with Q5 High-Fidelity DNA polymerase (ThermoFisher Scientific) in a 25 µl reaction volume with the following PCR conditions: 98^{o}C for 30 s, followed by 30 cycles of 98^{o}C for 30 s, 55^{o}C for 30 s, 72^{o}C for 1 min 45 s and a final extension at 72^{o}C for 2 min The amplified fragment was gel purified using a NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel, MN) and concentrated to ~10 µl using a CentriVap concentrator (Labconco). A 10 µl A-tailing reaction was prepared with 3 µl of purified PCR product and incubated in the presence of 10 mM dATPs and Taq polymerase at 72^{o}C for 30 min. The resulting product was ligated overnight into the pGEM-T Easy vector (Promega) and Sanger sequenced to verify the insert. The open reading frame was then digested with *Bam*HI and *Xho*I and ligated overnight into the corresponding restriction sites of the bacterial expression vector pET28a (Novagen).

Recombinant protein expression in Escherichia coli and TPS assays

The constructed plasmid was transformed into *E. coli* BL21-CodonPLus(DE3) cells (Stratagene) and grown at 37^{o}C in 100 ml Luria-Bertani (LB) media supplemented with 50 µM Kanamycin until the optical density at 600 nm (OD_{600}) reached 0.5 to 0.7. Protein production was then induced with 0.5 mM isopropyl 1-thio-ß-D-galactopyranoside (IPTG) and incubated with shaking at 18^{o}C for 16 h. Cell pellets were washed with 10 mM Tris base and 50 mM potassium chloride, resuspended in 4 ml phosphate buffered saline (PBS, 50 mM sodium phosphate, 100 mM sodium chloride, 10% glycerol) supplemented with 1 mM dithiothreitol (DTT) and 0.5 mM phenylmethylsulfonyl fluoride (PMSF), and ruptured by sonication. Clarified extracts were mixed with equal parts of PBS and recombinant His(6x)-tagged proteins were partially purified by Ni^{2+} affinity chromatography according to the manufacturer's instructions (Qiagen). Partially purified proteins were then desalted on PD-10 desalting columns (GE) equilibrated with assay buffer (10 mM MOPSO, 10% glycerol [v/v] and 1 mM DTT, pH 7.0) and visualized by SDS-PAGE (10%, GenScript). Enzyme reactions (125 µl total volume) were prepared in a 10 ml screw cap vial (Supelco) by combing partially purified protein with 20 mM $MgCl_2$ and 60 µM commercially available prenyl diphosphate substrates including GPP, NPP, (*E,E*)-FPP, (*Z,Z*)-FPP and GGPP (Echelon Biosciences). Assay mixtures were analyzed using automated solid phase microextraction (SPME, AOC-5000 Shimadzu) following a 5 min incubation at 30^{o}C in the presence of a 100-µM polydimethylsiloxane fiber (Supelco). Collected volatiles where thermally desorbed for 4 min and analyzed using a gas chromatograph (240^{o}C injector port) coupled with a quadrupole mass spectrometer (GC-MS-QP2010S, Shimadzu). Extracts were separated with a 5:1 split on a 30 m x 0.25 mm ID x 0.25 µm Zebron capillary column (Phenomenex) using helium as the carrier gas (1.4 ml min^{-1} flow rate) and a temperature gradient of 5^{o}C min^{-1} from 40^{o}C (hold 2

min) to 220°C. Identification of borneol was confirmed by comparisons of retention time and mass spectra to those of an authentic borneol standard.

Generation of SvTPS04 RNAi lines

Using primers in Supplemental Table 1, a 400 bp region flanking the 3' region of the *SvTPS04* gene (200 bp in the ORF and 200 in the 3'-UTR) was amplified with Q5 High-Fidelity DNA polymerase (ThermoFisher Scientific) in a 25 µl reaction volume with the following PCR conditions: 98°C for 30 s, followed by 30 cycles of 98°C for 30 s, 55°C for 30 s, 72°C for 1 min and a final extension at 72°C for 2 min. The amplified fragment was gel purified using a NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel, MN) and concentrated to ~10 µl. A 10 µl A-tailing reaction was prepared with 3 µl of purified PCR product incubated in the presence of 10 mM dATPs and Taq polymerase at 72°C for 30 min. The resulting product was ligated overnight into the pGEM-T Easy vector (Promega) and Sanger sequenced to verify the insert. Plasmid (2 µg) was digested with restriction enzymes (*Bam*HI, *Xho*I) and the resulting ~400 bp fragment was purified as described above and ligated overnight into the corresponding restriction sites of the entry vector pENTR11 (ThermoFisher Scientific). An LR clonase reaction (ThermoFisher Scientific) was then performed with the monocot specific RNAi vector pANIC12a (ABRC Stock #1720) to generate the final construct. Plant transformation and regeneration was carried out by the Boyce Thompson Institute (Ithaca, NY) following established methods (Van Eck, 2018).

Borneol extraction and quantification by GC-MS

Fresh root material (1 g) was rinsed with deionized water, dried with tissue paper and immediately frozen in liquid nitrogen for processing. Samples were ground to a fine powder for 2 min in the presence of liquid nitrogen, weighed, transferred to 5 ml hexanes and mixed by vortex for 20 sec. The ground material was placed in an ultrasonic bath (Fisher Scientific) for 10 min and then pelleted by centrifugation. Following the collection of two fractions, 1-bromodecane was added for a final concentration of 20 ng/µl as an internal standard and extracts were dried over a $MgSO_4$ mock column and concentrated on ice to ~40 µl under a gentle stream of nitrogen. Extracts were analyzed using a gas chromatograph (240^{o}C injector port) coupled with a quadrupole mass spectrometer (GC-MS-QP2010S, Shimadzu). Extracts were separated with a 2:1 split on a 30 m x 0.25 mm ID x 0.25 µm Zebron capillary column (Phenomenex) using Helium as the carrier gas (1.4 ml min^{-1} flow rate) and a temperature gradient of 5^{o}C min^{-1} from 40^{o}C (hold 2 min) to 220^{o}C. Identification of borneol was confirmed by comparisons of retention time and mass spectra to those of an authentic borneol standard. Borneol quantification was attained by multipoint standard curve analysis with authentic borneol (Sigma-Aldrich) and normalization to gram fresh weight.

Determination of borneol stereochemistry

Root hexane extracts described above were re-analyzed by chiral separation on a Machery-Nagel fused silica capillary column (HYDRODEX-ß-3P) using a temperature program from 40^{o}C to 230^{o}C with a ramp speed of 2^{o}C/min and the instrument described above. Chiral compound analysis from TPS assays was achieved by incubating partially purified protein as described above with 60 µM GPP and 10 mM $MgCl_2$ for 1 h at 30^{o}C with the addition of a 1 ml hexane overlay. Following incubation, hexane fractions were dried over magnesium sulfate ($MgSO_4$), concentrated to ~40 µl, and 1 µl was analyzed by GC-MS using the methods described above. Identification of

borneol enantiomers was confirmed by comparisons of retention times to authentic standards (Sigma-Aldrich).

Total RNA extraction and RT-qPCR

Total RNA was extracted from root tissue described above in biological triplicate using the TRIzol reagent (Life Technologies, Carlsbad, CA) and the Qiagen Plant RNAeasy kit (Qiagen) in accordance with the manufacturer's protocol. RNA was treated for DNA contamination with DNAse (Qiagen), and used for first strand cDNA synthesis with SuperScriptII reverse transcriptase and oligo$(dT)_{18}$ primer (Invitrogen) according to the manufacturer's instructions.

RNA quantity was first normalized between samples and replicates to 2.5 µg based on denaturing gel electrophoresis and spectrophotometer measurements at 260 nm. The resulting cDNA was diluted to 100 ng/µl. Reactions were performed with 1 µl cDNA in a 20 µl reaction using Power SYBR Green PCR master mix (Applied Biosystems) and gene specific primers (Supplemental Table S1). PCR amplifications were done in biological and technical replicates with a CFX96 Touch real-time PCR detection system (Bio-Rad) with the following cycles: 95^{o}C for 10 min, followed by 40 cycles of 95^{o}C for 15 s, 50^{o}C 30 s and 60^{o}C for 1 min. Melt curve analysis was performed at the end of amplification to ensure specificity of each primer pair. Relative expression levels across tissues were calculated using the ΔΔCt relative quantification method and normalized to actin.

Bacterial gDNA extraction

DNA from unplanted soil and soil particles attached to roots were extracted using the MoBio Power Soil DNA extraction kit (Qiagen) following the manufactures instructions. After removal of soil particles, roots were rinsed (5x) with sterile DI water. Roots were then surface sterilized

by submerging in sterile water and subjecting to 10 minutes of high frequency bursts using an ultrasonic bath (Fisher Scientific) to wash away surface DNA following established methods by Del Giudice et al. (2008) and Edwards et al. (2015). Roots were checked for sterility by placing root material on LB agar plates and incubating at 28^{o}C for 16 h. Sterile roots were then pulverized under liquid nitrogen prior to DNA extraction used for soil as described above. Samples were sent to the Vincent J. Coates Genomics Sequencing Laboratory (University of Berkeley) for 16S sequencing (http://qb3.berkeley.edu/gsl/).

Gene identification and phylogenetic reconstruction

Gene expression data for *Setaria* borneol synthase genes was obtained from publically available transcriptome-based data on Phytozome (https://phytozome.jgi.doe.gov). FPKM values were compared across tissues/conditions by hierarchical clustering analysis in XLSTAT 2018 as described by Pelot et al. (2018).

Analysis of 16S rRNA data

Raw data provided by the Vincent J. Coates Genomics Sequencing Laboratory (University of Berkeley) were analyzed following the in house pipeline for QIIIME2 (APPENDIX 3, https://github.com/scubalaina/plants_lol/blob/master/README.md). Following data import, raw files were cleaned to remove and trim base pairs with low quality scores using the tool DADA2 (https://benjjneb.github.io/dada2/). Representative sequences were then clustered based on similarity with a sequence identity threshold of 99%. Resulting clustered operational taxonomic units (OTUs) were then classified and further processed using R. A UniFrac distance matrix was calculated from normalized OTU counts and utilized for NDMS plotting with ggplot2. Average sample composition at the phyla level was visualized from the normalized OTU count table using

a stacked barplot in QIIME2. Statistical enrichment analysis of samples was done by performing pair-wise comparisons between treatments and WT samples using a Mann-Whitney U test.

REFERENCES

Chen, X.J., Chen, H., Yuan, J.S., Kollner, T.G., Chen, Y.Y., Guo, Y.F., Zhuang, X.F., Chen, X.L., Zhang, Y.J., Fu, J.Y., Nebenfuhr, A., Guo, Z.J., and Chen, F. (2018). The rice terpene synthase gene OsTPS19 functions as an (*S*)-limonene synthase in planta, and its overexpression leads to enhanced resistance to the blast fungus *Magnaporthe oryzae*. *Plant Biotechnol. J.* 16**,** 1778-1787. doi: 10.1111/pbi.12914.

Del Giudice, L., Massardo, D.R., Pontieri, P., Bertea, C.M., Mombello, D., Carata, E., Tredici, S.M., Tala, A., Mucciarelli, M., Groudeva, V.I., De Stefano, M., Vigliotta, G., Maffei, M.E., and Alifano, P. (2008). The microbial community of *Vetiver* root and its involvement into essential oil biogenesis. *Environ. Microbiol.* 10**,** 2824-2841. doi: 10.1111/j.1462-2920.2008.01703.x.

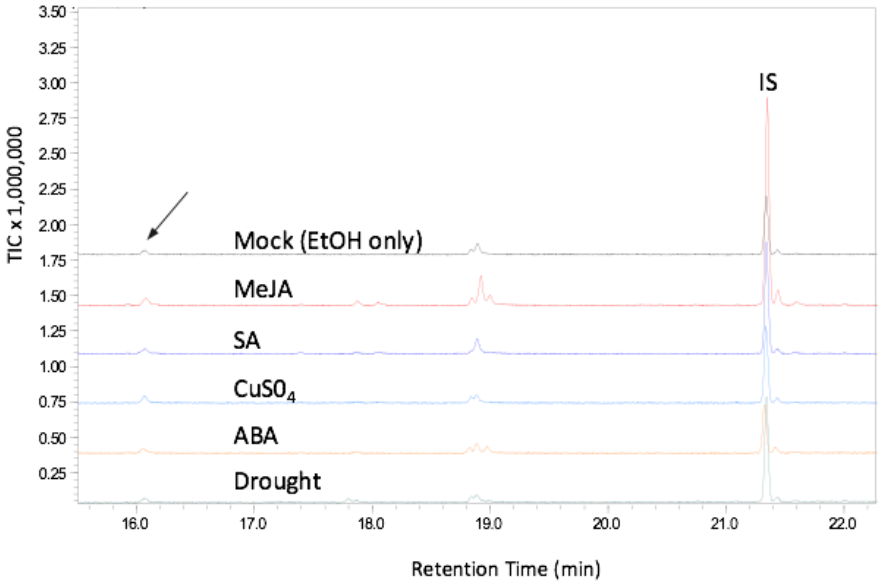
Edwards, J., Johnson, C., Santos-Medellin, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., Eisen, J.A., and Sundaresan, V. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* 112**,** E911-920. doi: 10.1073/pnas.1414592112.

Letunic, I., and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23**,** 127-128. doi: 10.1093/bioinformatics/btl529.

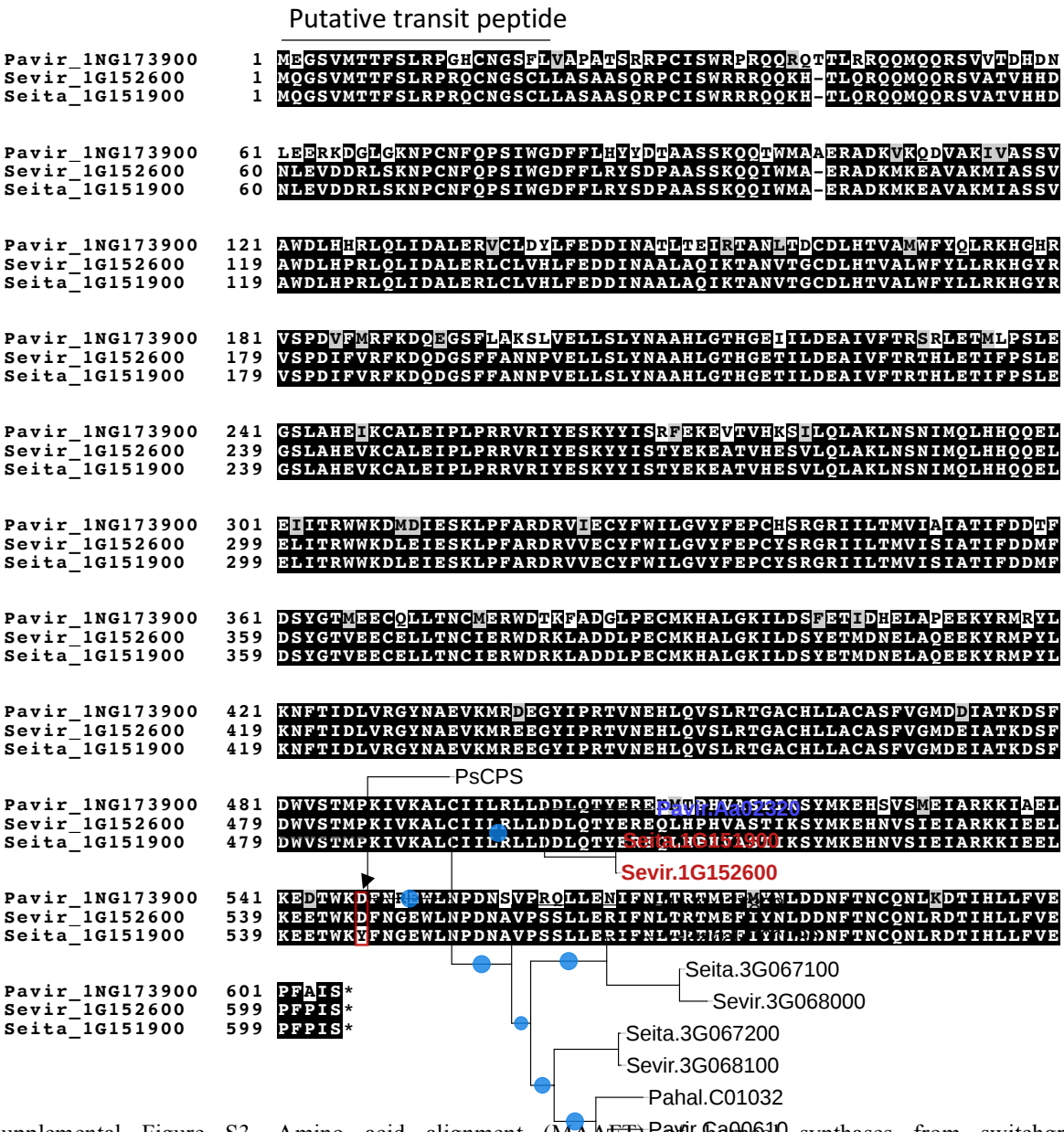
Pelot, K.A., Chen, R., Hagelthorn, D.M., Young, C.A., Addison, J.B., Muchlinski, A., Tholl, D., and Zerbe, P. (2018). Functional diversity of diterpene synthases in the biofuel crop switchgrass. *Plant Physiol.* 178**,** 54-71. doi: 10.1104/pp.18.00590.

Van Eck, J. (2018). The status of *Setaria viridis* transformation: *Agrobacterium*-mediated to floral dip. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00652.

SUPPLEMENTAL DATA

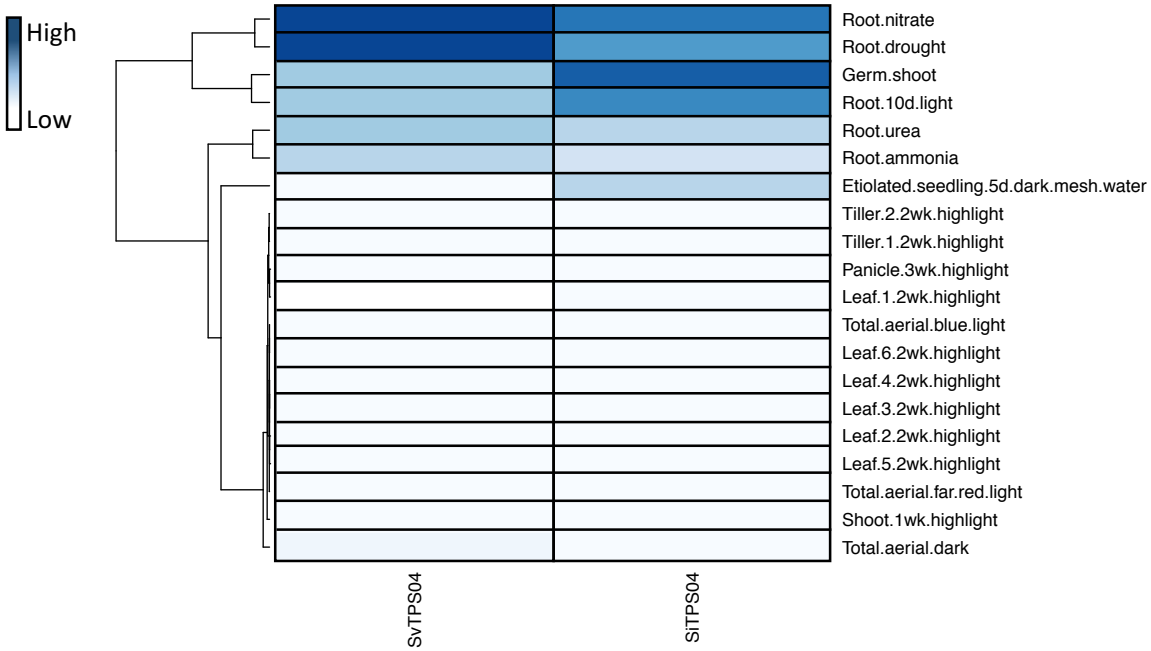


Supplemental Figure S1. GC-MS analysis of hexane extracted volatiles from roots treated with phytohormones and abiotic stressors. Samples were analyzed in biological triplicate and the arrow indicates the peak of borneol. IS = internal standard 1-bromodecane. MeJA- methyl jasmonate (1 mM), SA- salicylic acid (1 mM), ABA – abscisic acid (1 mM), $CuS0_4$ – copper sulfate (10 mM), Drought (14 days without water). Borneol identification was based on comparison to authentic borneol.

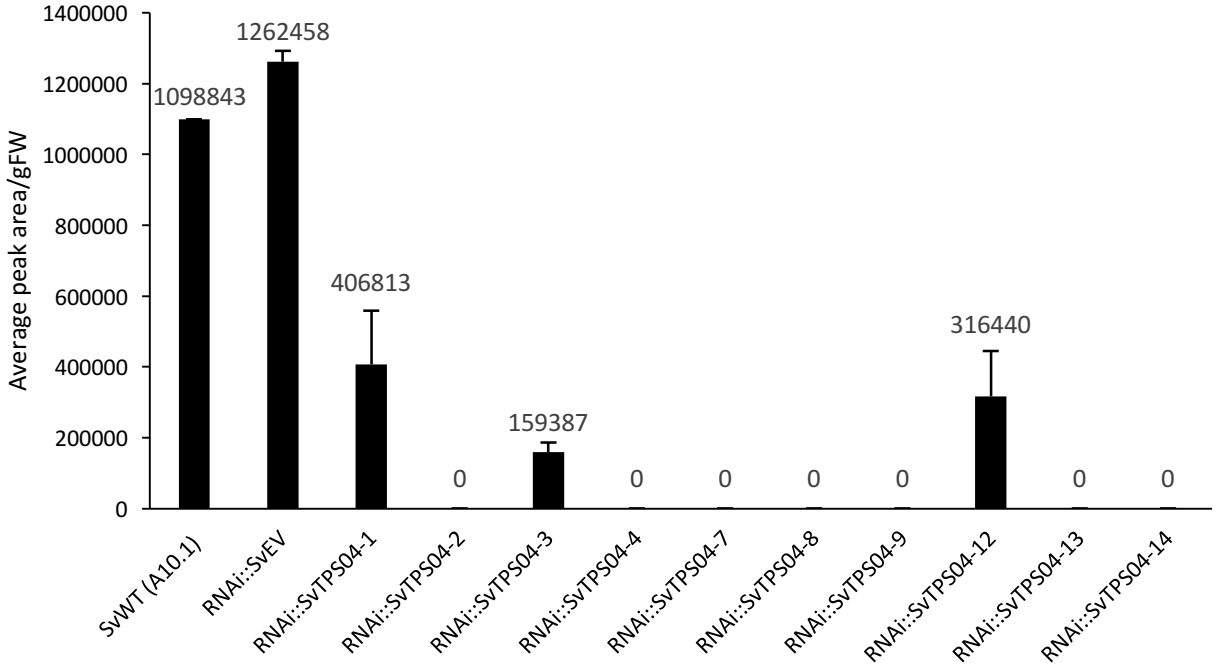


Supplemental Figure S3. Amino acid alignment (MAAFT) of borneol synthases from switchgrass (Pavir_1NG173900) and *Setaria* (Sevir_1G152600; Seita_1G151900). The red arrow indicates the single amino acid change between *S. viridis and S. italica*.

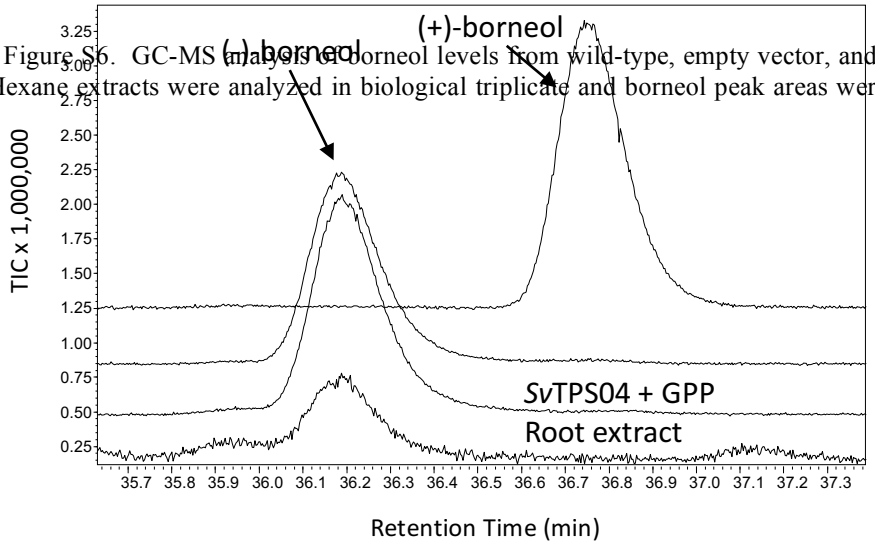
Supplemental Figure S2. Maximum-likelihood phylogenetic analysis of TPS-a type terpene synthases from Poaceae grasses. Blue circles indicate greater than 80% branch support over 500 bootstrap replications. The previously characterized borneol synthase from switchgrass is highlighted in blue. The putative borneol synthases from *Setaria* are highlighted in red. The *ent*-CPP synthase from *Picea sitchensis* (PsCPS) serves as an outgroup. Pavir: *Panicum virgatum*, Seita: *Setaria italica*, Sevir: *Setaria viridis*, Pahal: *Panicum Halii*.



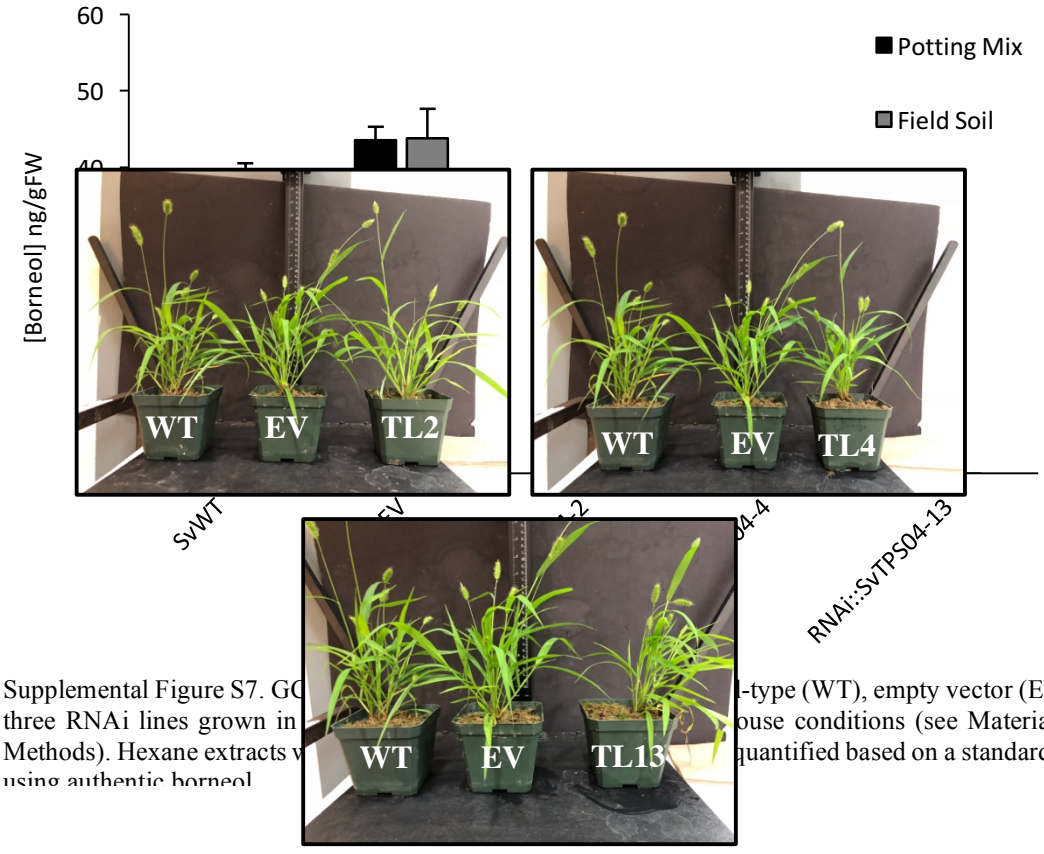
Supplemental Figure S4. Hierarchical cluster analysis of *Setaria TPS04* expression across tissues and treatments. The heat map compares relative transcript abundance in FPKM (Fragments Per Kilobase of transcript per Million mapped reads) from data sets at https://phytozome.jgi.doe.gov/.



Supplemental Figure S6. GC-MS analysis of borneol levels from wild-type, empty vector, and ten transgenic RNAi lines. Hexane extracts were analyzed in biological triplicate and borneol peak areas were averaged per sample.

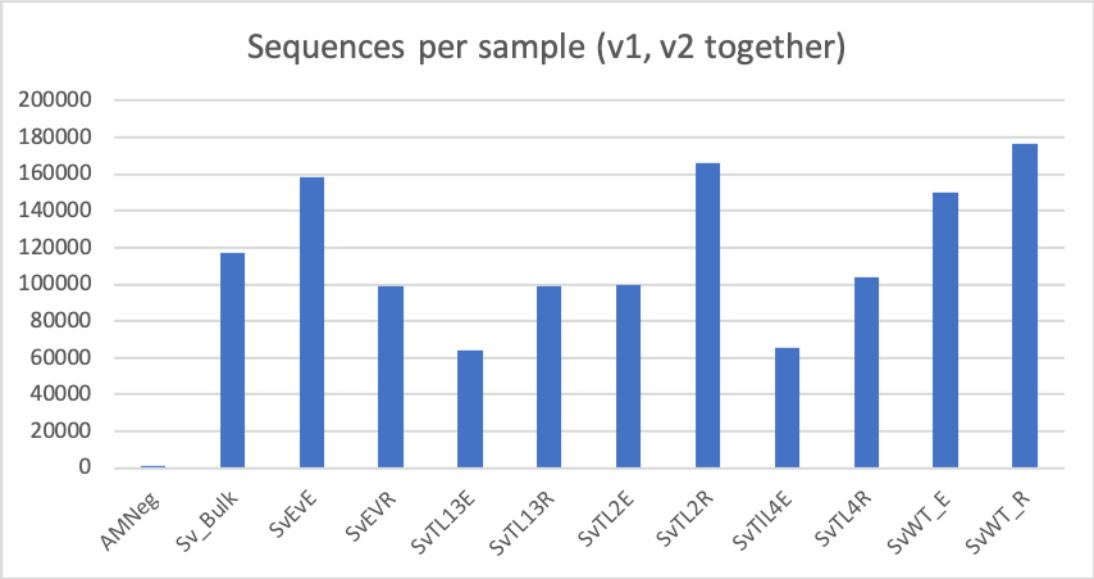


Supplemental Figure S5. Chiral GC-MS analysis of hexane extracted volatile products from the in vitro assay of purified *Sv*TPS04 with geranyl diphosphate (GPP) and borneol extracted from roots. Retention times of borneol peaks from extracts were compared to authentic standards (Sigma-Aldrich) to determine stereospecificity.

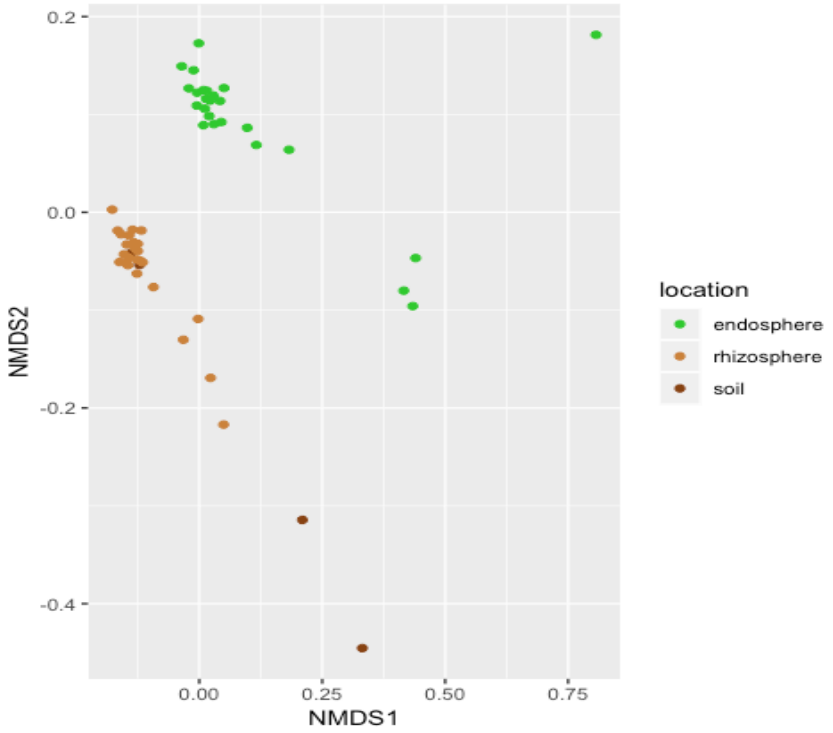


Supplemental Figure S7. GC
l-type (WT), empty vector (EV) and three RNAi lines grown in
ouse conditions (see Materials and Methods). Hexane extracts v
quantified based on a standard curve using authentic borneol.

Supplemental Figure 8. Comparison of wild-type (WT), empty vector (EV) and transgenic lines (TL) used in this study.



Supplemental Figure S9. Average read counts across all 16S samples determined by following the QIIME2 pipeline (https://qiime2.org).



Supplemental Figure S10. non-metric multidimensional scaling (NMDS) analysis of microbial communities associated with the endosphere and rhizosphere of soil grown *Setaria* plants. Unplanted bulk soil was included as a negative control.

Supplemental Table 1. Primers used in this study.

Primer Name	Primer Sequence 5' – 3'	Purpose
SvActinF	GTG CTT TCC CTC TAC GCC AGT G	Housekeeping Control Gene
SvActinR	ACC GCT GAG CAC AAT GTT ACC A	Housekeeping Control Gene
SvTPS04_TCF	GGATCCATGGCATCAGCAGCATCACAG	Full-length Gene Cloning
SvTPS04_TCR	CTCGAGCTAGGAGATAGGAAAGGGCTCC	Full-length Gene Cloning
SvTPS04geneUTRTCF	GGATCCAGATTTTAACGGCGAGTG	RNAi
SvTPS04geneUTRTCR	CTCGAGTTTTTTGAGTAAAATGTGACTGGTTGTC	RNAi

APPENDIX C: QIIME STANDARD OPERATING PROCEDURE FOR 16S DATA ANALYSIS

Download the data from the server. This will be a .tar file. Extract the .tar archive using the graphical manager. This is done by right clicking the file and selecting "Open with Archive Manager". You can then extract the archive by clicking the "Extract" button. You will be given an option to save the extracted files wherever you would like. Save them in a single folder with a meaningful name. Extract all of the .gz files that are produced by right clicking in the new folder and selecting "Open in Terminal" option. Then enter "gunzip *.gz" without the quotes. This will extract all of the .gz files to reproduce the .fastq files. You can then delete all of the .gz files using the command "rm *.gz"

To merge and filter the files we will use USEARCH. The reads from the Illumina machine are a demultiplexed forward and reverse read. These reads must be combined. The command to do so is as follows:

usearch -fastq_mergepairs forward.fastq -reverse reverse.fastq -fastq_truncqual 3 -fastq_maxdiffs 0 -fastqout merged.fastq

The -fastq_mergepair is followed by the forward read, and the -reverse flag is followed by the reverse read. fastq_truncqual is the quality score at which to truncate the reads. -fastq_maxdiffs is the maximum allowed differences between reads. -fastqout is followed by the name of the merged fastq file.

This can be entered into the terminal at the location of the fastq files. The easiest way is to set up a worksheet that produces all of the filenames and inserts them into the command. An example is "merge_and_filter_fastq.sh.csv" located in the same folder as this document. This long list of commands can simply be copied into the terminal or run as a shell script. The next step is to

filter the resulting files, discarding ones likely to have errors. The command is:

usearch -fastq_filter merged.fastq -fastq_maxee 0.5 -fastaout filtered.fasta

Here -fastq_filter is followed by the merged fastq file. -fastq_maxee refers to the maximum expected errors in that read. More information on expected errors can be found at: drive5.com/usearch/manual/expected_errors.html. The -fastaout command turns the fastq file into a fasta file with a given name.

Again, this can be done much more easily in a worksheet and copied into the terminal. Next, make the mapping file for the data. The format is the following:

#SampleID InputFileName Description
1S 1S_filtered.fasta Sample_1
2S 2S_filtered.fasta Sample_2
3S 3S_filtered.fasta Sample_3
…

This should be done in OpenOffice Calc or Excel. The columns should be tab separated values. The top row must be present and begin with #SampleID. The InputFileName and Description fields must also be present with Description as the final column, but other fields may be added which will be associated with any metadata entered in that column. The description column can be either the sample name or a more descriptive name. Next, since the Illumina machine already separated the samples and removed the barcodes, we must add sample IDs back into all of the reads. This is so we are able to place all the sequences into one file for processing and still determine what sequence came from what sample. The command is:

add_qiime_labels.py -m mapping_file -i input_directory -c InputFileName -o output_directory

-m refers to the mapping file. The -i flag is the input directory containing all of the fasta

files. The -c tells the program what column to check in the mapping file for the filenames. The -o is the output directory of the combined sequences file.

Next we can pick the OTUs. This is done by:

pick_open_reference_otus.py -o output_directory -i combined_seqs.fna -r $HOME/gg_13_8_otus/rep_set/97_otus.fasta -a -O 8 -f

The -o command specifies the output directory. The -i command is the input combined sequences file. The -r is the directory of the reference files and should not be changed. -a allows for multiple threads to run at once, and -O 8 forces 8 threads to run. These should not be changed. -f forces overwriting if the output directory already exists. This step can easily take 8 or more hours to complete.

The output can be summarized using the the following command:

biom summarize-table -i input.biom -o output.txt

This produces a summarized output of the numbers of OTUs per sample. This can be helpful in choosing sample depth in the next step.

The final step is doing the diversity analyses. This is done by calling:

core_diversity_analyses.py -o -i -m -e 254 -t -a

The -o and -i are the output and input files respectively. -m referrs to the mapping file. -e is the sampling depth; the number of samples it takes from each OTU for beta-diversity comparisons. This could be chosen based on information from the summariz-table command. -t is the tree file, and -a allows the program to start multiple threads. This may take some time. The results can be accessed most conveniently by opening the index.html in the output directory.