Grapevine terpenoids and their contribution to the flower volatilome

by

Samuel Jacobus Smit

Thesis presented in partial fulfilment of the requirements for the degree of
Master of Science

at
Stellenbosch University
Institute for Wine Biotechnology, Faculty of AgriSciences

Supervisor: Dr Philip Young
Co-supervisor: Prof Melané Vivier

March 2016

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2016

Summary

Terpenoids represent the largest group of plant natural products and show tremendous diversity in chemical structure and bioactivity. This diversity arises from the action of terpene synthases (TPS), the key enzymes that accept prenylated diposhpate precursors as substrates and convert them to mono-, sesqui- and diterpenoid hydrocarbons. The promiscuity of TPSs results in a multitude of products that readily undergo further enzymatic and non-enzymatic modifications/rearrangements. The grapevine (*Vitis vinifera*) genome contains an over-represented TPS-encoding gene family with extensive gene duplications. Domestication of grapevine and selection of diverse cultivars resulted in plants with varying capacities to produce terpenoids, as is evident in wines from so-called aromatic varieties. Grapevine TPSs (*VviTPS*) are, therefore, largely studied for their role in modulating the flavour and aroma profiles of wines from aromatic (for example, Gewürztraminer and Muscat) or spicy (Shiraz) cultivars. These aromas are largely contributed by monoterpenes that impart floral aromas, while the sesquiterpene lactone rotundone has been linked to a spicy/peppery aroma. These results suggest that cultivar differences, in terms of TPS metabolism, are present, yet limited knowledge on the biological *in planta* function of grapevine terpenoids exists. In other plant species the terpenoids show important ecological functions in specialised metabolic processes that contribute to plant fitness and/or adaptability, for example, in mediating defence/stress responses or as important infochemcials for attraction. Grapevine flowers show promise for exploring these *in planta* roles due to increased transcriptional activity and volatile emissions in some cultivars.

In silico characterisation of the grapevine TPS gene family revealed that *VviTPS* expression was upregulated in flower developmental stages. In general, it was found that sesqui-TPS expressed at inflorescence development while mono-TPS expressed during flowering. The complexity of grapevine terpenoid metabolism was explored through co-expression analysis. Genes identified through expression networks revealed candidate genes encoding for enzymes that potentially modify terpenoids. These enzymes included cytochrome P450s, glycosidases and glycosyltransferases, all with reported functions that modify terpenoids in terms of bioactivity, solubility and volatility. Grapevine flower terpenoids were characterised for diverse wine cultivars by means of chemical analytical methods and showed remarkable differences. All cultivars produced sesquiterpenes as major volatiles with valencene, 7-epi-α-selinene, farnesenes and β-caryophyllene presence and absence affecting the volatilome of the different cultivars. The results presented show that the different cultivars differ in their capacity to produce certain terpenoids. Functional characterisation of putative TPS-encoding genes in a heterologous yeast expression system was utilised to demonstrate that cultivar-specific mutations affect *VViTPS* functionality. Aberrant mutations resulting in premature stop codons

and/or altered protein structures affecting the catalytic site of the TPS were prevalent. A novel gene encoding for a *E*-β-farnesene synthase was isolated for the cultivar Muscat D'Alexandria and was functionally linked to the cultivar's unique volatilome which was dominated by *E*-β-farnesene (~60%). The results reported here suggest that grapevine flowers have unique TPS-encoding genes as a result of independent cultivar selection pressures that influence their terpenoid volatilome.

The overrepresented grapevine TPS family is of great biological and economic importance for cultivar-specific traits. Multiple levels of both transcriptional and post-transcriptional regulation allows for great diversity in terpenoid metabolism. By studying specific organ/developmental stages it was shown that unique TPS-encoding genes are involved cultivar-specific terpenoid metabolism.

Opsomming

Terpene verteenwoordig een van die grootste natuurlike metaboliete en toon ongelooflike diversiteit in terme van chemiese struktuur en aktiwiteit. Hierdie diversiteit kan toegedien word aan die terpeen sintase (TPS) ensieme wat prenieldifosfaat voorlopers omskakel na die onderskeie substrate wat lei tot mono-, di- en seskwiterpene. TPSe het die vermoë om sonder onderskeid op te tree en kan dus enkele of menige terpene produseer vanaf ‘n substraat wat elk verander kan word deur beide ensiematies en nie-ensiematiese prosesse. Die wingerd (*Vitis vinifera*) genoom toon ‘n enorme *TPS* geen familie met uitgebreide duplisering van gene. Inburgering van wingerd het gelei tot ‘n diverse versameling van kultivars, elk met ‘n unieke vermoë om terpene to produseer, wat lei tot die onderskeid van die sogenaamde aromatiese kultivars. Om hierdie rede word wingerd TPSe (VviTPS) grootliks bestudeer vir hul vermoë om by te dra tot die blom (bv. Gewürztraminer en Muscat) of peper (bv. Shiraz) geur in wyn. Hierdie onderskeie geure kan toegdien word aan die mono- en sesquiterpene, onderskeidelik. Dit stel voor dat daar kultivar verskille bestaan in terme van terpeen metabolisme. Die *in planta* rol van wingerd terpene is egter nog nie wel bekend nie. In ander plante is dit al gewys dat die terpene belangrike ekologiese funksies verrig as sekondêre metaboliete wat lei tot ‘n verhoogde vermoë vir aanpassing, byvoorbeeld vir beskerming/verdediging of insek aanloking. Wingerde blomme lyk belowend om die begenoemde *in planta* funksies te ondersoek weens verhoogde geen transkripsie en vrystelling van vlugtige metaboliete wat al getoon is in sekere kultivars.

In silico karakterisering van die wingerd TPS geen familie het gewys dat *VviTPS* uitdrukking toeneem tydens blomvorming. Oor die algemeen lyk dit asof die seskwi-TPS gene uitgedruk word tydens die vroëe blomvorming stadiums terwyl die mono-TPS gene tydens die blom tydperk uitgedruk word. Die metaboliese prosesse betrokke by terpeen vorming was ondersoek deur middel van mede-uitdrukking ontleding. Dit het gelei tot die identifikasie van kandidaat gene wat lei tot byvoorbeeld sitokroom P450 en suiker hidrolase of transferase ensieme wat terpene kan modifiseer. Hierdie ensieme is bekend vir hul vermoë om die aktiwiteit en/of oplosbaarhied van terpene te verander. ‘n Versameling van wingerd kultivars was gekarakteriseer in terme van hul vlugtige terpene. Die nege kultivars geanaliseer het elk ‘n unieke terpeen profiel getoon met valencene, 7-epi-α-selinene, farnesene en β-caryophyllene as kenmerkende seskwiterpene. Die kultivars het verskil in terme van die konsentrasie en teenwoordigheid/afwesigheid van terpene wat dus toon dat kultivars verskil in hul vermoë om terpene to produseer. Belowende TPS gene was gevolglik geïsoleer en uitgedruk in ‘n heteroloë gis sisteem. Karakterisering van die gene het getoon dat mutasies spesifiek tot sekere kultivars verantwoordelik is vir ‘n verskil in ensiem funksie. Hierdie mutasies het gelei to ‘n vroëe stop kodon en/of ‘n verandering in die proteïen se aktiewe setel of struktuur. ‘n Unieke geen wat kodeer vir ‘n E-β-farnesene sintase geen was geïsoleer van die kultivar Muscat D'Alexandria en kon gekoppel word aan die unieke vlugtige profiel (*E*-β-farnesene as hoof

produk) van hierdie kultivar. Dit is dus moontlik dat unieke *TPS* gene voor geselekteer is tydens inburgering en gevolglik gelei het tot ‘n veranderde vermoë om terpene to produseer.

Die groot terpeen familie van wingerd is dus vir beide biologiese en ekonomiese doeleindes van belang in terme van unieke kultivar eienskappe. Die verskillende aspekte van geen reguleering en metaboliet modifikasie verleen dus aan wingerd die vermoë om menige terpene te produseer. Deur te kyk na speifieke orgaan/ontwikkelings stadia is daar getoon die *TPS* gene spesifiek tot ‘n kultivar verantwoordelik is vir uiteenlopende metaboliese moontlikhede.

This thesis is dedicated to the memory of Pieter Smit,
for teaching me how to see the humour in life,
and for providing me with the means to pursue my dreams.

Biographical sketch

Samuel Jacobus (Cobus) Smit was born in Polokwane, South Africa on 5 January 1990 and was raised in Kuilsriver. He matriculated at Paul Roos Gymnasium on 2008 and commenced his studies at the University of Stellenbosch in 2009 where he enrolled for a BSc-degree in Molecular Biology and Biotechnology. After graduating in 2012, he pursued postgraduate studies, obtaining a HonsBSc-degree in Wine Biotechnology *Cum Laude* in 2013.

Acknowledgements

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

- **Dr Philip Young**, who acted as my supervisor, provided guidance, advice and critical discussions that helped develop my abilities as a scientist;

- **Prof Melané Vivier**, who acted as my co-supervisor provided advice, valuable discussions and inputs throughout my studies.

- My wife, **Sarah**, for the unconditional love, continued prayers, support, and encouragement to pursue my dream.

- My **family**, for their faith, support and encouragement.

- My **colleagues** at the IWBT for their assistance and inputs in various fields.

- **Mr Malcolm Taylor** for his assistance with GC-MS analyses and method optimizations.

- **Prof Mario Pezzotti** and his group at the **University of Verona** for their assistance with various aspects of this study

- **Prof Ben Burger** for his assistance and knowledge

- **Mr Malcolm McLean** who helped design and manufacture custom glassware

- The **Institute for Wine Biotechnology** for affording me the opportunity to further my studies and financial support.

- The **Harry Crossley Foundation** and **Stellenbosch University** for financial assistance

- My **heavenly Father** for the capacity and grace to pursue my dream.

Preface

This thesis is presented as a compilation of 6 chapters.

Chapter 1 **General introduction and project aims**

Chapter 2 **Literature review**
Grapevine terpenoids: The ecological importance of cultivar diversity

Chapter 3 **Research results**
In silico characterisation of the grapevine TPS gene family in inflorescence and during flowering

Chapter 4 **Research results**
Analytical profiling of the grapevine flower volatilome from nine different wine cultivars

Chapter 5 **Research results**
Isolation and functional characterisation of grapevine flower sesquiterpene synthases

Chapter 6 **General discussion and conclusions**

Table of Contents

Chapter 1

General introduction and project aims

Chapter 1 – General introduction and project aims

1.1 Introduction

Plants employ various metabolic processes that result in specialised (secondary) metabolites that allow them to adapt to both abiotic and biotic environmental conditions, increasing their fitness. Secondary metabolism is less conserved with a lower evolutionary pressure compared to primary metabolism. This has resulted in large gene families producing metabolites with diverse functions in order to cope with diverse ecological challenges (Gershenzon & Dudareva, 2007; Degenhardt *et al.*, 2009). One such a gene family is the terpene synthases (TPSs) that produce the chemically and structurally diverse class of compounds known as terpenoids (Bohlmann *et al.*, 1998). TPSs are commonly found in both prokaryotic and eukaryotic organisms, often with promiscuous activity by synthesising multiple products from a single enzyme, resulting in an estimated 50,000 naturally occurring compounds (Hemmerlin *et al.*, 2012).

Terpene metabolism is also complex; synthesis of the compounds occurs by two compartmentalised pathways that result in the same C_5 prenyl diphosphate precursors (Bloch *et al.*, 1959; Rohmer *et al.*, 1993). These C_5 precursors serve as universal building blocks for the diverse terpenoid substrates. The diversity of TPSs facilitates complex carbocation cascades with the number of carbons and double bonds in the various terpenoid substrates being exploited to produce diverse cyclic and acyclic terpene structures that include the C_{10} mono-, C_{15} sesqui- and C_{20} diterpenoids (Bohlmann *et al.*, 1998; Tantillo, 2011; Miller & Allemann, 2012). Furthermore, terpenoids readily undergo non-enzymatic rearrangements and subsequent enzymatic modifications (for example, oxidation or glycosylation), further increasing the diversity of products (Cheng *et al.*, 2007; Yauk *et al.*, 2014). The diversity of plant terpenoids therefore provides the plant with a large battery of compounds with beneficial characteristics to react to its environment and even modulate it. The diversity and reactivity of this compound class, however, significantly complicates their study and accurate profiling and quantification with analytical techniques.

Terpenoids have been linked to numerous ecological interactions in plants with their role as “infochemicals” being one of the most important. Terpenoid infochemicals facilitate plant-plant, plant-organ and plant-insect interactions that include priming responses against incoming herbivores/pathogens and attraction of insects for pollination (Pichersky & Gershenzon, 2002; Heil & Silva Bueno, 2007). Their role as attractants of insects that act as bodyguards allows for indirect plant defence responses against invading herbivores/insects (Mumm & Dicke, 2010). Terpenoids are effectively deployed as antimicrobial or insect deterring chemicals (for example, as constituents of resin oils) and are often stored in glandular trichomes for rapid deployment against herbivores/invaders (Zulak & Bohlmann, 2010; Bleeker *et al.*, 2012). If not accumulating in specialised structures such as trichomes, these compounds can be temporarily stored in the lipid

phase or directly emitted to the atmosphere (Ormeño *et al.*, 2011).

Due to the large gene families they occur in, TPSs are typically characterised by means of heterologous expression systems in prokaryotic (for example, *Escherichia coli*) or eukaryotic (for example, *Saccharomyces cerevisiae*) hosts (Fäldt *et al.*, 2003; Lücker *et al.*, 2004; Martin & Bohlmann, 2004; Nguyen *et al.*, 2012). These systems provided insights in terms of TPS enzyme diversity and promiscuity. To understand terpenoid metabolism fully, it is important to link TPS gene expression and regulation, enzyme activity and metabolites formed with most/all of these aspects presenting problems in complex biological systems such as plants. The best studied TPS gene families of plants are those of *Arabidopsis,* where almost all of the 32 putative TPS-encoding genes have been characterised (Aubourg *et al.*, 2002). However, this gene family is relatively small compared to that of *Eucalyptus* with 172 TPS-encoding genes (113 putatively functional) or grapevine with 152 (69 putatively functional) (Martin *et al.*, 2010; Külheim *et al.*, 2015).

1.2 The grapevine *TPS* gene family

The grapevine genome (Jaillon *et al.*, 2007) revealed one of the largest *TPS* gene families of sequenced plants (Martin *et al.*, 2010). Of the 152 loci that showed similarity to known TPSs, 69 were predicted to encode for functional TPSs, typically resulting in mono- or sesquiterpenoids, 63 were probable pseudogenes and 20 were deemed to be partial TPSs. In grapevine (and wine) the mono- and sesquiterpenoids are typically studied for their role in modulating wine flavour and aroma, with the former linked to floral aromas in wine from aromatic cultivars (for example, Gewürztraminer, Muscat and Riesling) (Skinkis *et al.*, 2008; Martin *et al.*, 2012), whereas the latter contributes to the peppery aroma of red wines (for example, Shiraz) (Wood *et al.*, 2008). These aroma compounds can be directly contributed by the berries or produced as a result of the vinification process by which microbial enzymes liberate conjugated terpenoids or use certain precursors for *de novo* metabolism of terpenoids (Carrau *et al.*, 2005; Cramer *et al.*, 2014).

TPS enzyme activity has been shown to be influenced by the secondary structure of the enzyme, with single amino acid changes potentially affecting substrate specificity and/or functionality (Yoshikuni *et al.*, 2006; Garms *et al.*, 2012). It is thus likely that cultivar diversity within the *TPS* gene family is a result of independent evolutionary events. Different cultivars show unique terpene contributions in wines, suggesting that specific genes are potentially involved in cultivar-specific terpenoids.

Apart from uncovering the overrepresented grapevine *TPS* gene family, the functional (*in vitro*) analysis of a large number of the encoded proteins, as well as the documented influences of terpenoids on wine aromas, style and typicity, very little is known about the biological role and significance of grapevine terpenoids in the different organs they occur in. Most studies of grapevine terpenoids have focused on the grape berry as the main organ of study. With a

sequenced genome and with increased omics tools available, these resources should allow for greater insights on the genetic/biochemical/metabolic processes of grapevine in general (Martinez-Zapater *et al.*, 2010), but very few genome-wide studies currently reflect the transcriptional events in grapevine organs other than berries.

A gene atlas (whole-genome microarray) has, however, been established from 54 grapevine tissues/organs in different developmental stages and provides one of the few whole-plant grapevine resources (Fasoli *et al.*, 2012). The atlas was constructed using *V. vinifera* cv. Corvina and a later study, employing next generation RNA sequencing, provided evidence that some varietal differences/unique features exist for this cultivar compared to that of the reference genome (Pinot noir) (Venturini *et al.*, 2013).

1.3 Rationale and scope of the study

The main aim of this study was to functionally characterise flower-specific TPSs. A few studies have shown that grapevine flowers: (i) have unique TPS expression and volatiles profiles (Martin *et al.*, 2009; Matarese *et al.*, 2014); (ii) contain terpenoids that can act as infochemicals for insect attraction, suggesting some ecological functions (Tasin *et al.*, 2005, 2006); and (iii) display cultivar-differences in terms of terpenoids, but this aspect has not been comprehensively characterised and provide significant scope for further study.

With this in mind the following research questions were formulated:

- Is the grapevine flower volatilome different between selected wine cultivars?
- Are there flower-specific *TPS* genes in grapevine and when/where/how are they expressed?
- Can the grapevine flower volatilome of the different cultivars be linked with specific TPS-encoding genes and enzyme activities?

Grapevine presents a unique system for studying TPS diversity due to domestication events that enriched for cultivar diversity (This *et al.*, 2006). The grapevine gene atlas provides an important resource that can be exploited for new information on gene families in specific organs or developmental stages.

The methodological approach of the study therefore includes three main “tools”: firstly, an *in silico* evaluation of the *VviTPS* gene expression patterns in all organs represented in the gene atlas. By utilising bioinformatical approaches the *TPS* gene family was characterised in terms of organ-specific gene expression patterns. In order to understand the biochemical role of grapevine flower terpenoids a reverse genetical/biotechnological approach, combined with analytical profiling, was planned to investigate the terpenoid flower volatiles of different wine cultivars and to functionally characterise TPS-encoding genes that could influence/regulate these volatiles.

The specific objectives of this study were:

i. Computational characterisation (*in silico* mining) of the predicted TPS-encoding gene family in grapevine to identify *VviTPS* flower expression patterns as well as co-expressing genes implicated in terpenoid metabolism;

ii. Extraction, analysis and comparison of volatile terpenoids through chemical characterisation of flowers from different wine cultivars. The cultivars were selected based on aromatic descriptors associated with the cultivar-specific wines; and

iii. Functional characterisation of selected grapevine flower-specific TPS-encoding genes by means of gene isolation and heterologous yeast expression analysis in a reverse genetical approach.

The thesis includes a concise literature review on terpenoids in plants (chapter 2), followed by three research chapters providing the results obtained from the objectives outlined above (chapters 3-5). The thesis is concluded by chapter 6, to provide an integrated summary and future perspectives.

1.4 References

Aubourg S, Lecharny a., Bohlmann J. Genomic analysis of the terpenoid synthase (AtTPS) gene family of *Arabidopsis thaliana*. Mol Genet Genomics. 2002;267(6):730–45.

Bleeker PM, Mirabella R, Diergaarde PJ, Vandoorn A, Tissier A, Kant MR. Improved herbivore resistance in cultivated tomato with the sesquiterpene biosynthetic pathway from a wild relative. Proc Natl Acad Sci USA. 2012;109(49):20124–9.

Bloch K, Chaykin S, Phillips A., De Waard A. Mevalonic acid pyrophospahte and isopentenylpyrphosphate. J Biol Chem. 1959;234(10):2595–604.

Bohlmann J, Meyer-Gauen G, Croteau R. Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc Natl Acad Sci U S A. 1998; 95(8):4126–33.

Carrau FM, Medina K, Boido E, Farina L, Gaggero C, Dellacassa E, et al. *De novo* synthesis of monoterpenes by *Saccharomyces cerevisiae* wine yeasts. FEMS Microbiol Lett. 1;243(1):107–15.

Cheng A-X, Lou Y-G, Mao Y-B, Lu S, Wang L-J, Chen X-Y. Plant terpenoids : biosynthesis and ecological functions. 2007;49(2):179–86.

Cramer GR, Ghan R, Schlauch KA, Tillett RL, Heymann H, Ferrarini A, et al. Transcriptomic analysis of the late stages of grapevine (*Vitis vinifera* cv. Cabernet Sauvignon) berry ripening reveals significant induction of ethylene signaling and flavor pathways in the skin. BMC Plant Biol. 2014;14(1):370.

Degenhardt J, Köllner TG, Gershenzon J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry. 2009;70(15-16):1621–37.

Fäldt J, Martin D, Miller B, Rawat S, Bohlmann J. Traumatic resin defense in Norway spruce (*Picea abies*): methyl jasmonate-induced terpene synthase gene expression, and cDNA cloning and functional characterization of (+)-3-carene synthase. Plant Mol Biol. 2003;51(1):119–33.

Fasoli M, Dal Santo S, Zenoni S, Tornielli GB, Farina L, Zamboni A, et al. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell. 24(9):3489–505.

Garms S, Chen F, Boland W, Gershenzon J, Köllner TG. A single amino acid determines the site of deprotonation in the active center of sesquiterpene synthases SbTPS1 and SbTPS2 from *Sorghum bicolor*. Phytochemistry. 2012;75:6–13.

Gershenzon J, Dudareva N. The function of terpene natural products in the natural world. Nat Chem Biol. 2007;3(7):408–14.

Heil M, Silva Bueno JC. Within-plant signaling by volatiles leads to induction and priming of an indirect plant defense in nature. Proc Natl Acad Sci USA. 2007;104(13):5467–72.

Hemmerlin A, Harwood JL, Bach TJ. A raison d'être for two distinct pathways in the early steps of plant isoprenoid biosynthesis? Prog Lipid Res. 2012;51(2):95–148.

Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449(7161):463–7.

Külheim C, Padovan A, Hefer C, Krause ST, Köllner TG, Myburg AA, et al. The *Eucalyptus* terpene synthase gene family. BMC Genomics. BMC Genomics. 2015;16(1):450.

Lücker J, Bowen P, Bohlmann J. *Vitis vinifera* terpenoid cyclases: Functional identification of two sesquiterpene synthase cDNAs encoding (+)-valencene synthase and (-)-germacrene D synthase and expression of mono- and sesquiterpene synthases in grapevine flowers and berries. Phytochemistry. 2004;65:2649–59.

Martin DM, Aubourg S, Schouwey MB, Daviet L, Schalk M, Toub O, et al. Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. BMC Plant Biol. 2010;10:226.

Martin DM, Bohlmann J. Identification of *Vitis vinifera* (-)-α-terpineol synthase by *in silico* screening of full-length cDNA ESTs and functional characterization of recombinant terpene synthase. Phytochemistry. 2004;65(9):1223–9.

Martin DM, Chiang A, Lund ST, Bohlmann J. Biosynthesis of wine aroma: transcript profiles of hydroxymethylbutenyl diphosphate reductase, geranyl diphosphate synthase, and linalool/nerolidol synthase parallel monoterpenol glycoside accumulation in Gewürztraminer grapes. Planta. 2012;236(3):919–29.

Martin DM, Toub O, Chiang A, Lo BC, Ohse S, Lund ST, et al. The bouquet of grapevine (*Vitis vinifera* L. cv. Cabernet Sauvignon) flowers arises from the biosynthesis of sesquiterpene volatiles in pollen grains. Proc Natl Acad Sci U S A. 2009;106(17):7245–50.

Martinez-Zapater JM, Carmona MJ, Diaz-Riquelme J, Fernandez L, Lijavetzky D. Grapevine genetics after the genome sequence: challenges and limitations. Aust J Grape Wine Res. 2010;16:33–46.

Matarese F, Cuzzola A, Scalabrelli G, D'Onofrio C. Expression of terpene synthase genes associated with the formation of volatiles in different organs of *Vitis vinifera*. Phytochemistry. 2014

Miller DJ, Allemann RK. Sesquiterpene synthases: passive catalysts or active players? Nat Prod Rep. 2012;29(1):60–71.

Mumm R, Dicke M. Variation in natural plant products and the attraction of bodyguards involved in indirect plant defense. Can J Zool. 2010;88(7):628–67.

Nguyen T-D, MacNevin G, Ro D-K. *De novo* synthesis of high-value plant sesquiterpenoids in yeast. 1st ed. Methods Enzymol. Elsevier Inc.; 2012.

Ormeño E, Goldstein A, Niinemets Ü. Extracting and trapping biogenic volatile organic compounds stored in plant species. Trends Anal Chem. 2011;30(7):978–89.

Pichersky E, Gershenzon J. The formation and function of plant volatiles: perfumes for pollinator attraction and defense. Curr Opin Plant Biol. 2002;5(3):237–43.

Rohmer M, Knani TM, Simonin P, Sutter B, Sahmt H, Nationale E, et al. Isoprenoid biosynthesis in bacteria : a novel pathway for the early steps leading to isopentenyl diphosphate. 1993;524:517–24.

Skinkis PA, Bordelon BP, Wood K V. Comparison of monoterpene constituents in Traminette, Gewürztraminer, and Riesling winegrapes. Am J Enol Vitic. 2008;59(4):440–5.

Tantillo DJ. Biosynthesis via carbocations: theoretical studies on terpene formation. Nat Prod Rep. 2011;28(6):1035–53.

Tasin M, Anfora G, Ioriatti C, Carlin S, De Cristofaro A, Schmidt S, et al. Antennal and behavioral responses of grapevine moth *Lobesia botrana* females to volatiles from grapevine. J Chem Ecol. 2005;31(1):77–87.

Tasin M, Bäckman AC, Bengtsson M, Ioriatti C, Witzgall P. Essential host plant cues in the grapevine moth. Naturwissenschaften. 2006;93:141–4.

This P, Lacombe T, Thomas MR. Historical origins and genetic diversity of wine grapes. Trends Genet. 2006;22(9):511–9.

Venturini L, Ferrarini A, Zenoni S, Tornielli GB, Fasoli M, Dal Santo S, et al. *De novo* transcriptome characterization of *Vitis vinifera* cv. Corvina unveils varietal diversity. BMC Genomics; 2013;14(1):41.

Wood C, Siebert TE, Parker M, Capone DL, Elsey GM, Pollnitz AP, et al. From wine to pepper: Rotundone, an obscure sesquiterpene, is a potent spicy aroma compound. J Agric Food Chem. 2008 28;56(10):3738–44.

Yauk Y-K, Ged C, Wang MY, Matich AJ, Tessarotto L, Cooney JM, et al. Manipulation of flavour and aroma compound sequestration and release using a glycosyltransferase with specificity for terpene alcohols. Plant J. 2014.

Yoshikuni Y, Ferrin TE, Keasling JD. Designed divergent evolution of enzyme function. Nature. 2006;440(7087):1078–82.

Zulak KG, Bohlmann J. Terpenoid biosynthesis and specialized vascular cells of conifer defense. J Integr Plant Biol. 2010;52(1):86–97.

Chapter 2

Literature Review

Grapevine terpenoids: The ecological importance of cultivar diversity

Chapter 2 – Grapevine terpenoids: The ecological importance of cultivar diversity

2.1 Introduction

Grapevine (*Vitis* spp) is cultivated worldwide for wine, table grape, raisin and other minor related industries. The International Organisation of Vine and Wine (OIV) estimates that an area of 7, 573 million hectares was under vine in 2014(OIV, 2015), making it the most cultivated crop worldwide. In addition, most of these vines are of the *V. vinifera* species, which consist of an estimated 5, 000 cultivated varieties (cultivars). This diversity is due to almost eight thousand years of domestication, with most cultivars being maintained today through vegetative propagation to maintain the unique/desirable varietal characteristics (This *et al.*, 2006). The genetic diversity of different cultivars is evident in the berry (colour, shape and size), leaf shape, plant growth (vigour) and taste (Keller, 2010a).

This cultivar diversity facilitates a variety of wine styles that are, to a large extent, region or country-specific (for example: Italian Barolo and French Bordeaux wines). Cultivar diversity allows for specific wines that are usually defined by their flavour and aroma profiles which include: floral, tropical, herbaceous and spicy wines. These factors determine the typicality (typicity) of a wine and are influenced by the terroir (namely: soil, landscape and climate) as well as the vinification process (for example: yeast and bacteria involved in fermentation) (Deloire *et al.*, 2005). These complex interactions between grape juice and micro-organisms allow for endless possibilities for the final wine product.

Secondary metabolites fulfil various plant-specific functions to ensure adaptability to stress conditions, effective defence against attackers (microbial pathogens, herbivores and insects) and interactions with the environment (Bennett & Wallsgrove, 1994). One such class of secondary metabolites is the terpenoids that consists largely of the C_{10} mono-, C_{15} sesqui- and C_{20} diterpenoids. Only limited studies have addressed how terpenoids in grapevine are utilised by the plant, with the majority of research having been focused rather on applications for the wine industry in terms of flavour and aroma compounds (Boss, 2011). This review focuses on grapevine terpenoids with specific reference to genetic and chemical diversity, metabolism and the disparity in knowledge pertaining to their ecological functions.

2.2 Biosynthesis of terpenoids

The terpenoid (also referred to as isoprenoid) family is one of the most diverse with up to 50,000 compounds described in nature, most being synthesised by plants and bacteria (Hemmerlin *et al.*, 2012). Terpenoid metabolism is compartmentalised to two pathways, with the most prevalent aspects of biosynthesis illustrated in Figure 2.1.

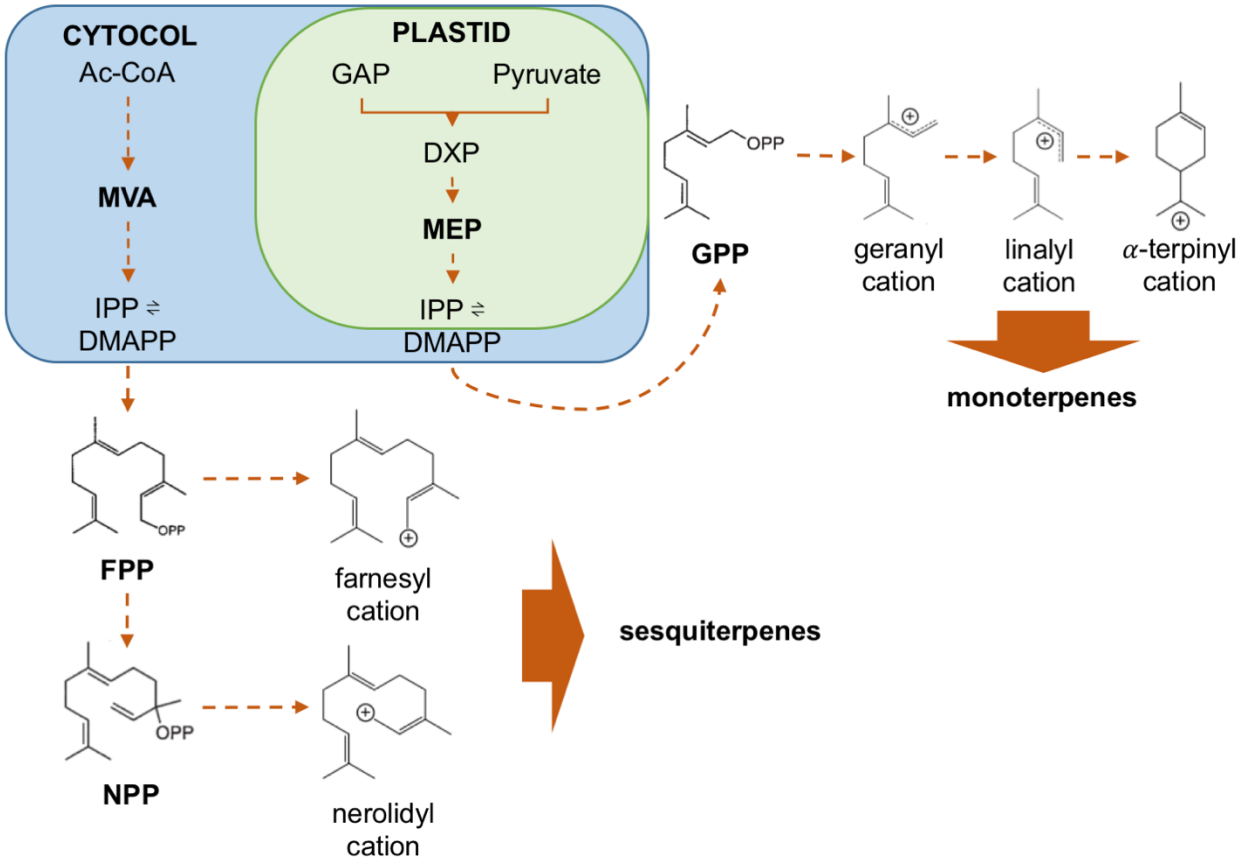


Figure 2.1 – Terpenoid biosynthesis proceeds via two pathways: cytosol localised MVA pathway where acetyl-coenzyme A (Ac-CoA) condensation results in mevalonate (MVA) and the plastidial MEP pathway where glyceraldehyde 3-phosphate (GAP) and pyruvate result in 1-deoxy-*D*-xylulose 5-phosphate (DXP) that is rearranged to form 2*C*-methyl-*D*-erythritol 4-phosphate (MEP). Both pathways result in ispentenyl disphosphate (IPP) and its allylic isomer dimethylallyl diphosphate (DMAPP) that serve as the building blocks to produce the monoterpenoid substrate geranyl diphosphate (GPP) and sesquiterpenoid substrates farnesyl diphosphate (FPP) and nerolidyl diphosphate (NPP). Cation intermediates from these respective substrates are illustrated.

The first pathway to be discovered was localised to the cytosol and involves the condensation of three acetyl-coenzyme A (Ac-CoA) molecules to mevalonate (MVA) and is commonly referred to as the MVA pathway (Bloch *et al.*, 1959; Pollard *et al.*, 1966). An alternative pathway was later discovered in bacteria where glyceralaldehyde 3-phosphate (GAP) and pyruvate serve as intermediates to yield 1-deoxy-*D*-xylulose 5-phosphate (DXP). DXP was subsequently subjected to intramolecular rearrangement and reduction to form 2*C*-methyl-*D*-erythritol 4-phosphate (MEP). Consequently, the pathway is currently referred to as the MEP- or DOXP pathway (Rohmer *et al.*, 1993; Hemmerlin *et al.*, 2012) and is localised to the plastids of plants (Lichtenthaler, 1999). Cross-talk between the MVA and MEP pathways is reported upon where under stressed conditions, greater adaptability is allowed for, as reviewed in Hemmerlin *et al.,* (2012).

Both the MVA and MEP pathways result in the C_5 intermediates isopentenyl diphosphate (IPP) and its allylic isomer dimethylallyl diphosphate (DMAPP) that serve as the universal building blocks for

the biosynthesis of terpenoids. The most basic terpene class is that of the hemiterpenes (C_5) that are formed from IPP and DMAPP. Prenyltransferases use these precursors to yield the prenyl diphosphate intermediates: geranyl diphosphate (GPP, C_{10}), farnesyl diphosphate (FPP, C_{15}) and geranylgeranyl diphosphate (GGPP, C_{20}). Terpene synthases (TPSs) catalyse the production of monoterpenes (C_{10}), sesquiterpenes (C_{15}) and diterpenes (C_{20}) from the respective substrates (reviewed by Tholl, 2006). FPP also serves as substrate for the larger C_{30} triterpenes while condensation of the two GGPPs produce the substrate for tertraterpene (C_{40}) biosynthesis (Bohlmann *et al.*, 1998).

2.2.1 TPS protein characteristics

TPSs share important structural features that can be used to identify TPS-encoding genes. The respective characteristics of TPS proteins are localised to two large domains, anmely: the N-terminal region (PF01397) and the C-terminal metal co-factor binding domain (PF03936) as defined in the Pfam database (Finn *et al.*, 2014). The presence/absence of certain motifs, plastidial transit peptides and sequence length results in seven TPS subfamilies (*TPS-a* through *TPS-g*) with grapevine TPSs (*VviTPSs*) being represented in five of these subfamilies (Bohlmann *et al.*, 1998; Martin *et al.*, 2010).

Generally it is found that the proteins of monoterpene synthases are longer than those of sesquiterpene synthases, due to the presence of an N-terminal plastidic transit peptide upstream of the first characteristic RRx_8W TPS motif in monoterpene synthases (Bohlmann *et al.*, 1998). All grapevine mono-TPS contain this motif with modified versions found in other types of *VviTPSs* (Martin *et al*., 2010). The N-terminal domain does not seem to be involved in catalysis but rather provides the scaffold necessary for the enzyme to function (such as: folding of the protein to allow for proper activity of the C-terminal active site) (Köllner *et al.*, 2004). The C-terminal domain contains the catalytic region of TPSs and is defined by two divalent metal binding motifs. The first is an aspartate-rich motif that is largely conserved in TPSs as DDxxD but modified versions like the EDxxD and DxxD motifs have been reported (reviewed in Chen *et al.*, 2011). These motifs along with the downstream NSE/DTE motif coordinate the binding of metal ions (divalent cations e.g. Mg^{2+} and Mn^{2+}) and water molecules to stabilise the active site for ionisation of the prenyl diphosphate precursors (Bohlmann *et al.*, 1998).

Predictions on TPS function using the primary structure are complicated due to the complexities of amino acid interactions in the active site. However, the three-dimensional (3D) structures of some TPSs have been determined which allows for new insights on the structure-function relationships of their activity and terpenoid biosynthesis. Residues in the active site as well as neighbouring residues affect the catalytic geometry of the enzyme and subtle residue changes in the C-terminal can result in altered substrate specificity or enzyme promiscuity, such as the ability to form single or multiple products (Greenhagen *et al.*, 2006; Yoshikuni *et al.*, 2006; Garms *et al.*, 2012). A comprehensive review by Gao *et al.*, (2012) addresses what has been learnt from 3D structures and the current

understanding of structure-function relationships of TPSs.

2.2.2 Terpenoid skeletal diversity

In order to understand the chemical diversity of terpenoids, it is crucial to understand the underlying enzymatic reactions and non-enzymatic mechanisms that facilitate the formation of diverse skeletal types. Skeletal diversity refers to the stereochemistry of terpenoids and how subtle changes in the chemical structure result in the different classes of terpenoids.

The condensation of IPP and DMAPP molecules results in the respective prenyl diphosphate precursors with subsequent ionisation and carbocation cascades (involving cyclizations, hydride shifts and rearrangements) which lead to carbocation intermediates. From GPP, the geranyl and linalyl cation intermediates result in cyclic monoterpenes, while the α-terpinyl cation leads to acyclic monoterpenes (Davis & Croteau, 2000). FPP, and its isomer nerolidyl diphosphate (NPP), are ionised to the respective farnesyl and nerolidyl cations that lead to the biosynthesis of sesquiterpenes (Figure 2.1). The larger carbon skeleton of FPP allows for an added double bond compared to GPP resulting in the greater structural diversity of sesquiterpenes. TPSs are thought to facilitate the mechanistic steps that lead to the carbocation intermediates as well as the subsequent reactions that result in a multitude of terpenoids (Bohlmann *et al.*, 1998; Davis & Croteau, 2000; Degenhardt *et al.*, 2009). Examples of selected carbon skeletons that can be used to classify the different types of sesquiterpenes are illustrated in Figure 2.2.

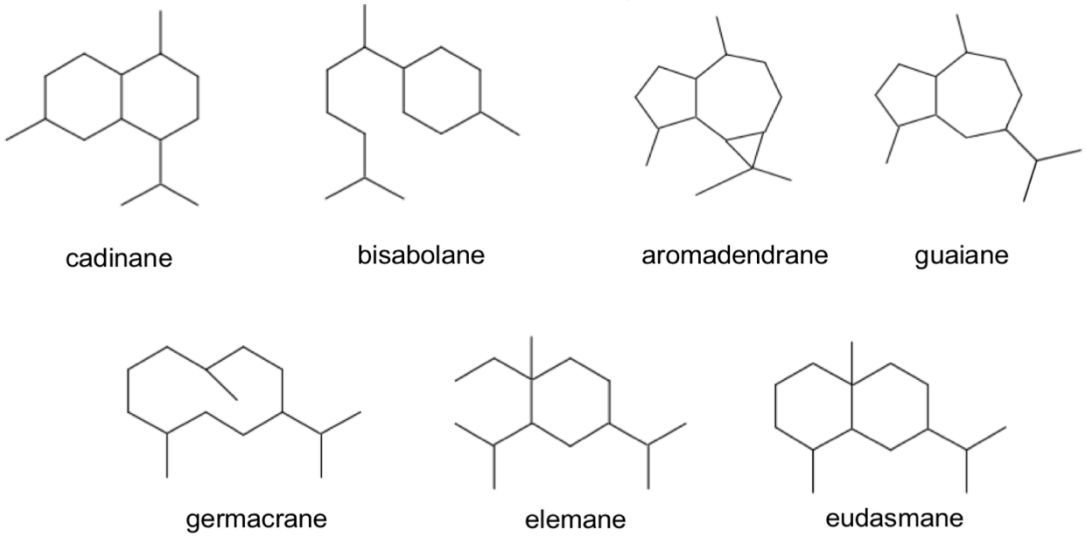


Figure 2.2 – Examples of sesquiterpenoid carbon skeletons used for classification of sesquiterpene types.

The diversity of sesquiterpenes is a result of a different reaction mechanism that proceeds from the farnesyl and nerolidyl cationic intermediates, as illustrated in Figure 2.3. From the farnesyl cation the two transoid *E*-humulyl and *E,E*-germacradienyl cations are formed with their associated sesquiterpenes (Figure 2.3 group A and B) (Davis & Croteau, 2000). The isomerisation of NPP from FPP is relatively slow and can be seen as a rate-limiting step for the synthesis of the derived products. However, once NPP is formed the subsequent reactions through the different cation intermediates proceed rapidly (Cane *et al.*, 1997; Miller & Allemann, 2012). The carbocation cascades via different

cation intermediates allow for the production of multiple sesquiterpenes through the action of a single TPS.

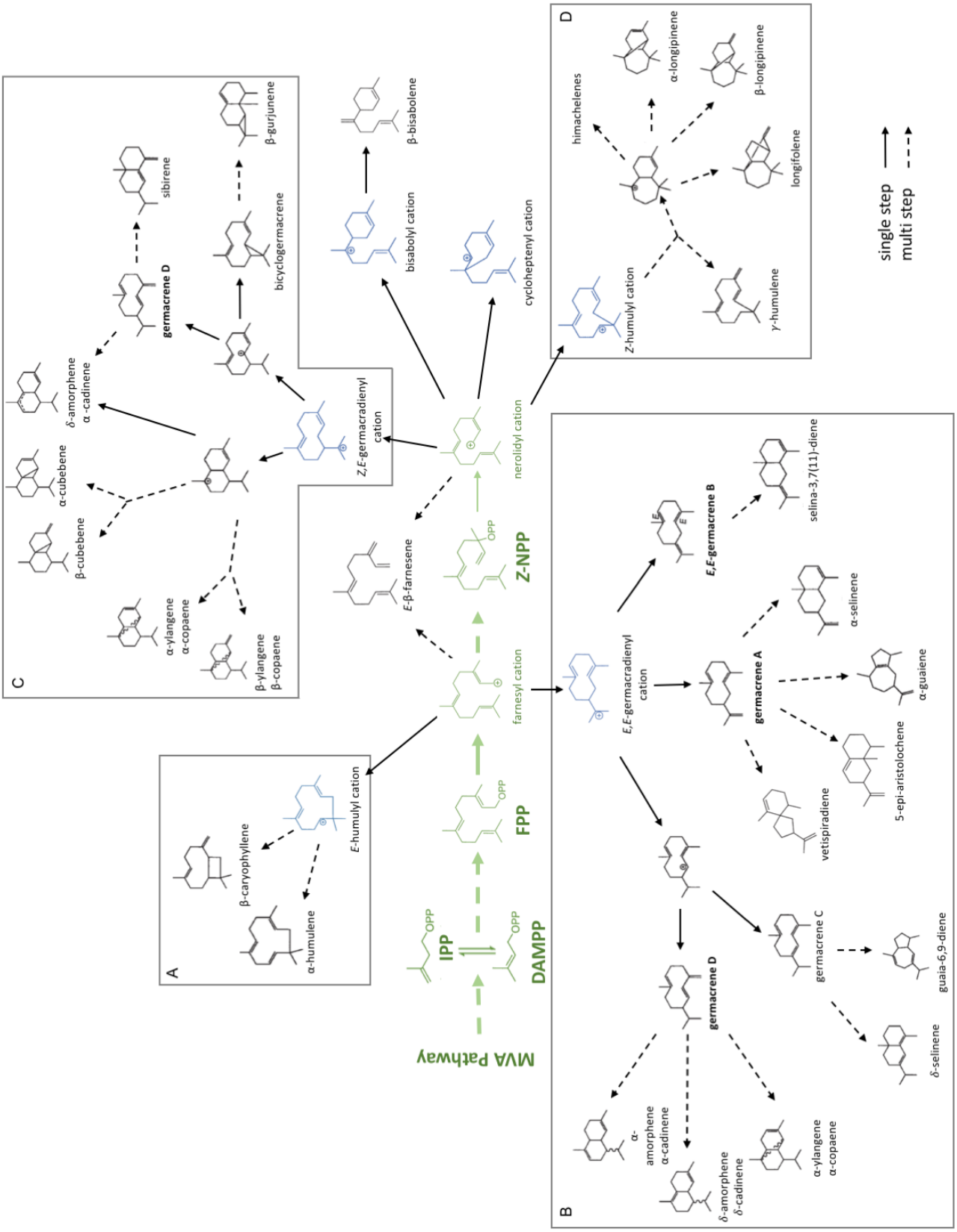


Figure 2.3 – The central MVA pathway that leads to the farnesyl and nerolidyl cation intermediates is shown in green. From these intermediates different carbocation cascades result in the different types of sesquiterpenoids (grouped A to D). Solid arrows indicate that a single mechanism is involved in the production of the proceeding structure, while dashed arrows show that multiple mechanistic steps are involved.

It was hypothesised by Bülow & Konig (2000) that the germacrenyl cations (group B and C of Figure 2.3) facilitate the biogenesis of a diverse group of sesquiterpenes from the germacrene A to E cation intermediates (Bülow & Konig, 2000). Germacrene A serves as intermediate for the production of elemane, guaiane and germacrane skeletal types (Figure 2.2) with germacrene E thought to result in the eudasmanes. When germacrene A is converted to germacrene E the production of selinanes is seen. NPP is converted to germacrene B and then germacrene D to allow for the biosynthesis of the cadinenes. Germacrene D can also be formed from FPP with A and C cations as intermediates (Bülow & Konig, 2000; Adia, 2005).

A TPS can be viewed as a driver towards one of the cationic intermediates with the inherent chemical properties of these intermediates being exploited by the enzyme to facilitate the production of multiple products. The examples used here serve only to illustrate sesquiterpene diversity and are by no means exhaustive. Numerous other mechanisms and intermediates are known that in some cases act in an organism/species specific manner. Tantillo, (2011) gives a comprehensive review on the carbocation cascades that lead to sesquiterpenes. A review by Wedler *et al.*, (2015) specifically explores the impact of these cascades on wine flavour and aroma.

2.2.3 Enzymatic and non-enzymatic terpenoid modifications amplify the diversity of biologically important compounds

Terpenoids are prone to modifications that alter their chemical properties in numerous ways which allow for increased diversity in terms of bio-activity as well as intricate mechanisms for the deployment of stored compounds. Some examples of these modifications are discussed to highlight the far reaching specialised metabolic processes that amplifies terpenoid diversity.

The chemical nature of terpenoids is in many cases toxic, which necessitates plant mechanisms to modify it in a manner that is not detrimental to the plant. One such an important mechanism is the conjugation of the terpenoids by the addition of a sugar moiety to facilitate solubility, transport and storage. The addition of a sugar moiety is catalysed by the action of glycosyltransferases, a ubiquitous group of enzymes that form multigene families in plants, resulting in terpene glycosides (conjugated terpenes) (Bönisch *et al.*, 2014). Many plant species have developed cellular structures known as glandular trichomes that serve as chemical reservoirs on the plant surface that release volatile organic compounds (VOCs) upon damage. *V. vinifera* is largely devoid of trichomes (i.e. lacks structures to store terpenoids) with only hair-like trichomes found on leaves to reduce water loss (Keller, 2010a). In order to be bioactive, stored terpene glycosides need to be liberated, usually by glycosidases that cleave the glycosyl moiety (Rivas *et al.*, 2013). Glycosidases can be deployed by the plant or in some cases are secreted by feeding insects, allowing for a localised release of toxic or deterring terpenoids (Morant *et al.*, 2008).

Cytochrome P450 (CYP) enzymes form one of the most comprehensive plant protein families (51,000 annotated in plants) having diverse biological/physiological functions. The CYP

monoxygenases have the ability to modify terpenoids by introducing oxygen into hydrophobic intermediates to increase solubility and alter biological activity (Nelson & Werck-Reichhart, 2011). The seemingly promiscuous ability to perform multiple oxidation reactions from diverse substrates allows for chemical tailoring (fine-tuning) of metabolites (for example, terpenoids) to be deployed in a concerted manner. This was shown in *Arabidopsis* where the enantiomers of linalool, produced by two different TPSs (*Arabidopsis* TPS10 and -14), were oxygenated/epoxidised by two distinct CYPs (*CYP71B31* and *CYP76C3*) that resulted in different, but overlapping (that is, produced by both CYPs), irregular terpenoids (homoterpenes) (Ginglinger *et al.*, 2013). Homoterpenes are a result of oxidative degradation of FPP and GPP by CYPs to yield the acyclic C_{11} and C_{16} terpenoid derivatives (Boland *et al.*, 1998). They are common constituents of herbivore-induced defence responses as well as important attractants in floral scents, therefore are thought to act as semiochemicals for various plant interactions (Tholl, 2006).

2.3 The ecological importance of plant terpenoids

Terpenoids are involved in diverse ecological roles that include attraction (Pichersky & Gershenzon, 2002), defence (Mumm & Dicke, 2010), stress responses and signalling (Frost *et al.*, 2007). The chemodiversity of terpenoids allows plants to adapt to various environmental stimuli by tailoring the release/synthesis of these products.

Plants can respond directly, by releasing compounds that affect or alter the behaviour and/or performance of an invading organism, or indirectly by attracting organisms that act as 'bodyguards' by predating on the herbivorous organism (Taiz & Zeiger, 2006; Kaplan *et al.*, 2008; Lucas-Barbosa *et al.*, 2011). A review by Mumm & Dicke (2010) discussed how different plant species regulate the metabolism of volatile organic compounds (VOCs) for the attraction of bodyguards. Mechanical wounding caused by herbivory or oviposition induces a myriad of pathways, such as the octadecanoic, MVA and MEP pathways that release specialised metabolites such as the fatty acid derived C_6 green leaf volatiles, monoterpenes and sesquiterpenes, respectively (Lucas-Barbosa *et al.*, 2011).

Certain plant species emit terpenoids constitutively in a diurnal pattern that generally coincides with the activities of certain insects. For example, *Arabidopsis* almost exclusively emits terpenoids from its flowers during the photoperiod while insect foraging and methyl jasmonate elicitors release terpenoids in a rhythmic manner that show a maximum in the day time (Chen *et al.*, 2003; Martin *et al.*, 2003; Miller *et al.*, 2005). These emissions are transcriptionally regulated with an upregulation of the biosynthetic pathways as well as specific TPS-encoding genes (Arimura *et al.*, 2004). The compartmentalisation of pathways along with finely tuned transcriptional regulation allows for broad yet rapid biosynthesis of terpenoids.

Plants need to respond to mutualistic and antogonistic insect interactions, and as a result, these

interactions have been studied extensively in flower parts due to the importance of insects for pollination. The attractive scents emitted by flowers, however, also attract insects that are not beneficial to the plant, hence the evolution of defencive volatiles that select for mutualists while repelling antagonists (Junker & Blüthgen, 2010a, 2010b). By tailoring the floral volatile composition plants can filter, or select, for specific interactions (Junker & Blüthgen, 2010b). The monoterpenoids and benzenoids, for example, are common constituents of floral volatiles for many species with the former (such as, linalool) thought to play a defensive role while the latter serves as an attractant (Junker *et al.*, 2011).

Plant volatiles facilitate various signalling interactions that can lead to a priming state (Frost *et al.*, 2008). Plant priming refers to a state of increased readiness that allows for a more rapid response to future infection or attack with volatile emissions serving as the cues for priming in local (at the site of attack) or systemic (on the same plant or between neighbouring plants) responses which facilitates both direct and indirect defences to herbivory (Baldwin *et al.*, 2006; Frost *et al.*, 2007; Heil & Silva Bueno, 2007).

These examples serve to illustrate the ecological importance of plant terpenoids. Comprehensive reviews by Moore *et al.* (2014) and Cheng *et al.* (2007) discuss these functions in more depth.

2.4 The role of terpenoids in grapevine

The grapevine TPS (*VviTPS*) gene family was characterised by Martin *et al.*, (2010) and until recently showed the greatest number (69) of putatively functional TPS genes. A study on the *Eucalyptus* TPS gene families for *E. grandis* and *E. globulus* revealed 113 and 106 putative genes, respectively (Külheim *et al.*, 2015). In addition to the putatively functional genes of grapevine, 20 partial and 63 probable genes were identified, resulting in 152 loci associated with *VviTPS* genes. This large number *VviTPS*-like genes is mostly due to extensive duplications of the genome which not only hampers the assembly of the genome, but also the functional characterisation of the gene family.

Phylogenetic analysis of the putative TPS-encoding genes shows that five of the seven known TPS-encoding gene subfamilies are present in grapevine. Of particular interest to this study are subfamilies *TPS-a*, *TPS-b* and *TPS-g*. *TPS-a* genes localised to chromosomes 18 and 19 encoding mainly of sesquiterpenes synthases, with the monoterpenes synthases of *TPS-b* localising to chromosomes 8, 12 and 13. Chromosome locations of genes belonging to *TPS-g* are largely unknown and consist mainly of the acyclic monoterpene synthases. Of the putative genes, 39 were functionally characterised through mainly *in vitro* bacterial assays (Martin *et al.*, 2010).

2.4.1 Terpenes and wine: Contributions to flavour and aroma

Grapevine terpenoids are largely studied in grape berries and wine due to their desired influence in wine flavour and aroma. These sought after contributions of terpenoids have resulted in a research

bias for grapevine where terpenoids are almost exclusively studied in grape berries and wine.

The transcriptional regulation of TPSs and subsequent volatiles in grape berries has been extensively studied for a wide range of cultivars (Kalua & Boss, 2009; Sweetman *et al.*, 2012; Matarese *et al.*, 2013; Cramer *et al.*, 2014). The varietal differences of wines are largely influenced by cultivar-specific terpenoids. The aomatic cultivars (for example, Muscat varieties, Riesling and Gewürztraminer) show a higher monoterpene content that can be associated with floral aromas, while the spicy or peppery aroma of Shiraz wines is attributed to the sesquiterpene rotundone (Siebert *et al.*, 2008; Skinkis *et al.*, 2008; Wood *et al.*, 2008; Vilanova *et al.*, 2013). Some of the reported terpenoids found in wine are a result of yeast (Drawert & Barton, 1978; King & Dickinson, 2000; Loscos *et al.*, 2007) and bacterial (Hernandez-Orte *et al.*, 2009) enzymes that liberate grape-derived conjugates. However, some yeast strains have been reported to produce terpenoids *de novo* in the absence of grape derived precursors (Carrau *et al.*, 2005), suggesting that wine terpenoids are not solely contributed by the plant. For a recent review on wine aroma compounds see Robinson *et al.* (2014).

2.4.2 Ecological roles of grapevine terpenoids

The ecological role of grapevine terpenes has largely been inferred from other species. Literature on *VviTPS* genes and volatiles that do not relate to grape berries and/or wine is limited and shows great disparity in terms of their function in the grapevine. Here a discussion is made on some of the literature that report on possible ecological roles of grapevine terpenoids.

Matarese et al. (2014) analysed *VviTPS* expression and volatiles of twelve different tissue types for the cultivar Moscato bianco. An organ-specific expression pattern for some of the *VviTPS* genes targeted was revealed, with a distinct terpene expression and volatile profile for flower organs. The authors targeted 23 *VviTPS* genes, as characterised by Martin *et al.* (2010), and were able to link some of the *in vitro* volatiles to those detected in different plant organs. An earlier study by Martin *et al.* (2009) reveals that the majority of volatiles of Cabernet Sauvignon flowers are sesquiterpenes that predominantly accumulate in the pollen grains.

Terpenoids are implicated with stress-related responses, with a potential anti-oxidant role reported for the leaves of *in vitro* cultured *V. vinifera* L cv. Malbec in response to UV-B radiation (Gil *et al.*, 2012). Chardonnay cell cultures challenged with the fungal pathogen *Phaecremonium parasiticum* show increased TPS activity through *de novo* synthesis with the sesquiterpene nerolidol increasing in response to the fungal elicitor (Escoriaza *et al.*, 2013). Chemical elicitors are commonly used to simulate stress or defence responses in plants, with a few studies showing success in grapevine (Belhadj *et al.*, 2006; D'Onofrio & Boss, 2010; Gómez-Plaza *et al.*, 2012). Elicitation with methyl jasmonate of detached *V. vinifera* cv. Moria Muskat leaves shows a transcriptional induction of the MEP and MVA pathways leading to compartmentalised production of mono- and sesquiterpene, respectively (Hampel *et al.*, 2005). D'Onofrio *et al.* (2009), in their study, show a similar response

using Cabernet Sauvignon cell cultures where jasmonate elicitors caused upregulation of genes controlling the flux through the terpenoid biosynthetic pathways and resulting in elevated sesquiterpene (β-caryophyllene, α-cubebene, α-copaene and δ-cadinene) volatiles.

Very few studies have investigated the interactions between grapevines and insects. In the case of the grapevine berry moths it has been shown that grapevine terpenoids resemble the infochemicals released by the moth. Reconstituted mixtures of the terpenoids released by grapevine are attractants to these moths and grapevine berry moth infestations are common in Europe, causing extensive plant damage due to oviposition (Tasin *et al.*, 2005, 2006; Anfora *et al.*, 2009). Furthermore, studies on grapevine rootstocks show that phylloxera *(Daktulosphaira vitifoliae*) infestations caused the MVA and MEP pathways to be upregulated, with the sesquiterpene β-caryophyllene being emitted as one of the major volatiles (Lawo *et al.*, 2011).

2.5 Grapevine flower terpenoids

Before the availability of the grapevine genome, sesquiterpenes had been associated with flowering with Lücker *et al.* (2004) isolating the sesquiterpene synthases *VvVal* and *VvGerD*, which produced valencene and germacrene D respectively. Transcript abundance reveals a low abundance in flower-related tissues for *VvVal* with increased transcripts occurring in late berry stages. Later investigations on flowers by Martin *et al.*, (2009) reveal that flowers of Cabernet Sauvignon produce high levels of sesquiterpenes in flowers at the inflorescence development stages. Expression analysis of flowers reveals that some of the *VviTPSs* characterised by Martin *et al.* (2010) are expressed in inflorescences and during flower bloom (Matarese *et al.*, 2014). Furthermore, it is shown in the *V. vinifera* cv. Corvina gene expression atlas (Fasoli *et al.*, 2012), that grapevine flowers have unique expression patterns that correlate with *VviTPSs*. The phenological processes involved with flowering are complex with intricate expression and hormonal regulation directing development. These findings suggest that the secondary metabolic processes involved in terpenoid production play an important role in flower organogenesis, however, the reasons therefore are yet undiscovered. The specific role of *VviTPSs* and the transcriptional switch(es) from inflorescence development to flowering (and the consequent impact on grapevine phenology) are still poorly understood.

2.5.1 The phenology of grapevine: Growth cycles and flowering

The annual phenology of grapevine can be divided into two cycles: vegetative and reproductive, as illustrated in Figure 2.4. This biennial cycle is complex and, therefore, only the most prevalent aspects of the phenology are discussed.

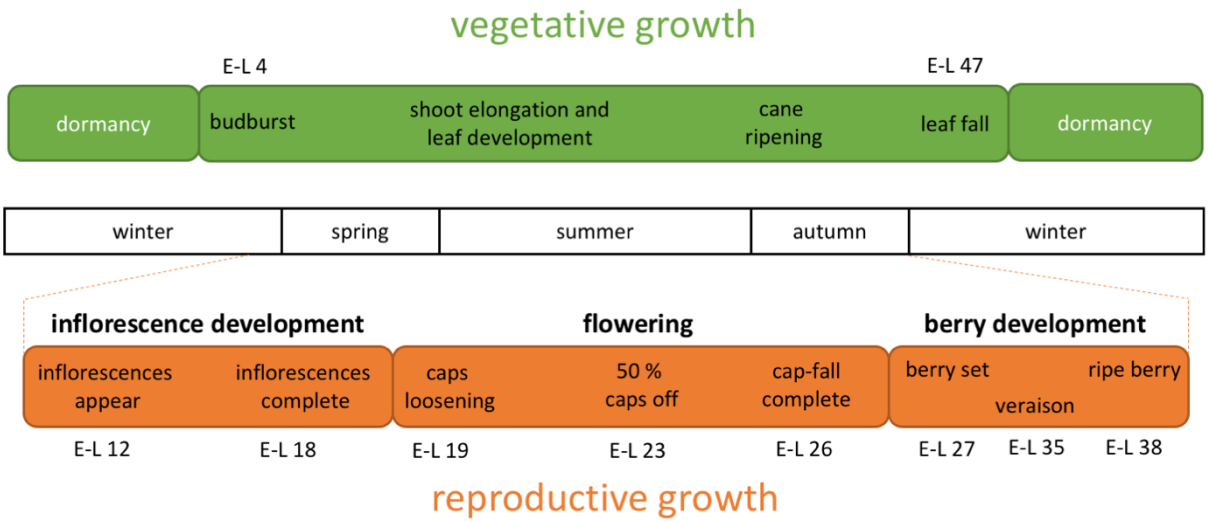


Figure 2.4 – The biennial growth cycle of grapevine. Vegetative growth is shown in green with reproductive growth in orange. Where applicable the associated E-L stage is indicated.

The developmental stages of grapevine is described in the modified Eichorn and Lorenz (E-L) system (Coombe, 1995). In brief, vegetative growth commences when dormancy is broken, with the bud starting to swell until the first leaf tissue becomes visible at budburst (E-L 4). Vegetative growth proceeds through stages of the development of leaves, elongation of the shoots and finally cane ripening. Cane ripening (browning of the cane) usually signifies the end of the vegetative growth phase with shoot elongation/leaf development ceasing. At this point the leaves start to senesce and the vine prepares for winter dormancy by storing reserves (E-L 41 to 47) (Galet, 2000).

After overwintering and a few days after budburst, the reproductive growth cycle starts. Differentiation of primary bud primordia (also known as the apical meristem) results in shoots and leaves with secondary buds (axillary meristems) forming. The axillary meristems can differentiate into lateral shoots, tendrils or inflorescences depending on the environmental conditions. The term inflorescence typically describes a cluster of flowers arranged on a single branch or stem (in grapevine a bunch of rachis). If conditions are inducive for flowering, axillary meristems differentiate into inflorescences instead of tendrils or shoots. The initiation of these organs is hormonally controlled by the antagonistic regulation of cytokinin and gibberellin, with the former promoting inflorescences while the latter promotes tendrils (Mullins *et al.*, 1992; Keller, 2010a; Díaz-Riquelme *et al.*, 2014). The inflorescences usually develop on the lateral shoots with the pattern of meristem differentiation to leaf, tendril and inflorescence being characteristic to each cultivar.

Grapevine flowering can be divided into three stages, namely: initiation of inflorescences, flower organogenesis and anthesis (flower bloom). The E-L system separates the inflorescence period and flower organogenesis stages (E-L 12 to E-L 18) from the last stage (Coombe, 1995). The E-L system uses the loosening of the calyptra (flower cap) at E-L 19 to describe the start of anthesis (flower opening or bloom) which continues to E-L 26 when all the calyptra have fallen and precedes fruit set. The calyptra can be considered the equivalent to flower petals in other flowering plants. Most

commercial *V. vinifera* cultivars have been selected for self-pollinating hermaphroditic flowers. Figure 2.5 illustrates the grapevine flower reproductive organs, namely: male stamens (consisting of the anthers and filaments) and the female pistil (consisting of the stigma, style and ovary) (Mullins *et al.*, 1992; Keller, 2010b).

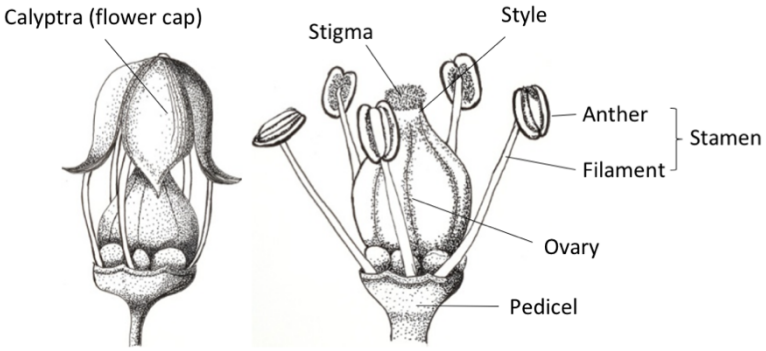


Figure 2.5 – Schematic representations of grapevine flowers and the respective flower organs

2.5.2 Flower TPSs in the context of grapevine omics

Grapevine 'omics' is expanding rapidly with various platforms being utilised to study the complex genetic and biochemical processes of *Vitis* spp. The availability of the *Vitis* genome serves as a catalyst for comprehensive studies utilising whole genome microarrays (Fasoli *et al.*, 2012; Cramer *et al.*, 2014; Díaz-Riquelme *et al.*, 2014) and RNA sequencing, RNA-seq, (Zenoni *et al.*, 2010; Sweetman *et al.*, 2012; Venturini *et al.*, 2013). However, the majority of these studies focused on grape berries at various ripening stages. Fasoli *et al.* (2012) and Díaz-Riquelme *et al.* (2014) give limited insights into the transcriptional processes involved in other grapevine organs. Correlations between transcript abundance and actual protein abundance, however, show great disparity as highlighted by Ghan *et al.* (2015). The numerous platforms available provide valuable insight in terms of protein abundance and gene expression when viewed in isolation. However, establishing a direct link between transcript abundance and protein/metabolites has so far been impossible. Extensive computational power and robust bioinformatics are required to move from correlation to causation in complex biological systems.

The aforementioned studies demonstrate that grapevine berries have complex metabolic networks that are reflected at both the transcriptional and metabolite level. The phenotypic variations observed between cultivars as a response to their genotype and environment undoubtedly translate to other organs, such as flowers, that undergo extensive transcriptional regulation during development. The complexity of organ specific metabolism is highlighted by the sub-cellular differences observed in berries where transcriptional and metabolic differences are observed between the skin, pulp and seeds of berries (Cramer *et al.*, 2014). Studies on specific gene families like those involved in carotenoid (Young *et al.*, 2012), terpenoid (Martin *et al.*, 2010), and stilbene metabolism (Vannozzi *et al.*, 2012) as well transcription factor families (Sweetman *et al.*, 2012) show the complexity of regulation when the associated metabolic processes are viewed in isolation. The

gene atlas provides evidence of sub-cellular transcription that is enforced by the metabolite profiles of berries. The extent of these processes in other organs will undoubtedly provide valuable insight but currently the omics resources necessary to fully explore these differences are lacking compared to the data available for berries.

2.6 Summary

It is not yet clear why grapevine shows such a large and duplicated TPS gene family or how (or even if) it is utilised by the plant. The terpenoid metabolic pathways and promiscuous nature of TPSs allows for the deployment of a variety of compounds for defence or as attractant infochemicals. The reported roles in response to pathogen attacks show great promise due to the relatively high disease susceptibility seen in grapevine. Numerous terpenoids are reported to be antimicrobial, but this role has not been established in grapevine. Expanding our understanding of grapevine terpenoids will allow for new insights into which TPSs can act as molecular markers for breeding programmes.

Other plant species have shown diverse functions for terpenoids but no species offers the genetic diversity we see in grapevine. The thousands of cultivars for grapevine were subjected to unique selection pressures during domestication that undoubtedly resulted in cultivar-specific functionalisation of the volatile terpenes. Studies on wine provide some insight into this, with monoterpenoids linked to floral aromas and sesquiterpenes to pepper notes in Shiraz wines. The actual underlying genetic factors that contribute to these differences are still largely unknown. Grapevine therefore presents a unique opportunity in terms of genetic and cultivar diversity to study the nuances that affect the cultivar-specific volatilome.

2.7 References

Adia AM. Isolation and structure elucidation of sesquiterpenoids from the essential oils of some liverworts (*Hepaticae*). University of Hamburg; 2005.

Anfora G, Tasin M, de Cristofaro A, Ioriatti C, Lucchi A. Synthetic grape volatiles attract mated *Lobesia botrana* females in laboratory and field bioassays. J Chem Ecol. 2009;35(9):1054–62.

Arimura G, Huber DPW, Bohlmann J. Forest tent caterpillars (*Malacosoma disstria*) induce local and systemic diurnal emissions of terpenoid volatiles in hybrid poplar (*Populus trichocarpa* × *deltoides*): cDNA cloning, functional characterization, and patterns of gene expression of (−)-germacr. Plant J. 2004;37(4):603–16.

Baldwin IT, Halitschke R, Paschold A, von Dahl CC, Preston CA. Volatile signaling in plant-plant interactions: 'talking trees' in the genomics era. Science. 2006;311(5762):812–5.

Belhadj A, Saigne C, Telef N, Cluzet S, Bouscaut J, Corio-Costet M-F, et al. Methyl jasmonate induces defense responses in grapevine and triggers protection against *Erysiphe necator*. J Agric Food Chem. 2006;29;54(24):9119–9125.

Bennett RN, Wallsgrove RM. Secondary metabolites in plant defence mechanisms. New Phytol. 1994;127:617–33.

Bloch K, Chaykin S, Phillips A., De Waard A. Mevalonic acid pyrophospahte and isopentenylpyrphosphate. J Biol Chemsitry. 1959;234(10):2595–604.

Bohlmann J, Meyer-Gauen G, Croteau R. Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc Natl Acad Sci U S A. 1998;95(8):4126–33.

Boland W, Gibler A, Gilbert M, Feng Z. Biosynthesis of and homoterpenes in higher plants; stereo-chemistry of the C-C-bond cleavage reaction. Tetrahedron. 1998;54(49):14725–36.

Bönisch F, Frotscher J, Stanitzek S, Rühl E, Wüst M, Bitz O, et al. A UDP-glucose:monoterpenol glucosyltransferase adds to the chemical diversity of the grapevine metabolome. Plant Physiol. 2014;165(2):561–81.

Boss PK. Vines to wine – linking fruit quality to wine flavour and aroma. CSIRO, Australia. 2011.

Bülow N, Konig W. The role of germacrene D as a precursor in sesquiterpene biosynthesis: investigations of acid catalyzed, photochemically and thermally induced rearrangements. Phytochemistry. 2000 Sep;55(2):141–68.

Cane DE, Chiu HT, Liang PH, Anderson KS. Pre-steady-state kinetic analysis of the trichodiene synthase reaction pathway. Biochemistry. 1997;36(96):8332–9.

Carrau FM, Medina K, Boido E, Farina L, Gaggero C, Dellacassa E, et al. *De novo* synthesis of monoterpenes by *Saccharomyces cerevisiae* wine yeasts. FEMS Microbiol Lett. 2005;243(1):107–15.

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. 2011;66(1):212–29.

Chen F, Tholl D, D'Auria JCD, Farooq A, Pichersky E, Gershenzon J. Biosynthesis and emission of terpenoid volatiles from *Arabidopsis* flowers. Plant Cell. 2003;15:481–494.

Cheng A-X, Lou Y-G, Mao Y-B, Lu S, Wang L-J, Chen X-Y. Plant terpenoids : biosynthesis and ecological functions. 2007;49(2):179–86.

Coombe BG. Adoption of a system for identifying grapevine growth stages. Aust J Grape Wine Res. 1995;1:100–10.

Cramer GR, Ghan R, Schlauch K a, Tillett RL, Heymann H, Ferrarini A, et al. Transcriptomic analysis of the late stages of grapevine (*Vitis vinifera* cv. Cabernet Sauvignon) berry ripening reveals significant induction of ethylene signaling and flavor pathways in the skin. BMC Plant Biol. 2014;14(1):370.

D'Onofrio C, Boss PK. Optimization of aroma production in grape cell suspensions induced by chemical elicitors. Quad di Sci Vitic ed Enol. 31. Torino; 2010. p. 55–58.

D'Onofrio C, Cox A, Davies C, Boss PK. Induction of secondary metabolism in grape cell cultures by jasmonates. Funct Plant Biol. 2009;36(4):323–38.

Davis EM, Croteau R. Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. Top Curr Chem. 2000;209:53–95.

Degenhardt J, Köllner TG, Gershenzon J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry. Elsevier Ltd; 2009;70(15-16):1621–37.

Deloire A, Vaudour E, Carey V, Bonnardot V, Van Leeuwen C. Grapevine responses to terroir: a global approach. J Int des Sci la Vigne du Vin. 2005;39(4):149–62.

Díaz-Riquelme J, Martínez-Zapater JM, Carmona MJ. Transcriptional analysis of tendril and inflorescence development in grapevine (*Vitis vinifera* L.). PLoS One. 2014;9(3):e92339.

Drawert F, Barton H. Biosynthesis of flavor compounds by microorganisms. 3. Production of monoterpenes by the yeast *Kluyveromyces lactic*. J Agric Food Chem. 1978;26(3):765–6.

Escoriaza G, Sansberro P, García-Lampasona S, Gatica M, Bottini R, Piccoli P. *In vitro* cultures of *Vitis vinifera* L. cv. Chardonnay synthesize the phytoalexin nerolidol upon infection by *Phaeoacremonium parasiticu*m. Phytopathol Mediterr. 2013;52(2):289–97.

Fasoli M, Dal Santo S, Zenoni S, Tornielli GB, Farina L, Zamboni A, et al. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell. 2012;24(9):3489–505.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2014;42(D1):D222–30.

Frost CJ, Appel HM, Carlson JE, De Moraes CM, Mescher MC, Schultz JC. Within-plant signalling via volatiles overcomes vascular constraints on systemic signalling and primes responses against herbivores. Ecol Lett. 2007;10:490–498.

Frost CJ, Mescher MC, Carlson JE, De Moraes CM. Plant defense priming against herbivores: getting ready for a different battle. Plant Physiol. 2008;146(3):818–24.

Galet P. General Viticulture. English ed. Collect. Avenir Oenologie. Chaintré, France : Oenoplurimédia; 2000.

Gao Y, Honzatko RB, Peters RJ. Terpenoid synthase structures: a so far incomplete view of complex catalysis. Nat Prod Rep. 2012;29(10):1153–75.

Garms S, Chen F, Boland W, Gershenzon J, Köllner TG. A single amino acid determines the site of deprotonation in the active center of sesquiterpene synthases SbTPS1 and SbTPS2 from *Sorghum bicolor*. Phytochemistry. Elsevier Ltd; 2012;75:6–13.

Ghan R, Van Sluyter SC, Hochberg U, Degu A, Hopper DW, Tillet RL, et al. Five omic technologies are concordant in differentiating the biochemical characteristics of the berries of five grapevine (*Vitis vinifera* L.) cultivars. BMC Genomics; 2015;16(1):946.

Gil M, Pontin M, Berli F, Bottini R, Piccoli P. Metabolism of terpenes in the response of grape (*Vitis vinifera* L.) leaf tissues to UV-B radiation. Phytochemistry. 2012;77:89–98.

Ginglinger J-F, Boachon B, Höfer R, Paetz C, Köllner TG, Miesch L, et al. Gene coexpression analysis reveals complex metabolism of the monoterpene alcohol linalool in *Arabidopsis* flowers. Plant Cell. 2013;25(11):4640–57.

Gómez-Plaza E, Mestre-Ortuño L, Ruiz-García Y, Fernández-Fernández JI, López-Roca JM. Effect of benzothiadiazole and methyl jasmonate on the volatile compound composition of *Vitis vinifera* L. Monastrell grapes and wines. Am J Enol Vitic. 2012;63(3):394–401.

Greenhagen BT, O'Maille PE, Noel JP, Chappell J. Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. Proc Natl Acad Sci U S A. 2006;103(26):9826–9831.

Hampel D, Mosandl A, Wüst M. Induction of *de novo* volatile terpene biosynthesis via cytosolic and plastidial pathways by methyl jasmonate in foliage of *Vitis vinifera* L. J Agric Food Chem. 2005;53:2652–7.

Heil M, Silva Bueno JC. Within-plant signaling by volatiles leads to induction and priming of an indirect plant defense in nature. Proc Natl Acad Sci U S A. 2007;104(13):5467–72.

Hemmerlin A, Harwood JL, Bach TJ. A raison d'être for two distinct pathways in the early steps of plant isoprenoid biosynthesis? Prog Lipid Res. 2012;51(2):95–148.

Hernandez-Orte P, Cersosimo M, Loscos N, Cacho J, Garcia-Moruno E, Ferreira V. Aroma development from non-floral grape precursors by wine lactic acid bacteria. Food Res Int. 2009;42:773–781.

Junker RR, Blüthgen N. Dependency on floral resources determines the animals' responses to floral scents. Plant Signal Behav. 2010a;5(8):1014–6.

Junker RR, Blüthgen N. Floral scents repel facultative flower visitors, but attract obligate ones. Ann Bot. 2010b;105(5):777–82.

Junker RR, Gershenzon J, Unsicker SB. Floral odor bouquet loses its ant repellent properties after inhibition of terpene biosynthesis. J Chem Ecol. 2011;37(12):1323–31.

Kalua CM, Boss PK. Evolution of volatile compounds during the development of cabernet sauvignon grapes (*Vitis vinifera* L.). J Agric Food Chem. 2009;57(9):3818–30.

Kaplan I, Halitschke R, Kessler A, Sardanelli S, Denno RF. Constitutive and induced defenses to herbivory in above- and below-ground plant tissues. Ecology. 2008;89(2):392–406.

Keller M. Botany and Anatomy. Sci Grapevines Anat Physiol. Elsevier; 2010a. p. 1–47.

Keller M. Phenology and Growth Cycle. Sci Grapevines Anat Physiol. Elsevier; 2010b. p. 49–83.

King A, Dickinson JR. Biotransformation of monoterpene alcohols by *Saccharomyces cerevisiae*, *Torulaspora delbrueckii* and *Kluyveromyces lactis*. Yeast. 2000;16(6):499–506.

Köllner TG, Schnee C, Gershenzon J, Degenhardt J. The variability of sesquiterpenes emitted from two *Zea mays* cultivars is controlled by allelic variation of two terpene synthase genes encoding stereoselective multiple product enzymes. Plant Cell. 2004;16:1115–1131.

Külheim C, Padovan A, Hefer C, Krause ST, Köllner TG, Myburg AA, et al. The *Eucalyptus* terpene synthase gene family. BMC Genomics; 2015;16(1):450.

Lawo NC, Weingart GJF, Schuhmacher R, Forneck A. The volatile metabolome of grapevine roots: First insights into the metabolic response upon phylloxera attack. Plant Physiol Biochem. 2011;49(9):1059–1063.

Lichtenthaler HK. The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. Ann Rev Plant Physiol Plant Mol Biol. 1999;50(1):47–65.

Loscos N, Hernandez-Orte P, Cacho J, Ferreira V. Release and formation of varietal aroma compounds during alcoholic fermentation from nonfloral grape odorless flavor precursors fractions. J Agric Food Chem. 2007;55(16):6674–84.

Lucas-Barbosa D, van Loon JJ a, Dicke M. The effects of herbivore-induced plant volatiles on interactions between plants and flower-visiting insects. Phytochemistry. 2011;72(13):1647–54.

Lücker J, Bowen P, Bohlmann J. Vitis vinifera terpenoid cyclases: Functional identification of two sesquiterpene synthase cDNAs encoding (+)-valencene synthase and (-)-germacrene D synthase and expression of mono- and sesquiterpene synthases in grapevine flowers and berries. Phytochemistry. 2004;65:2649–59.

Martin D, Gershenzon J, Bohlmann J. Induction of volatile terpene biosynthesis and diurnal emission by methyl jasmonate in fiolage of Norway spruce. Plant Physiol. 2003;132:1586–1599.

Martin DM, Aubourg S, Schouwey MB, Daviet L, Schalk M, Toub O, et al. Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. BMC Plant Biol. 2010;10:226.

Martin DM, Toub O, Chiang A, Lo BC, Ohse S, Lund ST, et al. The bouquet of grapevine (*Vitis vinifera* L. cv. Cabernet Sauvignon) flowers arises from the biosynthesis of sesquiterpene volatiles in pollen grains. Proc Natl Acad Sci U S A. 2009;106(17):7245–50.

Matarese F, Cuzzola A, Scalabrelli G, D'Onofrio C. Expression of terpene synthase genes associated with the formation of volatiles in different organs of *Vitis vinifera*. Phytochemistry. Elsevier Ltd; 2014 Jul 7;

Matarese F, Scalabrelli G, D'Onofrio C. Analysis of the expression of terpene synthase genes in relation to aroma content in two aromatic *Vitis vinifera* varieties. Funct Plant Biol. 2013;40(6):552.

Miller B, Madilao LL, Ralph S, Bohlmann J. Insect-induced conifer defense. White pine weevil and methyl jasmonate induce traumatic resinosis, *de novo* formed volatile emissions, and accumulation of terpenoid synthase and putative octadecanoid pathway transcripts in *Sitka* spruce. Plant Physiol. 2005;137:369–382.

Miller DJ, Allemann RK. Sesquiterpene synthases: passive catalysts or active players? Nat Prod Rep. 2012;29(1):60–71.

Moore BD, Andrew RL, Kulheim C, Foley WJ. Explaining intraspecific diversity in plant secondary metabolites in an ecological context. New Phytol. 2014;201:733–750.

Morant AV, Jørgensen K, Jørgensen C, Paquette SM, Sánchez-Pérez R, Møller BL, et al. Beta-glucosidases as detonators of plant chemical defense. Phytochemistry. 2008;69(9):1795–1813.

Mullins MG, Bouquet A, Williams LE. Biology of the Grapevine. Cambridge, UK: Cambridge University Press; 1992.

Mumm R, Dicke M. Variation in natural plant products and the attraction of bodyguards involved in indirect plant defense. Can J Zool. 2010;88(7):628–67.

Nelson D, Werck-Reichhart D. A P450-centric view of plant evolution. Plant J. 2011;66(1):194–211.

Pichersky E, Gershenzon J. The formation and function of plant volatiles: perfumes for pollinator attraction and defense. Curr Opin Plant Biol. 2002;5(3):237–43.

Pollard CJ, Bonner J, Haagen-Smit AJ, Nimmo CC. Metabolic transformation of mevalonic acid by an enzyme system from peas. Plant Physiol. 1966;41:66–70.

OIV, The OIV's assessment of the world vitiviniculture situations [Online]. 2015; http://www.oiv.int/oiv/info/enpublicationsstatistiques [Accessed Decmeber 2015].

Rivas F, Parra A, Martinez A, Garcia-Granados A. Enzymatic glycosylation of terpenoids. Phytochem Rev. 2013;12:327–39.

Robinson AL, Boss PK, Solomon PS, Trengove RD, Heymann H, Ebeler SE. Origins of grape and wine aroma. Part 2. Chemical and sensory analysis. Am J Enol Vitic. 2014;65(1):25–42.

Rohmer M, Knani TM, Simonin P, Sutter B, Sahmt H, Nationale E, et al. Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. 1993;524:517–24.

Siebert TE, Wood C, Elsey GM, Pollnitz AP. Determination of rotundone, the pepper aroma impact compound, in grapes and wine. J Agric Food Chem. 2008;56(10):3745–8.

Skinkis PA, Bordelon BP, Wood K V. Comparison of monoterpene constituents in Traminette, Gewürztraminer, and Riesling winegrapes. Am J Enol Vitic. 2008;59(4):440–5.

Sweetman C, Wong DC, Ford CM, Drew DP. Transcriptome analysis at four developmental stages of grape berry (*Vitis vinifera* cv. Shiraz) provides insights into regulated and coordinated gene expression. BMC Genomics. 2012;13(1):691.

Taiz L, Zeiger E. Plant Physiology. Sinauer Associates, Incorporated; 2006.

Tantillo DJ. Biosynthesis via carbocations: theoretical studies on terpene formation. Nat Prod Rep. 2011;28(6):1035–53.

Tasin M, Anfora G, Ioriatti C, Carlin S, De Cristofaro A, Schmidt S, et al. Antennal and behavioral responses of grapevine moth *Lobesia botrana* females to volatiles from grapevine. J Chem Ecol. 2005;31(1):77–87.

Tasin M, Bäckman AC, Bengtsson M, Ioriatti C, Witzgall P. Essential host plant cues in the grapevine moth. Naturwissenschaften. 2006;93:141–4.

This P, Lacombe T, Thomas MR. Historical origins and genetic diversity of wine grapes. Trends Genet. 2006;22(9):511–9.

Tholl D. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Curr Opin Plant Biol. 2006;9:297–304.

Vannozzi A, Dry IB, Fasoli M, Zenoni S, Lucchin M. Genome-wide analysis of the grapevine stilbene synthase multigenic family: genomic organization and expression profiles upon biotic and abiotic stresses. BMC Plant Biol. 2012;12(1):130.

Venturini L, Ferrarini A, Zenoni S, Tornielli GB, Fasoli M, Dal Santo S, et al. *De novo* transcriptome characterization of *Vitis vinifera* cv. Corvina unveils varietal diversity. BMC Genomics. BMC Genomics; 2013;14(1):41.

Vilanova M, Genisheva Z, Graña M, Oliveira JM. Determination of odorants in varietal wines from international grape cultivars (*Vitis vinifera*) grown in NW Spain. South African J Enol Vitic. 2013;34(2):212–222.

Wedler H, Pemberton R, Tantillo D. Carbocations and the complex flavor and bouquet of wine: mechanistic aspects of terpene biosynthesis in wine grapes. Molecules. 2015;20:10781–92.

Wood C, Siebert TE, Parker M, Capone DL, Elsey GM, Pollnitz AP, et al. From wine to pepper: Rotundone, an obscure sesquiterpene, is a potent spicy aroma compound. J Agric Food Chem. 2008;56(10):3738–44.

Yoshikuni Y, Ferrin TE, Keasling JD. Designed divergent evolution of enzyme function. Nature. 2006;440(7087):1078–1082.

Young PR, Lashbrooke JG, Alexandersson E, Jacobson D, Moser C, Velasco R, et al. The genes and enzymes of the carotenoid metabolic pathway in *Vitis vinifera* L. BMC Genomics. 2012;13(1):243.

Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M, Malerba G, et al. Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-seq. Plant Physiol. 2010;152:1787–95.

Chapter 3

Research results

In silico characterisation of the grapevine *TPS* gene family in inflorescence and during flowering

Chapter 3 – *In silico* characterisation of the grapevine *TPS* gene family in inflorescence and during flowering

3.1 Introduction

The 'omics' revolution of the last few decades has resulted in a plethora of large datasets that are publicly available. Computational or *in silico* techniques are commonly applied for the prediction and discovery of new information through the mining of these datasets. The suffix 'omics' is used to described large scale analysis on a protein, transcription or genome level of an organism. Various specialised fields have been created under the 'omics' blanket which facilitates the use of available resources to address new questions. For grapevine, as in most organisms, the first such resource is the genome sequence. In 2007, two groups independently published draft grapevine genome sequences, but with distinct differences: Jaillon *et al.* (2007) sequenced a near homozygous (~90%) inbred *V. vinifera* L. cv. Pinot noir clone (PN40024), whereas Velasco *et al.* (2007) sequenced a heterozygous commercial Pinot noir clone (ENTAV 115). Both groups predicted roughly 30 000 protein-encoding genes, but the complexities involved in assembling the heterozygous genome resulted in PN40024 being selected as the reference genome.

Diverse cultivars and research focuses have resulted in a somewhat fragmented picture in terms of grapevine omics. The reference genome forms the base for most next-generation experiments but herein lies a great limitation in teasing out cultivar specific differences. A grapevine gene expression atlas by Fasoli *et al.* (2012) gives a comprehensive overview of transcriptional events of grapevine. This microarray-based analysis shows that the plant has complex transcription networks that ultimately direct the plant towards maturation. A wealth of information was generated for organ and developmental specific tissues covering roughly 91 percent of the then predicted genes.

The diverse cultivars and research focusses resulted in a somewhat fragmented picture in terms of grapevine omics. The reference genome forms the basis for most next-generation experiments but herein lies a great limitation in teasing out cultivar specific differences. The grapevine gene atlas was done using the Corvina cultivar, but with expression analysis on a microarray designed from Pinot noir reference genome. Later studies by the same group using RNA sequencing and *de novo* assembly of the transcriptome show that in Corvina alone there are unique splice isoforms, novel protein encoding genes and 180 private genes (unique to Corvina) that were not previously found in the grapevine genome (Venturini *et al.*, 2013). This is to be expected seeing that considering that selection for genetic differences have been made for 8000 years which have resulted in an estimated 10 000 different cultivars worldwide (This *et al.*, 2006; Keller, 2010; Myles *et al.*, 2011).The compounding factors (genomic, somatic and phenotyptic plasticity) that influence cultivar diversity and limited molecular resources/tools necessitates a creative approach when

studying gene families across cultivars. In this chapter we aimed to 'mine' the grapevine gene atlas through *in silico* tools to study the grapevine TPS gene family. Although the cultivar differences are immense, the assumption was made that there would be enough homology to extrapolate the findings to the cultivars studied, as discussed in later chapters. The expression differences for TPS genes, their expression pattern similarities (organ and development specific) as well as co-expressing genes that could influence terpenoid volatiles were identified. The information gained in this chapter forms the basis for all following chapters with the workflows used forming a blueprint for future studies.

3.2 Materials and methods

3.2.1 Identification of TPS genes for *in silico* analysis

Seventy *V. vinifera* predicted TPS genes for *in silico* analysis were selected (Figure 3.1, green section). This list consisted of the functional/putatively functional genes characterised or identified by (Martin *et al.*, 2010), and corresponding gene accessions were retrieved from the V1 genome annotation repository available through CRIBI (http://genomes.cribi.unipd.it/grape/). Genes that had no V1 annotation were cross-referenced with CRIBI V2 resulting in a predicted/functional annotation for all 70 genes.

3.2.2 *In silico* expression and clustering analysis

RMA-normalised expression data for the identified putative TPS encoding genes were retrieved from the grapevine (*Vitis vinifera* L. cv Corvina) gene atlas (described in Fasoli *et al.* 2012) via GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/; accession GSE36128). The mean and standard deviation (of biological repeats, n=3) were calculated using Statistica (version 12) (Dell Inc., Tulsa, USA) and compared across flower developmental stages or flower organs (Figure 3.1, orange section). Expression pattern analysis and clustering were performed using the average for the three biological repeats.

Gene expression patterns were analysed based on Pearson-correlation coefficients (PCC) and clustered with Cluster 3.0 (de Hoon *et al.*, 2004) using the standard software parameters (Figure 3.1, blue section). Data was log2 scaled and mean-centered. No filtering was applied so that invariant genes were included in the analysis to allow for visualisation of the gene family as a whole. Hierarchical clustering was performed with average linkage and uncentred correlation as the similarity metric. Java Treeview (Saldanha, 2004) was used for visualisation of the clustering (tree) output.

3.2.3 Gene co-expression network construction

The grapevine gene expression atlas GSE36128 (Fasoli *et al.*, 2012) dataset served as the 'hook' to identify genes that co-express with TPS genes of interest ('baits'). By utilising an R script (Itkin

et al., 2013) we searched for genes in the hook dataset that co-express with our gene baits (Figure 3.1, red section). Pairwise similarity using the PCC between the bait and all genes in a hook dataset with an assigned *r*-value indicating the level of similarity was calculated. The *r*-value was used to filter the highest ranking co-expressing genes (namely: genes with the highest statistical significance/PCC). This cut-off was determined by the complexity of the hook dataset and the resulting number of co-expressing genes. The bait:hook interaction was visualised as a co-expression network (Gene:Gene Co-espression Network, GGCN) using the Cytoscape (version 3.2.1) software package (Shannon *et al.*, 2003). Gene baits served as source nodes connected by an edge with all their co-expressing genes (target nodes).

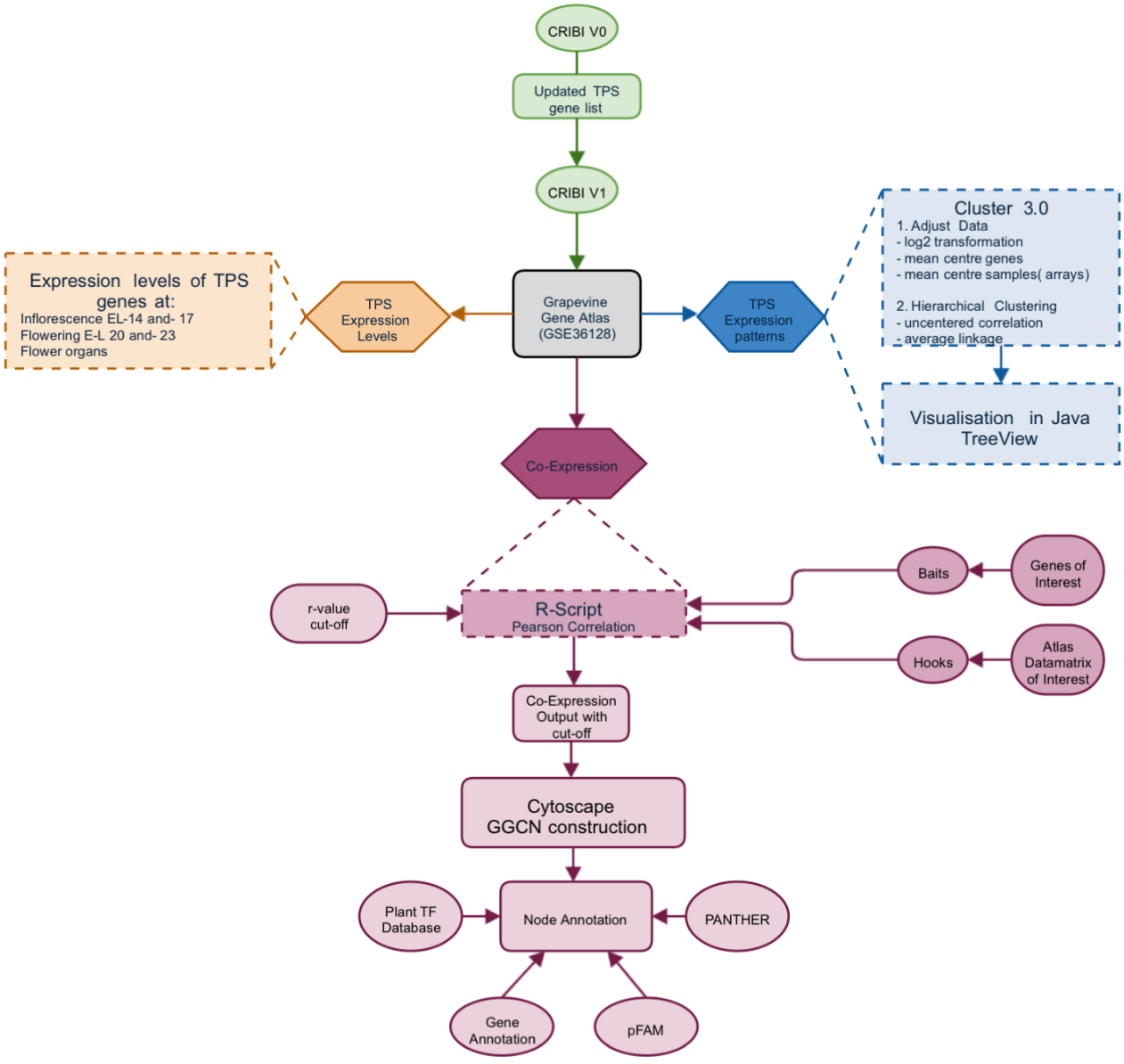


Figure 3.1 – Workflow summarizing the methods used for the *in silico* analysis of the grapevine TPS gene family. The different aspects of the analysis were grouped by colour. Green: migration of gene IDs to those used in CRIBI. Orange: Expression of TPS genes at specific stages. Blue: Expression pattern clustering. Red: Gene co-expression network analysis.

3.2.4 Annotation of nodes to identify putative gene functions

Gene classification of network nodes were performed using PANTHER (http://pantherdb.org/) (Thomas *et al.*, 2003; Mi *et al.*, 2005, 2013). The statistical overrepresentation test with default settings was used for *V. vinifera* with PANTHER version 10.0 (2015-05-15 release) and GO Ontology database (2015-08-06 release) annotations to determine protein class and gene ontology (GO) classifications, respectively, with Bonferroni correction. Statistically significant outputs for the different classifications were retrieved to determine network gene representations. Nodes were, furthermore, annotated using the available gene functions and protein family (pFam) descriptors (Finn *et al.*, 2014) from the CRIBI database (http://genomes.cribi.unipd.it/grape/). *V. vinifera* transcription factors were annotated using the Plant Transcription Factor database version 3.0 (http://plntfdb.bio.uni-potsdam.de/v3.0/).

3.3 Results

3.3.1 *In silico* expression pattern clusters of TPS genes

The atlas dataset used for gene expression clustering (GEC) excluded samples relating to post-harvest-withering (PHW) due to the unique transcriptome identified for this winemaking style (Fasoli *et al.,* 2012). Hierarchical clustering was performed in two dimensions which allows for clustering across samples and genes. Three gene expression clusters (GEC) were identified as illustrated in Figure 3.2. In GEC-1 a clear grouping of 16 genes (10 sesqui- and 6 mono-TPS) is seen that shows the highest correlation with the two flowering stages, E-L 20 and 23, and the four flower organs. GEC-2, consisting of 30 genes (thirteen mono-, eleven sesqui-, two di-TPS and four others genes that show homology to iridoid TPSs), was less defined with a general horizontal trend correlating with berry ripening and berry-related tissues/organs. Subtle differences across genes result in multiple vertical groupings for this cluster. The main samples that resulted in GEC-3 are correlated to inflorescence and tendrils that share a vertical pattern across 24 genes (18 mono-, five sesqui- and one di-TPS).

The sixteen genes of GEC-1 all share high expression levels for flowering stages and organs but can be divided into three sub-clusters that show the nuances in expression patterns (Figure 3.3). GEC-1A has two additional hotspots that correlate with rachis, at various stages of berry ripening and root tissue. GEC-1B consists of genes that correlate mainly with flower organs and their physiological stages. The unique expression pattern and difference in expression intensity for flower stages and organs compared to GEC-1A and -1B resulted in the third sub-cluster that consists of only two genes. These two genes are very similar in sequence and are probable duplications based on sequence homology.

GEC-2, consisting of 30 genes, was less defined with a general horizontal trend correlating with berry ripening and berry-related tissues/organs. Subtle differences across genes result in multiple

vertical groupings for cluster two. Most genes of GEC-2 show no significant expression correlation to the flower related samples that are the focus of this study. We however see a high expression trend for pericarp, skin and flesh tissue at véraison, mid-ripe and ripe stages of berry development.

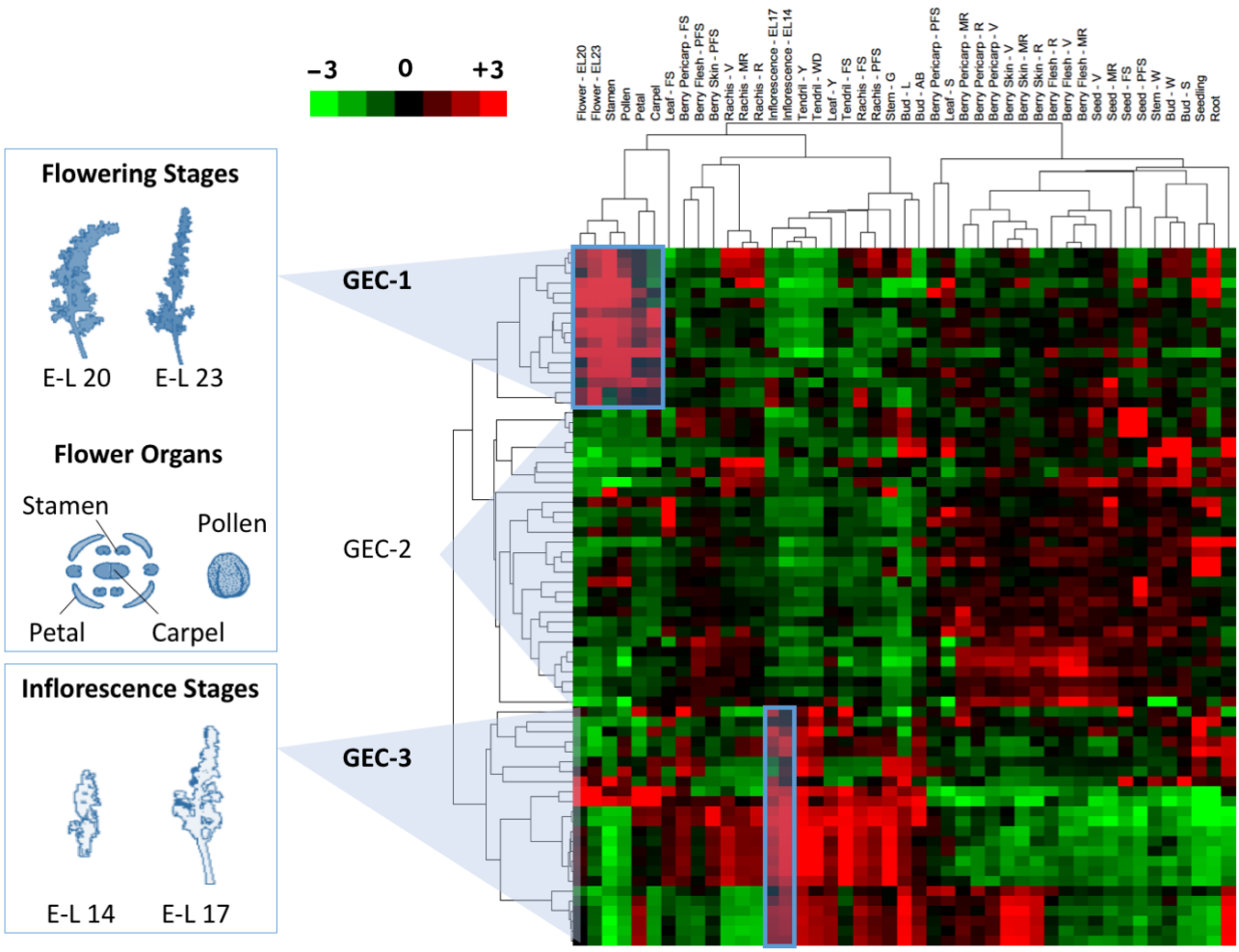


Figure 3.2 – Gene expression pattern clustering through pairwise PCC similarities. Positive correlations are shown in red, negative correlations in green with the intensity of the colour indicating the magnitude of expression. The dendograms show pairwise similarity of the expression pattern to gene/sample neighbours with similarity being directly proportional to the length of the branch. Horizontal similarity indicates genes with a similar expression pattern across stages/samples while vertical similarity shows samples/stages that express similar genes. Genes the show similar patterns across stages/samples were clustered with the main samples/stages resulting in these clustered indicated as GEC-1 to -3. Developmental stages and organs associated with GEC-1 and -3 is shown graphically with their associated E-L stages.

The main samples that resulted in GEC-3 are correlated to inflorescence and tendrils that share a vertical pattern across 24 genes. Multiple hot spots for genes with similar horizontal expression patterns allow for sub-clustering of the genes into three groups, Figure 3.4. GEC-3A separates from the other sub-clusters due to an organ-specific transcriptional upregulation (hotspot) correlating with inflorescence and tendril samples. Two genes positioned between GEC-3A and -3B could not be grouped into either sub-cluster because it showed high expression correlation across a broad range of samples that include flower stages, organs, inflorescence stages and

tendrils. GEC-3B shows a hotspot for samples that relate to young developing tissues that range from buds to leaves, tendrils, rachis and berries at or just after fruit-set. GEC-3C consists of a hotspot spanning inflorescence and tendrils to include samples at different plant development stages and berry pericarp from véraison onwards.

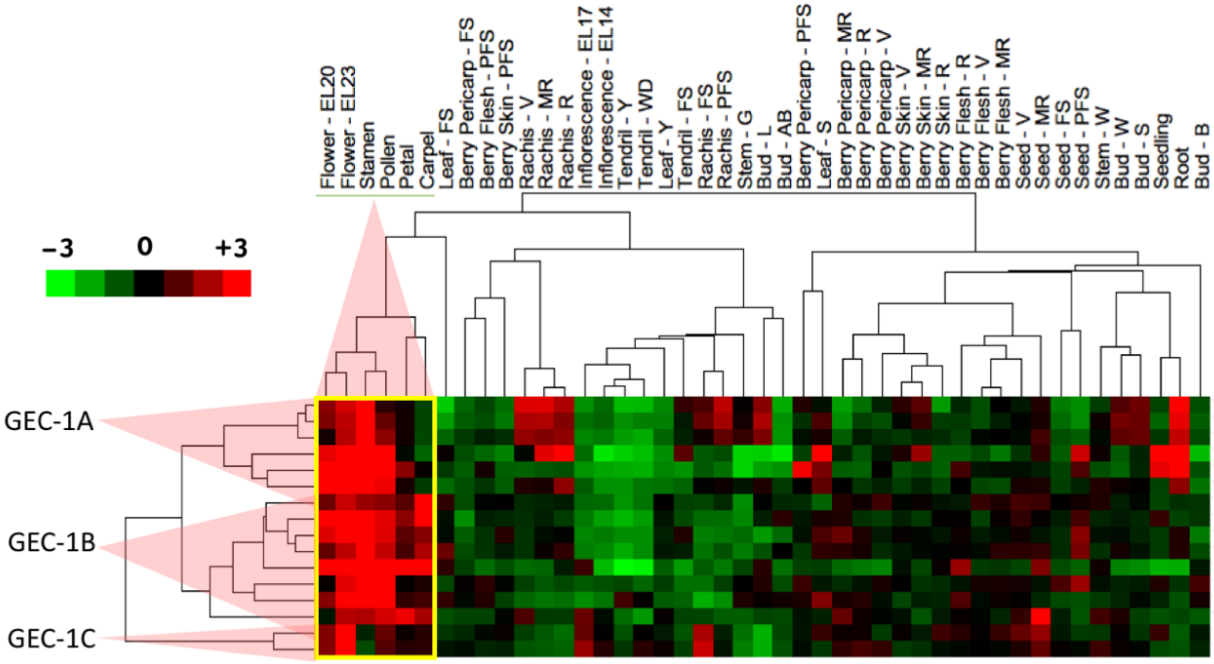


Figure 3.3 – An expanded view of GEC-1 with sub-clusters of genes indicated as GEC-1A to –C.

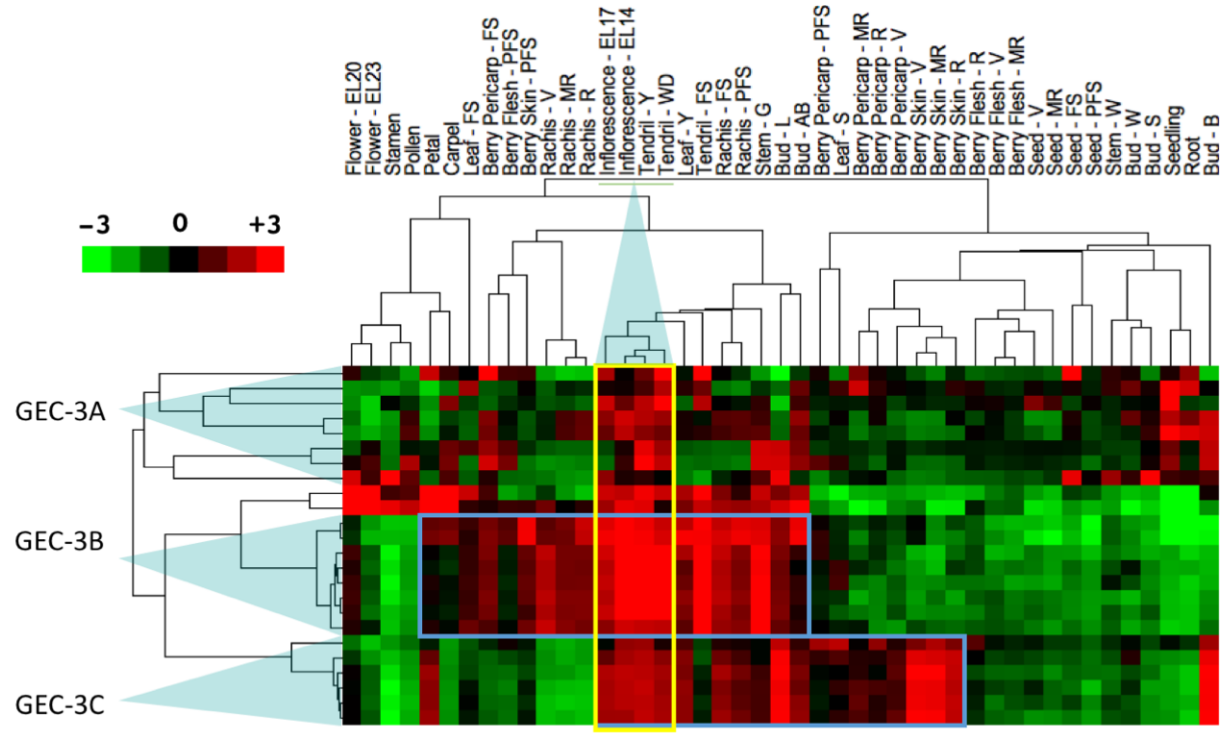


Figure 3.4 – An expanded view of GEC-3 with sub-clusters of genes indicated as GEC-3A to –C. Expression hotspots correlating to inflorescences/tendrils are shown in the yellow square. Samples/stages that contribute to the hotspots that seperate GEC-3A and –C are shown in blue squares.

3.3.2 Inflorescence and flowering specific expression patterns

We analysed the *in silico* expression patterns for GEC-1 and -3 genes, 40 in total, in a flowering and inflorescence-specific subset of the gene atlas. This subset consisted of the inflorescence (E-L14 and 17) and flowering (E-L20 and 23) stages as well as the four flower organs samples. Pairwise gene clustering resulted in two broad clusters GEC-4 and GEC-5, as before but greater resolution in terms of expression differences between stages and organs is seen as shown in Figure 3.5. This pairwise comparison also resulted in a rearrangement of nearest neighbours with GEC-4 consisting of 21 genes and GEC-5 of 19. GEC-6 was seen as a sub-branch from GEC-4 and consisted of genes that could not fit into GEC-4 or -5 exclusively due to expression similarities with both. In GEC-4 it is noted that the male flower organs separate from the female organs with greater expression levels seen in male organs.

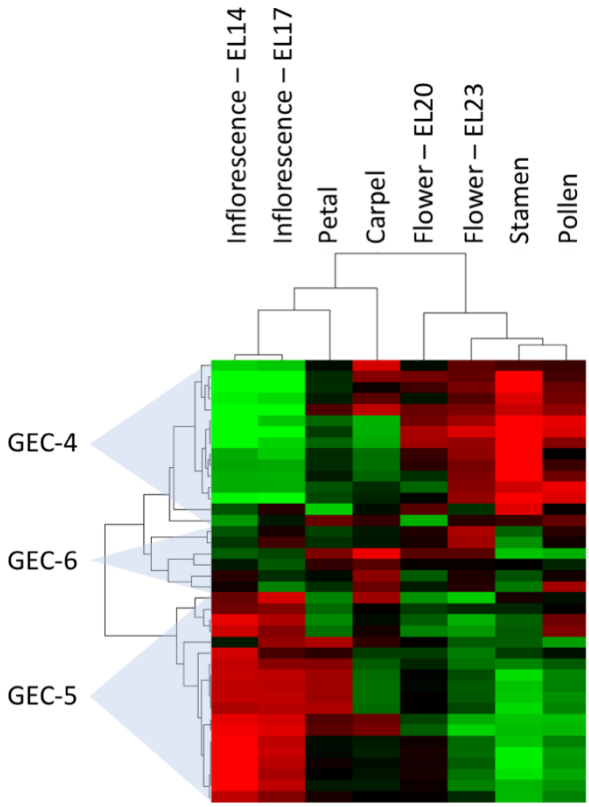


Figure 3.5 – Genes that showed high expression correlation to inflorescences/flowers were reanalysed resulting in GEC-4 to -6. GEC-4 genes correlate with flowering and flower organs with GEC-5 showing correlation with inflorescence stages. GEC-6 showed no clear pattern.

3.3.3 Gene atlas expression levels for TPS genes of interest at inflorescence and flowering stages

Gene atlas expression data for TPS genes that group to GEC-4, -5 and -6 were retrieved for the inflorescence and flowering stages as biological repeats (n=3). Calculated means and standard deviations (SD) were visualised as line graphs (Additional File, Figure 1 and 2)

3.3.4 Co-expression network analysis and enrichment for stage specific co-expressors

3.3.4.1 Global gene atlas GGCNs

GECs from the expression analysis were evaluated for co-expression that resulted in a GGCN for the respective clusters. GEC-1 and -3 were shown to consist of genes that are specifically related to inflorescence and flowering and were therefore analysed in depth. A network for GEC-2 was constructed (GGCN-2) as part of the workflow used in this study but will not be discussed as it was shown to be berry-ripening specific. The hook dataset used for PCC excluded samples related to PHW. Networks were constructed for co-expressing genes above the *r*-value cut-off at 0.8. GGCN-1 (Figure 3.6A) resulted in 450 nodes connected by 1537 edges for the 16 baits queried. GGCN-3 (Figure 3.6B) had 24 baits and comprised of 217 nodes connected by 398 edges. Even though GGCN-3 had more baits queried it showed fewer co-expressing genes that met the statistical cut-off. For each network the respective TPS baits were coloured as pink squares (Figure 3.6A and –B).

Fifteen of the sixteen baits for GGCN1 localised to a central node of co-expressing genes with the outer radius being genes that co-express with a specific bait. GGCN-3 consisted of six co-expression sub-networks (GGCN-3A to -F), with each sub-network showing a number of co-expressing genes unique to the baits of that sub-network. Sub-networks A and B contained ten and eight baits respectively. Sub-network C showed co-expression with a bait not part of GEC-3 (gold square). Sub-networks E and F were unique with single baits showing specific co-expressing genes.

3.3.4.2 Network enrichment for co-expressing genes specific to flowering or inflorescence

GEC-4, -5 and -6 genes were subsequently used as baits to identify co-expressing genes that are specific to either flowering or inflorescence. GEC-4 and -6 were grouped together as flowering-specific baits, while GEC-5 genes were deemed to be inflorescence-specific. PCC were determined for the inflorescence, flowering and flower organ subset and all co-expressing genes with an *r*-value cut-off above 0.95 were used to construct a flowering- (GGCN-4) and an inflorescence- (GGCN-5) specific network. GGCN-4 had 411 nodes with 803 edges and GGCN-5 had 913 nodes with 2841 edges (Additional File, Figure 3). Cross-comparison between the global atlas co-expression networks (GGCN-1 and -3) and the enriched networks (GGCN-4 and -5) resulted in a highly flowering/inflorescence specific list of co-expressing genes. Cross-comparison was achieved by merging these four networks to construct GGCN-6 as illustrated in Figure 3.7. This network consisted of two sub-networks (GGCN-6A and -6B) with edge colouring showing which parent networks contribute to these sub-networks. GGCN-1 and -4 formed sub-network A while GGCN-5 dominated sub-network B with GGCN-3 being distributed between both sub-networks.

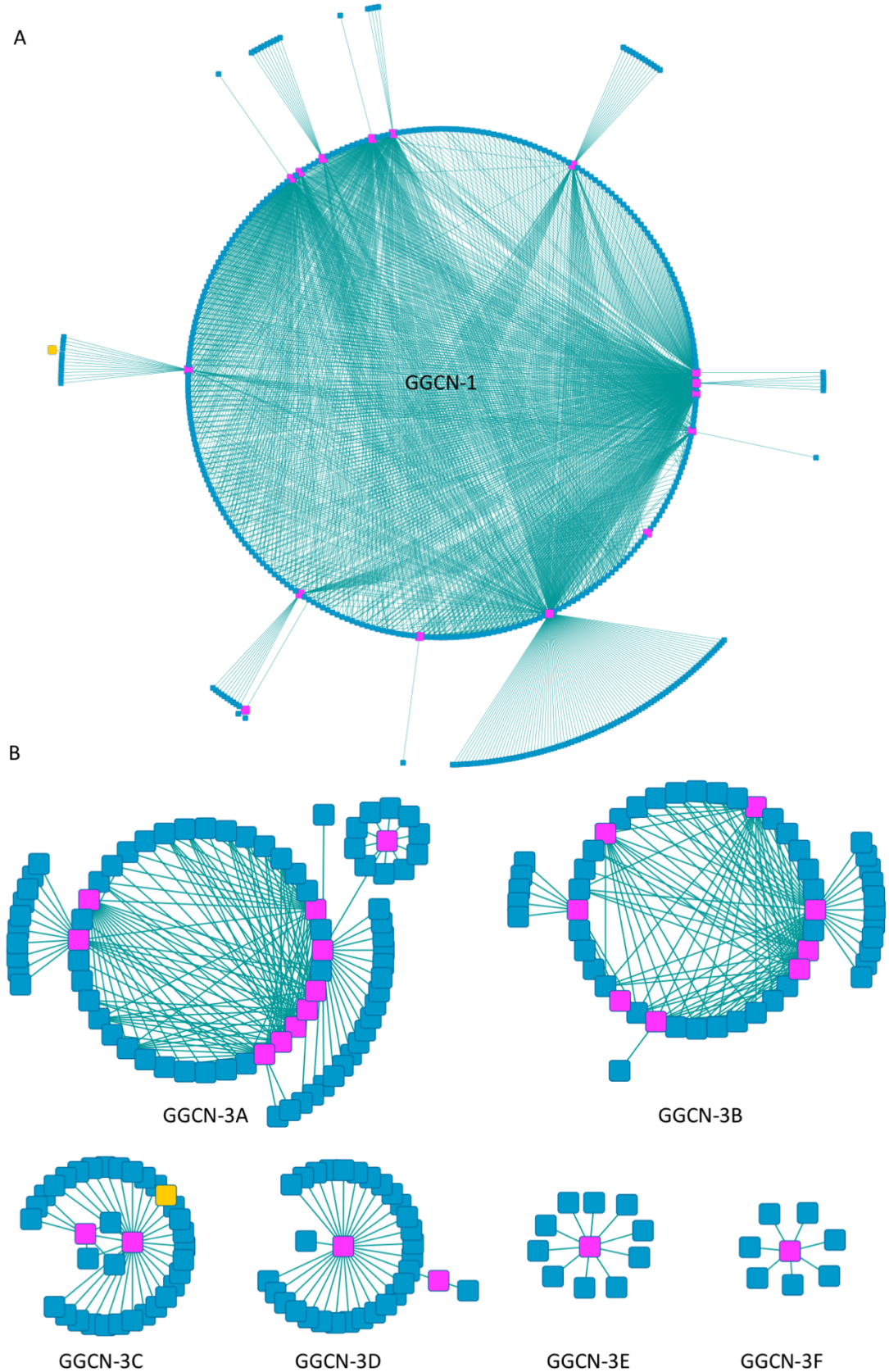


Figure 3.6 – Gene co-expression networks (GGCN) for TPS gene baits. (**A**) GGCN-1 was constructed using genes from GEC-1 as baits with (**B**) showing GEC-3 gene baits for the construction of GGCN-3. GGCN-3 shows six sub-networks (3A to -F) of co-expressing genes associated with specific baits. Gene baits from the respective GECs are shown as pink squares. The yellow square of GGCN-3C represents a TPS gene that was part of the 70 target genes but not a gene bait for the network (i.e. a candidate TPS gene co-expressing with the baits of GEC-3).

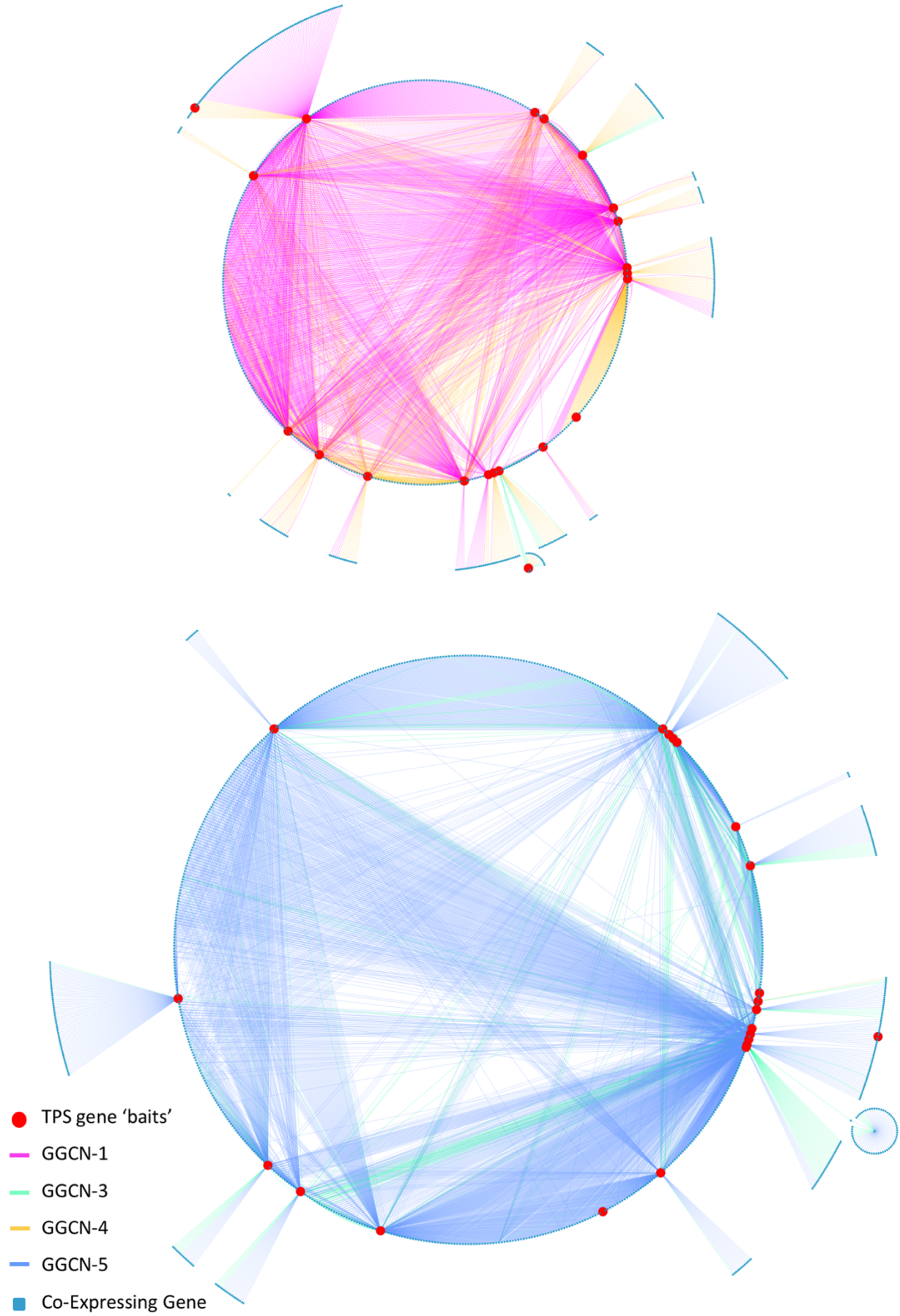


Figure 3.7 – Merging of GGCN-1, -3, -4 and -5 resulted in two sub networks, A and B, for GGCN-6. The legend reflects the edge colours for the parent networks used in the construction of GGCN-6 as well as the node colours for TPS gene baits and co-expressing genes.

3.3.4.3 Gene ontology (GO) and predicted protein classifications of co-expressed genes

GGCN-6A consisted of 752 nodes with 2383 edges while -6B had 1007 nodes with 3196 edges. PANTHER classifications for protein classes and gene ontologies (GO) for molecular and biological processes, that showed enrichment with a 0.05 *p*-value cut-off, were retrieved. The percentage of genes relative to the number of nodes per network that showed no classification for the three different analyses is shown in (Additional File, Figure 4). These genes have unknown/unidentified functions and are likely novel genes. The respective network nodes that could be classified for the three analyses are shown in (Additional File, Figure 5 and 6).

Incorporating multiple databases for the custom node annotation file allowed for more specific identification of gene function that was not evident in the PANTHER classifications. From GGCN-6 103 genes were identified that were exclusively correlated with flowering and flower organs while 65 genes were only correlated to inflorescence stages. Of these genes only 39 were uncharacterised, unnamed or hypothetical proteins. We excluded our gene baits from this enrichment to reveal only co-expressing genes of interest. Ten TPS-like genes were unique to flowering while six were unique to inflorescence. These genes did not form part of the 70 baits queried in the networks as they were predicted to be tandem gene duplications or predicted as non-functional. Flowering contained five glycosyl hydrolase gene family members and inflorescence three. Cytochrome P450s and ATP-binding cassette (ABC) transporters where identified in both inflorescence and flowering. Eight shikimate *O*-hydroxycinnamoyltransferase and five multicopper oxidase genes were found to co-express in flowering stages.

Global gene atlas TFs that correlate with flowering/flower organs or inflorescence revealed a large number that co-expressed with multiple TPS-encoding genes. GGCN-5 showed 104 TFs that co-expressed with nineteen TPS genes while GGCN-4 had seventeen TFs associated with ten TPS genes (Figure 3.8A). The TFs could be grouped into 42 different families with their distribution between the networks shown in Figure 3.8B. Subsequent enrichment in GGCN-6 revealed four TFs belonging to the C2C2-CO-like, MYB, NAC and orphans (TFs that cannot be grouped) families, that were highly specific to flowering while four bHLH, two TCP and one AUX/IAA TF were uniquely associated with inflorescence.

The consideration of PCC is that it identifies patterns and excludes absolute expression levels. To negate this limitation, we retrieved the absolute expression levels from the gene atlas for co-expressing genes of interest. Differential expression between inflorescence (E-L 14 and 17) and flowering (E-L 20 and 23) stages for relevant genes from GGCN-6A and -6B are shown in Figure 3.9.

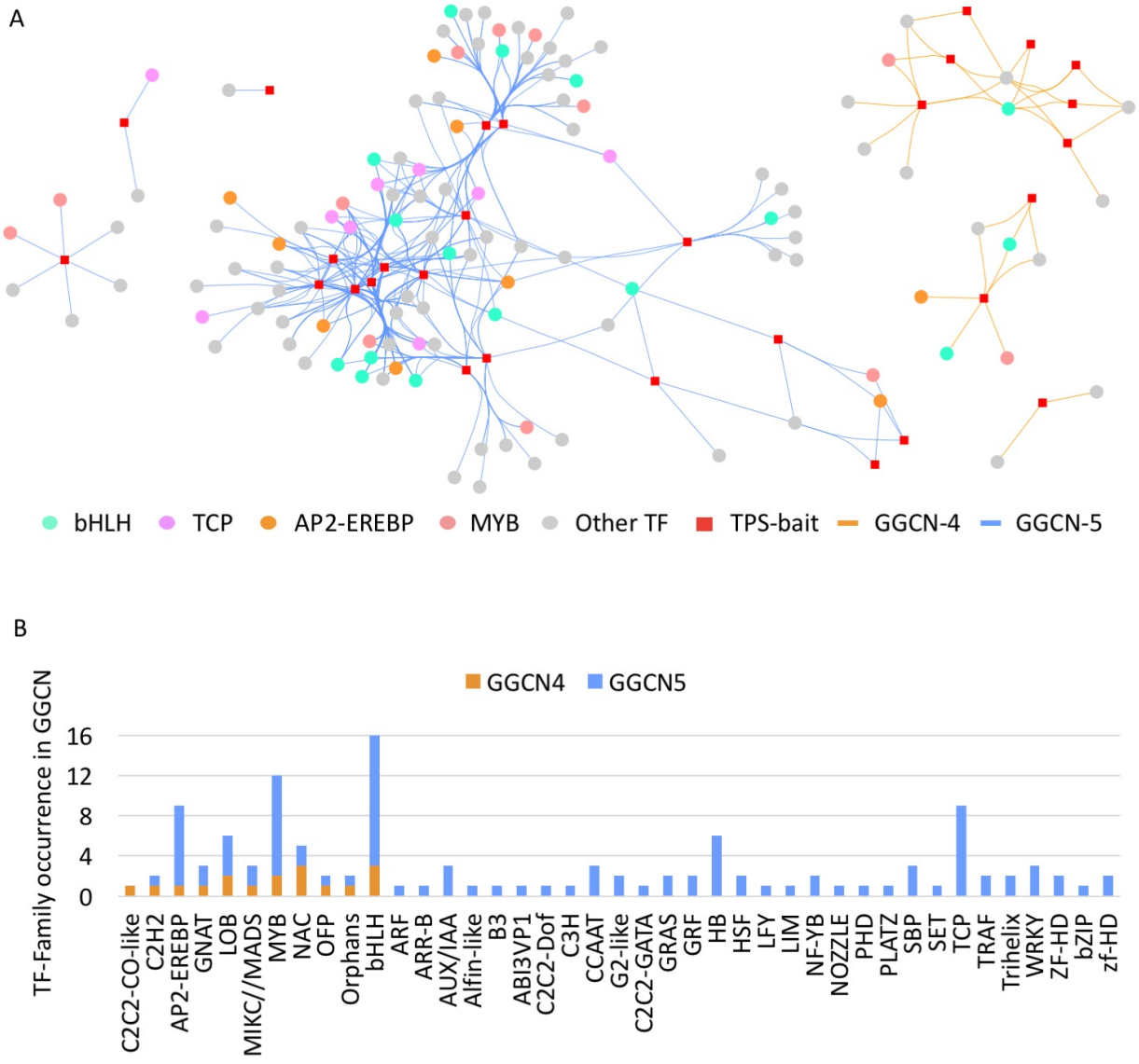


Figure 3.8 – TFs that co-express with gene baits of GGCN-4 and -5 were extracted and visualised in (**A**). Edge colouring was used to separate the two networks, with gene baits as red squares. Over-represented TF families were coloured with the number of TFs of that family shown in (**B**).

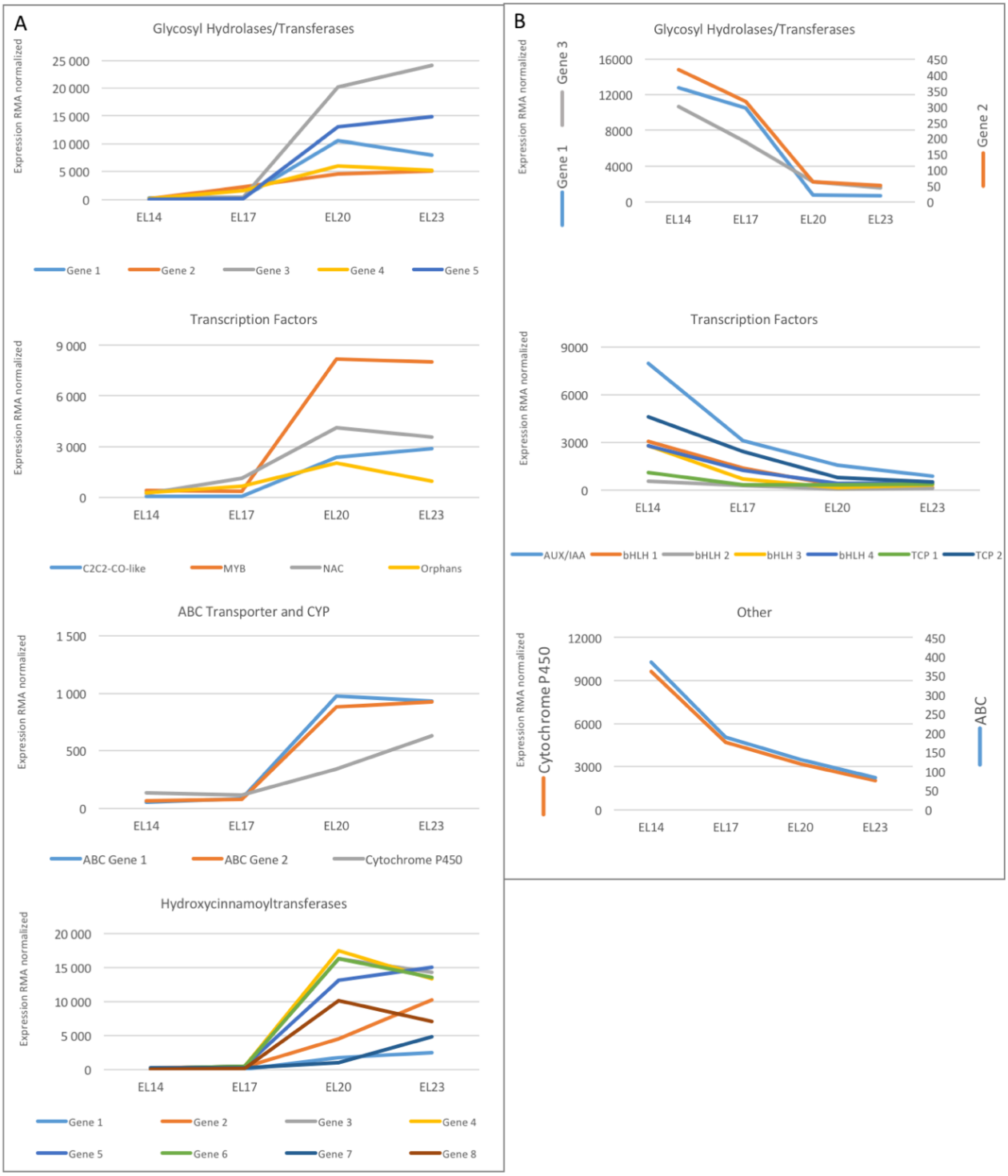


Figure 3.9 – Co-expressing genes of interest that are differentially expressed between the inflorescence and flowering stages. Genes that show increased expression in (A) flowering (E-L 20 and -23) and (B) inflorescence stages (E-L 14 and -17) are shown. Expression levels are the RMA normalised values reported in the grapevine gene atlas GSE36128. A gene number was assigned when multiple genes associated with a specific function was found.

3.4 Discussion

3.4.1 TPS expression is upregulated in flowers and inflorescences and specific TPSs are differentially expressed during these developmental stages

Plant volatile terpene biosynthesis is complex and highly species-specific with a spatiotemporal regulation of transcripts affecting the volatile emissions. Large scale cross-genome comparisons have shown that terpenoid metabolism evolved and diversified in flowering plants with sub-functionalisation events giving rise to species-specific volatiles (Hofberger *et al.*, 2015). By nature of the two pathways involved in mono- and sesquiterpene biosynthesis, compartmentalisation to the cytosol or plastids, respectively, allows for transcriptional regulation and control over substrate specificity. The deployment of terpenoid volatiles in flowers suggests that they act as attractants, and in most plants, this would be true. In the case of grapevine, however, where the flower is hermaphroditic, the need to attract a pollinator has become redundant. Grapevine breeding/selection over the centuries has selected for aroma and taste qualities and therefore it is no surprise that many research groups focus on the impact of terpenes on wine aroma and other sensorial properties.

We identified 70 TPS genes that were either already characterised by Martin *et al.* (2010) or predicted to be functional based on sequence homology with isolated genes. The high sequence homology between genes, with relatively few nucleotide changes resulting in a gene with altered functionality, gave rise to a diverse gene family with numerous volatile possibilities. By nature of *in silico* data mining we used available data to identify patterns and trends to generate hypotheses on gene functions. The *in silico* gene expression clusters were based on pairwise gene expression pattern similarities and resulted in a clear separation between the flowering (GEC-4 and -6) and inflorescence (GEC-5) physiological stages (Figure 3.5). This suggests that the members of this over-represented gene family have separate functions that can be linked to the physiological development of flowers. Gene descriptors for the grapevine genome as well as other information from various databases showed that our list of TPS genes are dominated by mono- and sesquiterpene-like genes, with the former being prominent at inflorescence development while the latter were associated more with flowers at anthesis.

Absolute expression levels showed considerable biological variation for some of the genes but as a general trend the sesquiterpenes that cluster with the flowering stages, and related organs, had increased expression levels compared to the monoterpenes (Additional File, Figure 2). One gene, mono-TPS08, was the exception to this trend with a consistently high expression level in the flowering stages. This gene showed high sequence similarity with known isoprene synthases. Isoprene is the building block for all other terpenes but as a volatile itself it has important biological functions. It has been shown that abiotic factors such as heat and sunlight irradiation has an effect on isoprene emissions in *Populus alba* while other poplar species showed that isoprene synthase

is transcriptionally regulated and volatilised during young leaf development (Mayrhofer *et al.*, 2005; Sasaki *et al.*, 2005). It was also suggested by Mayrhofer *et al.* (2005) that isoprene synthase is under diurnal and seasonal regulation in response to changes in light and temperature. Isoprene has furthermore been implicated in various protective functions ranging from protection against membrane damage to plant-insect interactions (Lothawornkitkul *et al.*, 2008; Unsicker *et al.*, 2009). The inflorescence specific gene cluster (GEC-5) consisted mainly of monoterpene genes (Figure 3.5). Absolute expression levels of mono-TPS genes were increased at inflorescence stages compared to the sesqui-TPS genes but generally at lower levels than the flowering genes of GEC-4 and -6 (Additional File, Figure 1).

These patterns are in agreement with other plant species where they have found specific genes involved in compartmentalised regulation of TPS transcription and volatile emission. Mono- and sesquiterpenes are known to be important plant volatiles and numerous genes have been isolated and characterised from various plants. A review by Unsicker *et al.* (2009) discussed the importance of these volatiles in vegetative plant tissues. Reproductive organs, flowers and fruits, have seemingly specialised genes involved in terpene biosynthesis that fulfil a wide range of functions. The monoterpene *S*-linalool was one of the first acyclic monoterpenes characterised and is shown to be emitted by female organs of *Clarkia breweri* flowers (Dudareva *et al.*, 1996). It was found that the sesquiterpene nerolidol and monoterpene linalool are produced by nearly identical genes with substrate availability due to compartmentalisation resulting in these distinct volatiles in snapdragon flowers (Nagegowda *et al.*, 2008). These volatiles in conjunction with the monoterpenes myrcene and (*E*)-β-ocimene are the main volatiles responsible for snapdragon flower volatiles (Dudareva *et al.*, 2003).

A recent study in the Muscato bianco cultivar shows that 23 of the *V. vinifera* TPS gene characterised by Martin *et al*. (2010) (and part of the study presented here), have specific expression patterns that are organ-specific, confirming that sub-functionalisation results in temporal gene expression (Matarese *et al.*, 2014). The flower organs used by the aforementioned authors are not well described in terms of physiological development but it is clear that the samples were from inflorescence and flower stages. It was shown through RT-qPCR that certain TPS transcripts were higher in inflorescences than flowers but the link between these transcripts and the volatiles detected at these stages were problematic due to the complex nature of plant systems. In some cases, transcripts were present at high levels but no corresponding volatile could be quantified while for the caryophyllene synthases multiple gene transcripts were correlated with a single volatile.

The role of mono- and sesquiterpene volatiles of grapevine flowers is not well characterised, with fragmented information from certain cultivars alluding to an insect interaction role. Entomological assays showthat male European grapevine moths (*Lobesia botrana*) respond to a reconstituted mixture of volatiles resembling that of *V. vinifera* cv. Solaris flowers. (*E*)-β-caryophyllene was

shown to be one of the important attractants and could attract the moths when used on its own (von Arx *et al.*, 2011). A similar study using cv. Casana green berries showed that the two sesquiterpene (*E*)-β-caryophyllene and (*E*)-β-farnesene and the hometerpene (*3E*)-4,8-dimethyl-1,3,7-nonatriene in highly specific ratios imitate the pheromones of the female berry moth and therefore attract the male moth (Tasin *et al.*, 2007). Oviposition by these moths result in damage to plant organs and subsequently create an infection point for the fungal necrotroph *Botrytis cinerea* and these volatiles are, therefore, seen as unwanted attractors (Tasin *et al.*, 2005).

3.4.2 Genes co-expressing with TPSs and the potential metabolic implications

Through gene co-expression analysis and gene classification enrichment for nodes that correlate with flowering (GGCN-6A, Figure 3.7) and inflorescence (GGCN-6B, Figure 3.7), a large proportion of co-expressors had no protein or GO classification.

We were able to show a strong TPS transcriptional response for most of the TPS-encoding genes targeted (gene baits) with most of these targets showing co-expression with other TPS-encoding genes that were either part of our target list or TPS-like genes that were predicted to be non-functional. The putatively non-functional genes were identified from the genome but cultivar-specific evolution could have retained some of these genes as functional copies. Herein lies the limitation of using a homozygous genome to make species-wide conclusions. Furthermore, the gene atlas used in this study was for the cultivar Corvina with the microarray used for the atlas being designed on the Pinot noir reference genome. *De novo* transcriptome characterisation of Corvina later showed significant varietal diversity with so-called private genes thought to contribute to cultivars specific characteristics (Venturini *et al.*, 2013). This suggests that co-expressing genes identified in the Corvina gene atlas are likely conserved between cultivars and that cross-cultivar gene homology (such as, Pinot noir versus Corvina), could identify genes of interest for future studies.

3.4.2.1 'Guilt-by-association': Transcription factors implicated in the regulation of terpene metabolism

Numerous TFs were identified from various TF families that correlate to specific flower physiological stages (Figure 3.8). The extensive cell division and organ differentiation processes occurring during inflorescence necessitates a complex metabolic network to control the cell proliferation burst experienced at this stage (Díaz-Riquelme *et al.*, 2014). The bHLH family of TFs was over represented and their role in controlling the cell cycle, chromatin assembly, cell proliferation and various other processes has been reported (Heim *et al.*, 2003). Interestingly, an AUX/IAA transcription factor that co-expressed at the inflorescence stages was identified. This TF family is involved in the repression of genes regulating primary/early auxin responses. The repression of these responses is likely due to the change in photoperiod that is beneficial for flowering where a down-regulation of auxin promotes the differentiation of inflorescence primordia

(Salisbury, 1955). Two members if the TCP TF-family also showed co-expression at inflorescence. Members of this family were shown in *Arabidopsis* to be expressed in rapidly growing flower primordia and are thought to be involved in promoting cell division (Cubas *et al.*, 1999). The GO biological process classifications for inflorescence (Additional File, Figure 6) supported the identification of these TFs due an over-representation of genes involved in functions like post-embryonic, tissue, shoot and reproductive organ developmental processes.

Four TFs of different families co-expressed during flowering: (1) C2C2-CO-like is part the zinc-finger domain containing TFs that act on CONSTANS (CO) like genes that are thought to control the induction of flowering in *Arabidopsis* (Lagercrantz & Axelsson, 2000); (2) the NAC TF-family with diverse functions in defence/stress responses and developmental programming (comprehensive reviewed in Olsen *et al.* 2005); (3) the MYB TF-family have diverse functions including secondary metabolism; and (4) the R2R3, a sub-family of MYBs, have been associated with the regulation of phenylpropanoid biosynthesis that result in hydroxycinnamic acid metabolites (Jin & Martin, 1999). Furthermore, it is shown that bHLH TFs form ternary complexes with MYB and WD40 TFs to regulate flavanoid biosynthesis (Hichri *et al.*, 2011). In rose (*Rosa hybrida)* flowers, the over-expression of the *Arabidopsis PAP1*, a MYB TF encoding gene, resulted in increased volatile emissions that originate from both phenylpropanoid and terpenoid biosynthetic pathways (Zvi *et al.*, 2012). The GO molecular functions associated with flowering (Additional File Figure 5) show significant enrichment for processes linked to metabolite modification via hydrolase activities. Numerous genes are linked to transport functions with an apparent increase in redox reactions. It is thus possible that the TFs identified in this study act as global regulators of flower volatiles synthesised through various metabolic pathways controlled by a specialised network of genes to facilitate metabolite modification and transport.

3.4.2.2 Co-expressing genes that could affect flower volatile biosynthesis and emissions

We found eight shikimate hydroxycinnamoyltransferase-like genes that co-express in flower stages. These genes are likely involved in phenylpropanoid biosynthesis that can result in a diverse range of metabolites that include stilbenes, flavonoids and anthocyanins (Vogt, 2010). The occurrence of potential glycosylation enzymes is of particular interest due to their role in the translocation of sugar moieties that alter the chemical properties of plant metabolites. For terpenoids, this is a necessity to ensure that the plant is not adversely affected by toxic terpenes or to change terpene solubility which facilitates transport (Rivas *et al.*, 2013). GO enrichment showed 30 candidate genes with glycosyl hydrolase characteristics with subsequent enrichments of GGCNs showing putative genes that perform antagonistic functions through either the addition of a sugar (glycosyltransferases), or through the cleavage of the sugar moiety through hydrolysis (glycosidases).

A recent study in grapevine identified the first *Vitis* glycosyltransferase which shows an increased expression at inflorescences and flowering stages compared to the leaves and roots, with a substrate specificity that was higher for monoterpenes than sesquiterpenes. Furthermore, it was shown through transcript analysis that most of the gene targets also expressed in young berries with a decrease in expression as ripening progresses (Bönisch *et al.*, 2014). We found that more glycosyl transferase type genes co-express with monoterpenes at inflorescence while the hydrolases are associated more with sesquiterpenes at flowering. Monoterpenes are more readily detected in grape berries and it is thought that glycosidases from yeasts during vinification release volatiles that contribute to the floral aromas found in wines from aromatic cultivars (Loscos *et al.*, 2007). It is, thus, reasonable to postulate that monoterpene volatiles are not important in flowering and are, therefore kept in the conjugated form for a yet-to-be identified biological role.

The global gene atlas co-expression networks (GGCN-1 and -3) showed numerous cytochrome P450 (CYP) genes with subsequent enrichment of the networks resulting in two genes that correlate highly with inflorescence and flowering, respectively. CYPs form the largest enzyme family in plants with numerous functions and activities that result in a plethora of modified metabolites. In plants, it is thought that their main function is to contribute to chemical defence, through amongst other functions, the oxidation of metabolites. These functions, structural diversity, enzymatic mechanisms and genetic evolution are explained in a comprehensive review by Werck-Reichhart & Feyereisen, (2000). Numerous groups have focussed on the interaction between CYPs and terpenoids. The first such an interaction was shown for CYP82G1 where the geranyllinalool C_{20}-precursor involved in diterpene biosynthesis serve as substrate resulting in the synthesis homoterpenes (C_{11} and C_{16} volatiles) as a response to insect herbivory (Lee *et al.*, 2010). Ginglinger *et al.* (2013) showed that CYP71B31 and CYP76C3 co-expressed with two linalool synthase monoterpenes in *Arabidopsis*. These two synthases produced the (+)-R and (-)-L enantiomers of linalool respectively with *in vitro* assays showing that the aforementioned CYPs use the enantiomers differentially resulting in specific oxidised terpenoids. Furthermore, the study showed specific compartmentalisation of the terpene and CYP proteins in flower organs, with the CYPs being present in the endoplasmic reticulum (ER) that surrounds plastids, facilitating a probable interaction with the plastidial-derived terpene substrates. These CYPs show a preference for linalool, compared to other monoterpenes, and it was proposed that oxidised linalool conjugates perform a function not related to flower volatile emissions (Ginglinger *et al.*, 2013). The two CYPs identified in the co-expression analysis presented here, therefore, provide promising candidates for the study of the interaction with grapevine terpenoids.

3.5 Conclusion

The expression patterns and related expression levels for TPS genes allowed us to identify candidate genes that are more likely transcriptionally involved in specific flower developmental stages.

Mono-TPS are associated with inflorescence stages with the co-expressing genes of these stages suggesting that monoterpenes are conjugated by glycosyltransferases and/or oxidised by CYPs to perform probable protective function(s) against biotic and abiotic stresses rather than functions related to volatile emissions for attraction.

Sesqui-TPS encoding genes showed correlation with flowering stages, with specific localisation to male flower organs, and are thought to be involved in flower volatile biosynthesis. We identified candidate glycosidase genes that could be involved in the cleavage of sesquiterpene-glycosyls to ensure volatilisation. The presence of phenylpropanoid biosynthesis genes suggests that flower volatile metabolism is upregulated, with an identified MYB transcription factor showing promise as a global regulator for volatile biosynthesis.

3.6 References

von Arx M, Schmidt-Büsser D, Guerin PM. Host plant volatiles induce oriented flight behaviour in male European grapevine moths, *Lobesia botrana*. J Insect Physiol. 2011;57(10):1323–1331.

Bönisch F, Frotscher J, Stanitzek S, Rühl E, Wüst M, Bitz O, et al. A UDP-glucose:monoterpenol glucosyltransferase adds to the chemical diversity of the grapevine metabolome. Plant Physiol. 2014;165(2):561–581.

Cubas P, Lauter N, Doebley J, Coen E. The TCP domain: A motif found in proteins regulating plant growth and development. Plant J. 1999;18(2):215–22.

Díaz-Riquelme J, Martínez-Zapater JM, Carmona MJ. Transcriptional analysis of tendril and inflorescence development in grapevine (*Vitis vinifera* L.). PLoS One. 2014;9(3):e92339.

Dudareva N, Cseke L, Blanc VM, Pichersky E. Evolution of floral scent in *Clarkia*: novel patterns of *S*-linalool synthase gene expression in the *C. breweri* flower. Plant Cell. 1996;8(7):1137–1148.

Dudareva N, Martin D, Kish CM, Kolosova N, Gorenstein N, Fäldt J, et al. (*E*)-beta-ocimene and myrcene synthase genes of floral scent biosynthesis in snapdragon: function and expression of three terpene synthase genes of a new terpene synthase subfamily. Plant Cell. 2003;15(5):1227–1241.

Fasoli M, Dal Santo S, Zenoni S, Tornielli GB, Farina L, Zamboni A, et al. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell. 2012;24(9):3489–505.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2014;42(D1):D222–230.

Ginglinger J-F, Boachon B, Höfer R, Paetz C, Köllner TG, Miesch L, et al. Gene coexpression analysis reveals complex metabolism of the monoterpene alcohol linalool in *Arabidopsis* flowers. Plant Cell. 2013;25(11):4640–57.

Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC. The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. Mol Biol Evol. 2003;20(5):735–47.

Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. J Exp Bot. 2011;62(8):2465–83.

Hofberger J a., Ramirez AM, Bergh E Van Den, Zhu X, Bouwmeester HJ, Schuurink RC, et al. Large-scale evolutionary analysis of genes and supergene clusters from terpenoid modular pathways provides insights into metabolic diversification in flowering plants. PLoS One. 2015;10(6):e0128808.

de Hoon MJL, Imoto S, Nolan J, Miyano S. Open source clustering software. Bioinformatics. 2004;20(9):1453–1454.

Itkin M, Heinig U, Tzfadia O, Bhide AJ, Shinde B, Cardenas PD, et al. Biosynthesis of antinutritional alkaloids in Solanaceous crops is mediated by clustered genes. Science. 2013;341(6142):175–179.

Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449(7161):463–467.

Jin H, Martin C. Multifunctionality and diversity within the plant MYB -gene family. 1999;577–585.

Keller M. Botany and Anatomy. Sci Grapevines Anat Physiol. 2010;1–47.

Lagercrantz U, Axelsson T. Rapid evolution of the family of CONSTANS LIKE genes in plants. Mol Biol Evol. 2000;17(10):1499–507.

Lee S, Badieyan S, Bevan DR, Herde M, Gatz C, Tholl D. Herbivore-induced and floral homoterpene volatiles are biosynthesized by a single P450 enzyme (CYP82G1) in *Arabidopsis*. Proc Natl Acad Sci. 2010;107(49):21205–10.

Loscos N, Hernandez-Orte P, Cacho J, Ferreira V. Release and formation of varietal aroma compounds during alcoholic fermentation from nonfloral grape odorless flavor precursors fractions. J Agric Food Chem. 2007;55(16):6674–6684.

Lothawornkitkul J, Paul ND, Vickers CE, Possell M, Taylor JE, Mullineaux PM, et al. Isoprene emissions influence herbivore feeding decisions. Plant Cell Environ. 2008 Oct;31(10):1410–1415.

Martin DM, Aubourg S, Schouwey MB, Daviet L, Schalk M, Toub O, et al. Functional annotation, genome organization and phylogeny of the grapevine (Vitis vinifera) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. BMC Plant Biol. 2010;10:226.

Matarese F, Cuzzola A, Scalabrelli G, D'Onofrio C. Expression of terpene synthase genes associated with the formation of volatiles in different organs of *Vitis vinifera*. Phytochemistry. 2014

Mayrhofer S, Teuber M, Zimmer I, Louis S, Fischbach RJ, Schnitzler J-P. Diurnal and seasonal variation of isoprene biosynthesis-related genes in grey poplar leaves. Plant Physiol. 2005;139(1):474–484.

Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S, et al. The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res. 2005;33:D284–288.

Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. 2013;41(D1):D377–386.

Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK. Genetic structure and domestication history of the grape. 2011.

Nagegowda D a, Gutensohn M, Wilkerson CG, Dudareva N. Two nearly identical terpene synthases catalyze the formation of nerolidol and linalool in snapdragon flowers. Plant J. 2008;55(2):224–239.

Olsen AN, Ernst HA, Leggio L Lo, Skriver K. NAC transcription factors: structurally distinct, functionally diverse. Trends Plant Sci. 2005;10(2):79–87.

Rivas F, Parra A, Martinez A, Garcia-Granados A. Enzymatic glycosylation of terpenoids. Phytochem Rev. 2013;12:327–39.

Saldanha AJ. Java Treeview - Extensible visualization of microarray data. Bioinformatics. 2004;20(17):3246–3248.

Salisbury F. The dual role of auxin in flowering. Plant Physiol. 1955;327–334.

Sasaki K, Ohara K, Yazaki K. Gene expression and characterization of isoprene synthase from *Populus alba*. FEBS Lett. 2005;579(11):2514–2518.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape : a software environment for integrated models of biomolecular interaction networks. 2003;(Karp 2001):2498–2504.

Tasin M, Anfora G, Ioriatti C, Carlin S, De Cristofaro A, Schmidt S, et al. Antennal and behavioral responses of grapevine moth *Lobesia botrana* females to volatiles from grapevine. J Chem Ecol. 2005;31(1):77–87.

Tasin M, Bäckman A-C, Coracini M, Casado D, Ioriatti C, Witzgall P. Synergism and redundancy in a plant volatile blend attracting grapevine moth females. Phytochemistry. 2007;68(2):203–9.

This P, Lacombe T, Thomas MR. Historical origins and genetic diversity of wine grapes. Trends Genet. 2006;22(9):511–9.

Thomas PD, Campbell MJ, Kejariwal A. PANTHER: a library of protein families and subfamilies indexed by function. Genome …. 2003;2129–41.

Unsicker SB, Kunert G, Gershenzon J. Protective perfumes: the role of vegetative volatiles in plant defense against herbivores. Curr Opin Plant Biol. 2009;12(4):479–85.

Velasco R, Zharkikh A, Troggio M, Cartwright D a, Cestaro A, Pruss D, et al. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS One. 2007;(12):e1326.

Venturini L, Ferrarini A, Zenoni S, Tornielli GB, Fasoli M, Dal Santo S, et al. *De novo* transcriptome characterization of *Vitis vinifera* cv. Corvina unveils varietal diversity. BMC Genomics; 2013;14(1):41.

Vogt T. Phenylpropanoid biosynthesis. Mol Plant. 2010;3(1):2–20.

Werck-Reichhart D, Feyereisen R. Cytochromes P450: a success story. Genome Biol. 2000;1(6):REVIEWS3003.

Zvi MM Ben, Shklarman E, Masci T, Kalev H, Debener T, Shafir S, et al. *PAP1* transcription factor enhances production of phenylpropanoid and terpenoid scent compounds in rose flowers. New Phytol. 2012;195(2):335–345

Chapter 4

Research results

Analytical profiling of the grapevine flower volatilome from nine different wine cultivars

Chapter 4 – Analytical profiling of the grapevine flower volatilome from nine different wine cultivars

4.1 Introduction

The co-expression results of the previous chapter suggest that terpenoid metabolism is highly regulated in flower organs with a great potential for both enzymatic and non-enzymatic modifications. The overrepresented expression of *VviTPSs* at inflorescence development and flowering is supported by the *in planta* characterisation of Cabernet Sauvignon flowers (Martin *et al.*, 2009). It is shown that flowers have a diurnal emission pattern with inflorescences expressing a valencene synthase (*VvValCS)* that localises to the pollen grains, with subsequent emissions of valencene and 7-epi-α-selinene. It is proposed that grapevine floral scent biosynthesis is unique due to the formation of terpenoids in male organs before bloom, with emissions increasing at anthesis. Analysis of 23 terpenoid synthases in the cultivar Muscato bianco supports the inflorescence localised expression of *VviTPSs* (Matarese *et al.*, 2014). These two cultivars, however, differed in terms of their major volatiles with Cabernet Sauvignon emitting valencene while Muscato bianco emitted β-caryophyllene. These findings suggest that grapevine flowers are important organs in terms of grapevine terpenoid metabolism and that cultivar differences do exist.

In order to understand terpenoid metabolism, characterisation of the flower scent from different plants is necessary. In the current study, nine different commercially important wine cultivars, with relevant aroma descriptors in wine, were selected for chemical characterisation. The results presented show that cultivar-specific terpenoids are prevalent and suggests that the aromatic differences seen in wines are even greater in grapevine flowers.

4.2 Materials and Methods

4.2.1 *V. vinifera* flower material

All plant material was obtained from a Vititec mother block in Stellenbosch (33°57'35.1"S 18°51'43.6"E). Specific clones from nine *V. vinifera* cultivars (Table 4.1) were selected based on descriptors related to flavour and aroma attributes of the resulting wines (http://vititec.com/product-category/scion/). Flowers were sampled between 08:30 am and 10:00 am on 23 October 2014. A composite sample of six to eight flowers was collected at stage 17/18 according to the modified Eichhorn and Lorenz (E-L) system (Coombe, 1995). Samples were kept on ice until they could be frozen using liquid nitrogen. Thereafter, the frozen flowers were removed from the rachis, milled to a fine powder and stored at -80°C.

Table 4.1 – Cultivar clones selected for characterisation of flower volatiles. Cultivar abbreviations used in this study are indicated along with the aromatic of wine used to identify potential aromatic cultivars.

Cultivar	Abbreviation	Clone	Cultivar Colour	Vititec descriptors*
Chardonnay	CH	CY 76 F	white	very aromatic
Chenin blanc	CB	SN 624 A	white	floral
Muscat D'Alexandrie	MA	HP 28 B	white	muscat aromas
Pinot noir	PN	PN 5B	red	-
Pinotage	PI	PI 48 C	red	fruity/berry aromas
Sauvignon Blanc	SB	SB 11 R	white	aromatic
Shiraz	SH	SH 21 K	red	berry/smoky
Viognier	VG	VR 1024	white	prominent flavours/aromas
Weisser Riesling	WR	WR 110 A	white	spicy/floral aromas

* The Vititec descriptors are associated with wines from the respective clones as available from the clonal reference sheet (http://vititec.com/product-category/scion/).

4.2.2 Flower volatile analysis using head-space GC-MS analysis

A method used for grape berry analysis (Young *et al.*, 2015) was adapted to for flowers. Approximately 10 mg of frozen tissue was weighed off directly into a 20 mL GC vial containing 2 mL tartrate extraction buffer (5 g/L tartaric acid, 2 g/L ascorbic acid and 8 mg/L sodium azide). Anisole-D8 (prepared in methanol) was added as an internal standard at a final concentration of 0.05 mg/L. Samples were prepared in triplicate and the vials sealed with a screw cap GC vial.

Solid phase micro-extraction (SPME) of the vial head space (HS) was done using a 50/30 µm grey divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) fiber (Supelco, Bellefonte, PA) that underwent pre-conditioning at 270°C for 60 min in the GC injection port according to the manufacturer's specifications.

Sample vials were pre-incubated for 5 min at 45°C in the autosampler heating chamber. The heating chamber was maintained at 45°C and agitated at 250 rpm to allow for equilibration of compounds between the sample and headspace. The fibre was inserted through the septa and exposed to the analytes in the headspace for 30 min, while maintaining the agitation speed and temperature at 250 rpm and 45°C, respectively. Desorbtion of the analytes took place in the GC injection port, where after, the fibre was maintained for 20 min in order to prevent any carryovers.

An Agilent 6890N gas chromatograph (Agilent, Palo Alto, CA) system coupled to a CTC CombiPal Analytics auto-sampler and an Agilent 5975B inert XL EI/CI MSD mass spectrometer detector through a transfer line was used for the analyses. A Zebron 7HG-G009-11 ZB-FFAP capillary

column (30 m x 250 ID µm, 0.25 µm film thickness) (Phenomenex, USA) was used. The desorption temperature for the analytes was 250°C for 5 min with a 10:1 split. Helium served as the carrier gas, having an initial flow rate of 1 mL/min. Initial oven temperature was maintained for 2 min at 40°C, followed by a linear increase of 10°C/min to a final temperate of 240°C which was held for an additional 2 min. The total run time was 24 min and the transfer line temperature was 250°C.

4.2.3 Identification and quantification of volatiles

Authentic standards for identification and quantification of volatiles were purchased from Sigma-Aldrich: (+)-valencene (≥ 70 %), *E*-β-farnesene (≥ 90 %), β-caryophyllene (≥ 80 %) and α-humelene (≥ 96 %). Stock solutions of the standards were prepared in methanol. A calibration curve was prepared in 2 mL tartrate buffer containing 0.05 mg/L Anisole-D8 (internal standard).

MSD ChemStation software (G1701-90057, Agilent) was used to visualise peaks in the total ion count (TIC) mode according to retention times. Peak areas were integrated using the auto-integration function in combination with manual integration in order to insure that all peaks of interest were processed. When no authentic standard was available, the Wiley 275 mass spectral library was used for compound identification. Where authentic standards were not available, concentration was calculated semi-quantitatively using the response of (+)-valencene as reference.

4.2.4 Anti-oxidant activity assays

Thin-layer chromatography was combined with the stable free radical 2,2-diphenyl-1-picrylhydrazyl (DPPH) assay to determine the total anti-oxidant activity of compounds present in extracts from flower tissue (Marston, 2011; Devi *et al.*, 2013). Approximately 200 mg of flower tissue was weighed off, followed by a single step extraction as described for carotenoids and chlorophylls by Lashbrooke *et al.* (2010), with the exception of that of N-ethyldiisopropylamine (NED) which was not included since it would potentially interfere with the anti-oxidant potential of compounds present in the extract. The final ethyl-acetate partition was concentrated six times under nitrogen and spotted in duplicate onto ALUGRAM Xtra SIL G/UV_{254} pre-coated TLC sheets (Macherey-Nagel, Germany). Samples were spotted 1.5 cm from the bottom with a 2 cm gap between spots (1.5 cm from the sides) using a 10 x 10 cm sheet. To minimise spreading of the spot size, the extract was applied sequentially in 2 µL aliquots (to a total volume of 12 µL). The solvent was air-dried and samples separated in a TLC tank with toluene-ethyl acetate (93:7) as mobile phase (Horváth *et al.*, 2010). Authentic standards were diluted in ethyl acetate to 1:10 (v/v) and a total volume of 5 µL was spotted. Authentic standards for β-caryophyllene, (+)-valencene, α-humelene and *E*-β-farnesene were used as controls for relative identification of terpenes in the extracts and activity of terpenic compounds present in the extracts. After chromatographic separation the solvent front was measured for calculation of Rf values and the sheets air-dried.

Anti-oxidant determination was performed according to Wagner (1996), using an ethanolic vanillin-sulfuric acid reagent (50 mL absolute ethanol, 300 µL sulfuric acid and 1 g vanillin) for derivatization. The derivatizing reagent was sprayed onto the TLC sheet using an atomiser (Separations Scientific, South Africa) and the TLC plates were placed in a 90°C oven for 7 min (or until coloured bands were detected). Plates were photographed with a digital camera and the positions of the bands noted.

Total anti-oxidant activity was determined by spraying 0.625 mM DPPH prepared in 50% acetone onto the plates using a TLC atomiser. Light exposure was kept to a minimum during spraying and sheets were placed in the dark to develop at room temperature for 30 min. Positive anti-oxidant activity was noted as the presence of a pale yellow spot against a purple background. The position of the spot was noted relative to the authentic standards.

4.3 Results

4.3.1 Volatile terpene profiles of grapevine flowers

The GC-MS analysis revealed that the flowers of grapevine almost exclusively emit sesquiterpenes. Minor volatiles, such as C6 hexenals, ketones and alkanes were detected, but were not analysed further due to their low abundance. For the purpose of this study analytical parameters, described by Young *et al*. (2015), were adjusted to detect and quantify monoterpenes and sesquiterpenes.

A total of 26 sesquiterpene volatiles were detected between the nine cultivars, whereas no monoterpenes were detected. The presence or absence of these volatiles indicates distinct differences between cultivars (Table 4.2). Five volatiles were common to all cultivars, albeit at varying levels: β-caryophyllene, *E*-β-farnesene, α-humelene, α-selinene and α-farnesene.

Major volatiles that were identified are shown in Figure 4.1A, with their percentage contribution shown in Figure 4.1B. Cultivar similarity based on their volatile profiles was determined by hierarchical clustering (Figure 4.1B). The phylogeny shows that CB, SB, WR and SH are very similar in terms of their volatile composition. CH and PI formed a clade dominated by α-farnesene. These two cultivars, in combination with with PN, show the least diversity in terms of volatiles as well the the lowest concentrations. The volatile ratios for PN are similar to VG with both being dominated by β-caryophyllene. VG, however, produced (+)-valencene that was not found in PN.

Total ion count (TIC) traces were overlaid for representative traces (MA, PN and SH) to show peak differences in the retention time window where most of the flower sesquiterpenes were found. The differences in response intensity are clearly seen with the corresponding sesquiterpene structures shown in Figure 4.2. Most notable are the differences observed for *E*-β-farnesene, (+)-valencene and 7-epi-α-selinene, with PN flowers not producing the latter two compounds.

Table 4.2 – Cultivar-specific sesquiterpenoids identified in flowers of grapevine is shown (average of three technical repeats). Volatiles accurately identified using an authentic standard are indicated with an asterisk, with the major volatile for each cultivar shown in bold font.

	WR	VG	SH	PN	MA	PI	SB	CH	CB
1h-indene, 2,3,3a,4-tetrahydro-3,3a,6-trimethyl-1-(1-methylethyl)	–	2.12 (± 0.04)	2.48 (± 0.06)	–	–	–	2.09 (± 0.01)	–	–
beta-elemene	–	2.13 (± 0.04)	2.50 (± 0.07)	–	–	–	2.10 (± 0.05)	–	–
beta-caryophyllene*	14.94 (± 2.06)	**227.74 (± 29.64)**	24.74 (± 1.88)	85.58 (± 6.35)	8.54 (± 0.41)	9.34 (± 1.80)	15.06 (± 1.17)	3.00 (± 0.23)	65.92 (± 0.52)
(+)-sativene	–	–	2.08 (± 0.04)	–	–	–	–	–	–
(-)-isoledene	–	2.48 (± 0.13)	3.04 (± 0.13)	–	2.13 (± 0.02)	–	2.48 (± 0.09)	–	3.16 (± 0.03)
alpha-gurjunene	–	2.44 (± 0.10)	2.95 (± 0.13)	–	2.12 (± 0.03)	–	2.45 (± 0.09)	–	3.11 (± 0.05)
alloaromadendrene	6.11 (± 0.74)	5.44 (± 0.52)	10.86 (± 0.79)	–	3.57 (± 0.13)	–	7.22 (± 0.55)	–	8.99 (± 0.10)
trans-beta-farnesene*	19.89 (± 2.56)	54.62 (± 6.39)	9.55 (± 0.96)	27.49 (± 1.79)	**263.09 (± 13.41)**	16.41 (± 4.76)	26.38 (± 3.72)	25.67 (± 3.77)	28.61 (± 1.91)
alpha-humulene*	5.14 (± 0.23)	10.77 (± 0.86)	6.79 (± 0.27)	6.02 (± 0.23)	5.42 (± 0.17)	4.38 (± 0.31)	5.62 (± 0.19)	3.59 (± 0.07)	7.34 (± 0.04)
beta/gamma-selinene	2.91 (± 0.22)	3.95 (± 0.27)	3.88 (± 0.16)	2.45 (± 0.12)	3.39 (± 0.19)	2.43 (± 0.44)	2.97 (± 0.11)	–	3.32 (± 0.03)
beta/gamma-selinene	15.54 (± 1.68)	10.13 (± 0.81)	24.03 (± 1.69)	–	6.46 (± 0.35)	–	15.73 (± 1.33)	–	17.35 (± 0.06)
cis-caryophyllene	9.45 (± 1.04)	–	14.45 (± 1.06)	–	4.44 (± 0.19)	–	9.51 (± 0.88)	–	–
eremophilene	–	7.37 (± 0.52)	–	–	–	–	10.65 (± 4.07)	–	10.52 (± 2.24)
(+)-valencene*	68.04 (± 8.39)	82.02 (± 10.12)	**204.52 (± 13.85)**	–	46.93 (± 2.58)	–	**135.64 (± 13.64)**	–	**160.85 (± 1.99)**
e,e-alpha-farnesene	–	–	–	8.21 (± 0.52)	–	22.21 (± 6.44)	–	7.29 (± 0.9)	–
alpha-selinene	12.79 (± 1.00)	11.53 (± 0.92)	13.52 (± 1.23)	9.47 (± 0.60)	11.67 (± 0.69)	9.25 (± 2.26)	8.19 (± 2.19)	5.88 (± 0.59)	8.33 (± 0.24)
unidentified sesquiterpene	5.63 (± 0.59)	–	–	–	–	–	–	–	–
unidentified sesquiterpene	2.39 (± 0.07)	–	8.77 (± 0.56)	–	–	–	5.78 (± 0.33)	–	6.36 (± 0.10)
aromadendrene	–	–	3.63 (± 0.18)	–	2.21 (± 0.10)	–	2.86 (± 0.28)	–	3.07 (± 0.06)
alpha-farnesene	**93.19 (± 11.33)**	50.49 (± 5.19)	31.68 (± 3.31)	**51.98 (± 2.20)**	63.03 (± 3.65)	**141.96 (± 42.35)**	92.08 (± 12.08)	**78.21 (± 12.19)**	121.14 (± 7.39)
delta-cadinene	–	2.46 (± 0.07)	–	–	–	–	2.54 (± 0.16)	–	2.57 (± 0.08)
beta-cadinene	–	–	2.77 (± 0.09)	–	–	–	–	–	–
germacrene a	–	–	–	2.51 (± 0.33)	–	2.52 (± 0.49)	–	–	–
7-epi-alpha-selinene	89.81 (± 9.97)	53.02 (± 6.31)	141.35 (± 10.39)	–	31.49 (± 1.85)	–	87.86 (± 8.65)	–	100.53 (± 0.41)
beta-selinene	–	–	2.89 (± 0.04)	–	–	–	2.52 (± 0.02)	–	2.57 (± 0.02)

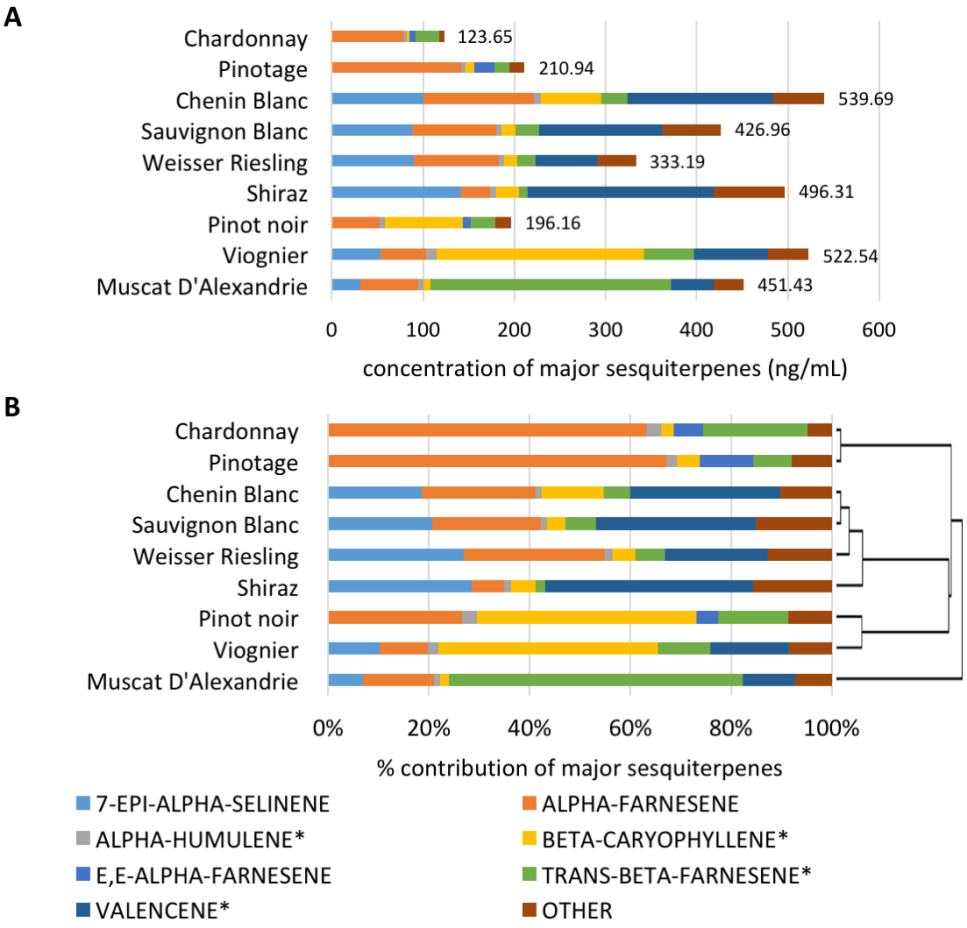


Figure 4.1 – Major volatiles for the respective cultivars are shown. The relative abundance of sesquiterpenoids and differences in total content is shown (A) with relative ratios as a percentage contribution to the total shown in (B). Volatiles quantified and accurately identified using an authentic standard are indicated with an asterisk

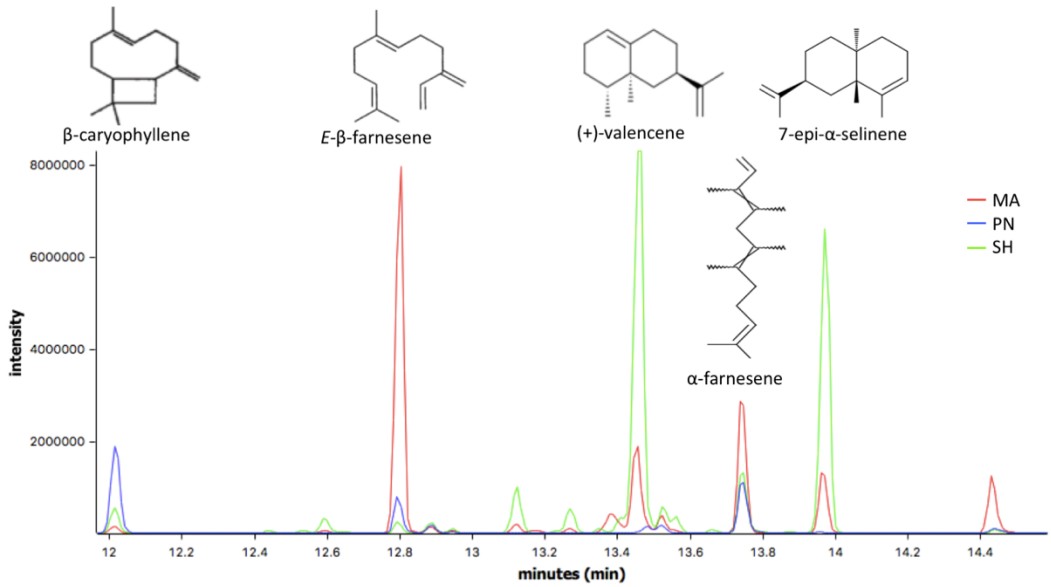


Figure 4.2 – Representative total ion count (TIC) traces are shown to illustrate cultivar differences. Muscat D'Alexandrie (MA), Pinot noir (PN) and Shiraz (SH) traces in the 12 to 15 minute window where most flower sesuiterpenoids were found is shown. Chemical structures of the major peaks are shown to illustrate skeletal diversity of the compounds.

4.3.2 Anti-oxidant activity of flower extracts

No clear differences were observed when cultivar-specific extracts were analysed. All extracts resulted in the formation of identical zones when sprayed with DPPH or derivitized with the ethanolic vanillin-sulfuric acid reagent. However, some of these zones could be correlated to those found for authentic standards. Figure 4.3 shows the representative extracts from the SH and VG cultivars compared to the authentic standards for (+)-valencene, β-caryophyllene and trans-β-farnesene with (A) showing DPPH staining while (B) shows the dervitized TLC plate. α-Humulene showed no definitive anti-oxidant zone and was excluded from the image. All cultivars showed three distinct bands (marked i – iii) that correlate with (+)-valencene, β-caryophyllene and trans-β-farnesene respectively. Derivitization of (+)-valencene resulted in numerous bands, but DPPH staining shows a single zone. Zones on the elution front indicate that compounds not separated by the chosen solvent also have anti-oxidant activity. Carotenes and chlorophylls also formed part of the extract and are likely responsible for the green/red zone seen at the sample loading area.

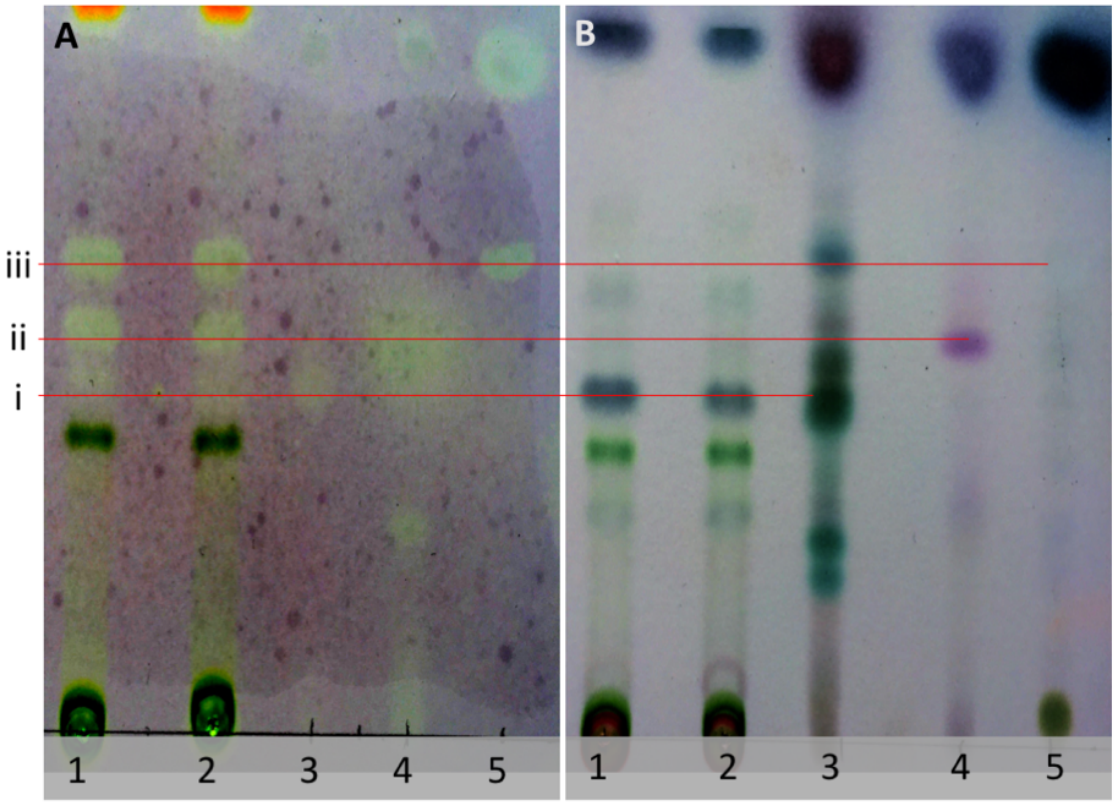


Figure 4.3 - TLC separation and antioxidant activity detection of terpenes. Extracts from flowers from Shiraz (1) and Viognier (2) can be compared to the authentic standards for (+)-valencene (3), β-caryophyllene (4) and trans-β-farnesene (5). **3A**, DPPH assay to determine anti-oxidant activity can be seen as pale yellow zones against a purple background. Derivitization of the terpenic compounds in **3B** can be seen as coloured bands.

4.4 Discussion

4.4.1 Different grapevine cultivars have distinct terpenoid metabolism

The analyses of flower sesquiterpene volatiles reveal unique cultivar differences. Both concentration (amplitude) and relative ratios contributed to the cultivar-specific volatiles. *In silico* analysis (Chapter 3) shows a significant upregulation of sesqui-TPS-encoding genes in flowering. Martin *et al*. (2009) have previously shown that the anthers are predominantly responsible for sesquiterpene biosynthesis at the onset of bloom.

The sampling time selected for the current study was chosen to capture volatiles at the developmental stage where inflorescence development was complete and flowering (bloom) had commenced. The lack of monoterpenes in our analysis was supported by the *in silico* analysis that showed that genes expressing monoterpene synthase were either not expressed at flowering (to the same degree as sesqui-TPS encoding genes) and/or that the monoterpene volatiles are in a conjugated form and therefore not volatile. We showed in Chapter 3 that glycosyltransferase encoding genes co-expressed with the TPSs, supporting the idea that some of the volatiles were conjugated with sugars, rendering them non-volatile with our method. It is possible that monoterpenes are present in flowers, but the analysis method used was optimised for free volatiles. Conjugated terpenoids could, therefore, be detected through a method by which conjugated terpenoids are volatilised, for example, by enzyme hydrolysis (Perestrelo *et al.*, 2012).

Previous studies have demonstrated that *in vitro* expression of *VvVal* resulted in (+)-valencene and 7-epi-α-selinene as major products (Lücker *et al.*, 2004). These volatiles were most abundant in Cabernet Sauvignon flowers along with *E,E*-α-farnesene (Martin *et al.*, 2009). Four cultivars (CH, SB, WR and SH) showed a similar volatile profile to those reported in Cabernet Sauvignon flowers (Martin *et al.*, 2009) but lacked *E,E*-α-farnesene as minor product. Shiraz has a distinct peppery aroma (most notable in in certain Australian wine styles) (Siebert *et al.*, 2008; Wood *et al.*, 2008). This aroma has been linked to the sesquiterpene ketone (-)-rotundone and it is shown that the aerial oxidation (namely, non-enzymatic) of α-guaiene results in the formation of this volatile and has an extremely low odour threshold (16 ng/L in wine) (Huang *et al.*, 2014). Wedler *et al*. (2015) propose a mechanism in which germacrene A serves as precursor for a carbocation cascade that facilitates the radical-initiated oxidation of α-guaiene to result in rotundone. Germacrene A is a common precursor for numerous sesquiterpenes (Bülow & Konig, 2000; Adia, 2005). It is, thus, likely that the TPS genes over-represented in grapevine flowers prefer the metabolic pathway where FPP proceeds to the *E,E*-germacradienyl cation via germacrene A as an intermediate for cyclic sesquiterpene diversity. The volatiles of SH suggest that this cultivar has a TPS gene(s) that facilitates the formation of cyclic sesquiterpenes rather than the acyclic compounds (such as farnesene isomers).

The current analysis identified three farnesene-type acyclic volatiles. Library identification was achieved for α-farnesene and its stereoisomer *E,E*,-α-farnesene while *E*-β-farnesene was accurately identified using an authentic standard. In poplar it has been shown that these volatiles are involved in herbivore-induced attraction of natural predators as an indirect plant-defense response (Arimura *et al.*, 2004; Frost *et al.*, 2007). However, studies show that grapevines emitting these volatiles induced oviposition, specifically attracting the female berry moth (*Paralobesia viteana*) (Cha *et al.*, 2008). Studies on grapevine berry moths *Lobesia botrana* and *P. viteana* demonstrate cultivar and species-specific attraction of the respective moths. It is difficult to conclude on any general trends since Solaris, a Swiss *V. vinifera* cultivar, was used to determine the mixture that attracts *L. botrana*; while a non-vinifera species (*V. riparia*) was used in the *P. viteana* attraction study (Cha *et al.*, 2008; von Arx *et al.*, 2011). The α-farnesene and *E*-β-farnesene isomers along with β-caryophyllene were present in flowers from all the cultivars but at vastly different ratios while *E,E*,-α-farnesene was present only in flower from CH, PI and PN. CH and PI flowers were dominated by α-farnesene (more than 60 percent of total volatiles) while PN and VG were dominated by β-caryophyllene at roughly 40 percent. *E,E*,-α-farnesene was found to be one of the major flower volatiles for the cultivars CH, SB, WR and SH. The presence of these volatiles suggests a cultivar-specific flower volatile composition that could fulfil similar evolutionary attractant roles.

Poplar tree species form unique mixtures of volatile sesquiterpenes that function as attractants, and are independent of the insect that induces the response (Danner *et al.*, 2011). Whether these volatiles are emitted as a similar broad insect attractor in grapevine, is unknown. From a commercial perspective, the attraction of berry moths is unwanted as the damage caused on grapes through oviposition facilitates secondary infections by fungal pathogens like *Botrytis cinerea* (Tasin *et al.*, 2005). Insect attraction is a complex response to stimulus with the mixture of volatiles playing an important role in the effectiveness of the response. Entomological assays with reconstituted mixtures show that sesquiterpenes act as important constituents for attraction, but only when applied in specific ratios with other volatiles like pentadecane, hexenyl acetate and hexanol (Tasin *et al.*, 2005; von Arx *et al.*, 2011). Our analysis shows traces of decanes and C_6 volatiles and further studies on grapevine flowers will probably identify further role players in flower volatile emissions.

MA (a known terpenic cultivar) shows a distinctly different flower volatile profile compared to the other cultivars analysed (Figure 4.1B). It produces high levels of sesquiterpenes, at a total concentration of 451.43 ng/mL, with *E*-β-farnesene contributing to nearly 60 percent of the volatiles and (+)-valencene, 7-epi-α-selinene and α-farnesene present as minor products. *E*-β-farnesene is shown to be one of the major aphid alarm pheromones which in turn acts as a kairomone (attractant) for predatory insects like ladybirds (*Adalia bipunctata*) (Francis *et al.*, 2004).

4.4.2 Different flower volitalomes resulted in near identical anti-oxidant potential

The major flower volatiles ((+)-valencene, β-caryophyllene and *E*-β-farnesene) identified through GC-MS show anti-oxidant activity that correlates to specific zones from grapevine flower extracts (Figure 4.3). The assay used in this study allowed for relatively quick screening of flower extracts and gives a qualitative indication of total anti-oxidant potential. This assay has been used with relative success to study monoterpene and essential oil anti-oxidant activity and is deemed to be an inexpensive and quick screening tool (Horváth *et al.*, 2010; Marston, 2011; Devi *et al.*, 2013). For the purpose of this study it was confirmed that compounds extracted from flowers correlate with sesquiterpene standards that show anti-oxidant activity, but the assay did not allow discrimination between the cultivars in terms of their anti-oxidant responses.

4.5 Conclusion

Based on the results presented it can be concluded that there are cultivar-specific metabolic processes regulating flower terpenoid biosynthesis. The diversity in flower volatiles suggests that it can be expected that cultivar specific sub-functionalisation of TPS-encoding genes facilitates a flux towards certain volatiles. Cultivar differences in flower volatile terpenes potentially influence plant-insect/insect-insect interactions and are promising for identifying key volatiles that might serve as kairomones (for biological control). Preliminary results also indicate that the major floral sesquiterpenes can act as anti-oxidants, which suggests, an involvement in redox homeostasis during flowering. The complex nature of grapevine flower volatile metabolism, and the cultivar variations seen, raises questions on the function and regulation of these versatile compounds. Compared to berry ripening, terpenoid biosynthesis is more pronounced in flowers; whether this has an impact on the terpenes found in berries and wine, is still not known.

The detection of sesquiterpenes in wine is uncommon with rotundone being an exception in only certain Shiraz wine styles. The up-regulation of sesqui-TPS-encoding genes in flowers shown *in silico* correlates well with the sesquiterpene emissions observed, confirming a direct link between transcription and emission. Investigating the underlying genes involved in sesquiterpene metabolism will shed light on the sub-functionalisation of TPSs and their impact on cultivar diversity.

4.6 References

Adia AM. Isolation and structure elucidation of sesquiterpenoids from the essential oils of some liverworts (*Hepaticae*). University of Hamburg; 2005.

Arimura G, Huber DPW, Bohlmann J. Forest tent caterpillars (*Malacosoma disstria*) induce local and systemic diurnal emissions of terpenoid volatiles in hybrid poplar (*Populus trichocarpa* × *deltoides*): cDNA cloning, functional characterization, and patterns of gene expression of (−)-germacr. Plant J. 2004;37(4):603–616.

von Arx M, Schmidt-Büsser D, Guerin PM. Host plant volatiles induce oriented flight behaviour in male European grapevine moths, *Lobesia botrana*. J Insect Physiol. Elsevier Ltd; 2011;57(10):1323–1331.

Bülow N, Konig W. The role of germacrene D as a precursor in sesquiterpene biosynthesis: investigations of acid catalyzed, photochemically and thermally induced rearrangements. Phytochemistry. 2000;55(2):141–168.

Cha DH, Nojima S, Hesler SP, Zhang A, Linn CE, Roelofs WL, et al. Identification and field evaluation of grape shoot volatiles attractive to female grape berry moth (*Paralobesia viteana*). J Chem Ecol. 2008;34:1180–9.

Coombe BG. Adoption of a system for identifying grapevine growth stages. Aust J Grape Wine Res. 1995;1:100–110.

Danner H, Boeckler GA, Irmisch S, Yuan JS, Chen F, Gershenzon J, et al. Four terpene synthases produce major compounds of the gypsy moth feeding-induced volatile blend of *Populus trichocarpa*. Phytochemistry. 2011;72:897–908.

Devi RS, Radhamany PM, Devi VG. Investigation of the antioxidant principles from *Psilanthus travancorensis* (WT. & ARN.) Leroy - An unexplored taxon of *Rubiaceae*. Int J Pharm Pharm Sci. 2013;5:13–7.

Francis F, Lognay G, Haubruge E. Olfactory responses to aphid and host plant volatile releases: (*E*)-β-farnesene an effective kairomone for the predator *Adalia bipunctata*. J Chem Ecol. 2004;30(4):741–55.

Frost CJ, Appel HM, Carlson JE, De Moraes CM, Mescher MC, Schultz JC. Within-plant signalling via volatiles overcomes vascular constraints on systemic signalling and primes responses against herbivores. Ecol Lett. 2007;10:490–8.

Horváth G, Jámbor N, Végh A, Böszörményi A, Lemberkovics É, Héthelyi É, et al. Antimicrobial activity of essential oils: the possibilities of TLC-bioautography. Flavour Fragr J. 2010;25(3):178–82.

Huang A-C, Burrett S, Sefton MA, Taylor DK. Production of the pepper aroma compound, (-)-rotundone, by aerial oxidation of α-guaiene. J Agric Food Chem. 2014;62(44):10809–15.

Lashbrooke JG, Young PR, Strever AE, Stander C, Vivier MA. The development of a method for the extraction of carotenoids and chlorophylls from grapevine leaves and berries for HPLC profiling. Aust J Grape Wine Res. 2010;16(2):349–60.

Lücker J, Bowen P, Bohlmann J. *Vitis vinifera* terpenoid cyclases: Functional identification of two sesquiterpene synthase cDNAs encoding (+)-valencene synthase and (-)-germacrene D synthase and expression of mono- and sesquiterpene synthases in grapevine flowers and berries. Phytochemistry. 2004;65:2649–59.

Marston A. Thin-layer chromatography with biological detection in phytochemistry. J Chromatogr A. 2011;1218:2676–83.

Martin DM, Toub O, Chiang A, Lo BC, Ohse S, Lund ST, et al. The bouquet of grapevine (*Vitis vinifera* L. cv. Cabernet Sauvignon) flowers arises from the biosynthesis of sesquiterpene volatiles in pollen grains. Proc Natl Acad Sci U S A. 2009;106(17):7245–50.

Matarese F, Cuzzola A, Scalabrelli G, D'Onofrio C. Expression of terpene synthase genes associated with the formation of volatiles in different organs of *Vitis vinifera*. Phytochemistry. 2014

Perestrelo R, Caldeira M, Câmara JS. Solid phase microextraction as a reliable alternative to conventional extraction techniques to evaluate the pattern of hydrolytically released components in Vitis vinifera L. grapes. Talanta. Elsevier B.V.; 2012 Jun 15;95:1–11.

Siebert TE, Wood C, Elsey GM, Pollnitz AP. Determination of rotundone, the pepper aroma impact compound, in grapes and wine. J Agric Food Chem. 2008;56(10):3745–8.

Tasin M, Anfora G, Ioriatti C, Carlin S, De Cristofaro A, Schmidt S, et al. Antennal and behavioral responses of grapevine moth *Lobesia botrana* females to volatiles from grapevine. J Chem Ecol. 2005;31(1):77–87.

Wagner H. Plant drug analysis: a thin layer chromatography atlas. 2nd ed. Berlin: Springer; 1996.

Wedler H, Pemberton R, Tantillo D. Carbocations and the complex flavor and bouquet of wine: mechanistic aspects of terpene biosynthesis in wine grapes. Molecules. 2015;20:10781–92.

Wood C, Siebert TE, Parker M, Capone DL, Elsey GM, Pollnitz AP, et al. From wine to pepper: Rotundone, an obscure sesquiterpene, is a potent spicy aroma compound. J Agric Food Chem. 2008;56(10):3738–44.

Young P, Eyeghe-Bickong HA, du Plessis K, Alexandersson E, Jacobson DA, Coetzee ZA, et al. Grapevine plasticity in response to an altered microclimate: Sauvignon Blanc modulates specific metabolites in response to increased berry exposure. Plant Physiol. 2015

Chapter 5

Research results

Isolation and functional characterisation of grapevine flower sesquiterpene synthases

Chapter 5 – Isolation and functional characterisation of grapevine floral sesquiterpene synthases

5.1 Introduction

The grapevine genome has been found to be highly heterozygous and having extensive gene duplications as a result of domestication (Velasco *et al.*, 2007). To reduce the complexity and difficulties in assembling the genome, the near homozygous in-bred PN40024 genome (Jaillon *et al.*, 2007) was selected as the reference genome instead of the heterozygous ENTAV-115 genome (Velasco *et al.*, 2007). However, the post genomics era has revealed that a single genome is not sufficient to explain genetic diversity between individuals of a species; hence the concept of pan-genomics.

The pan-genome (or supra-genome) is a representative genome that captures all the genes found within a phylogenetic clade (for example, species or genus). A pan-genome consists of the core genes that are present in all individuals of the clade as well as the genes that are variable between the individuals. The variable genes are considered dispensable and are predominantly responsible for the genetic diversity within a pan-genome (Saxena *et al*., 2014). This diversity is influenced by the gene structural variation (SV) between the representative individuals of a species. SV includes the genomic variations of insertions/deletions (InDels), inversions, translocations and copy number variants (CNVs). An extreme form of CNV ere the presence/absence variants (PAVs), which refer to genes/genomic regions present in one individual, but not in another (Saxena *et al.*, 2014; Golicz *et al.*, 2015). Core genes generally regulate the essential functions in an organism and are critical for the survival or characteristic aspects of the individuals in a clade. Dispensable genes usually allow for the evolution of non-essential genes that increase adaption and survivability and are, therefore, closely linked to genes involved in secondary/specialised metabolism. These groupings are not mutually exclusive, with genes involved in secondary metabolism being present in all individuals, for example, classified as core genes.

Martin *et al.* (2010) provided a framework for *V. vinifera* TPS genes when they characterised this overrepresented gene family in the PN40025 Pinot noir genome sequence (Jaillon *et al*., 2007). Extensive duplication events, pseudogenes and the remnants of retrotransposable elements are associated with this gene family, suggesting that the family contains extensive SVs. These authors provided further insights in terms of functional genes by isolating a total of thirty-nine functional genes from three cultivars. The authors alluded to cultivar variants but only discussed a representative gene from one of the three cultivars used for isolation. However, the potential for cultivar specific SVs and their potential influence on the TPS gene family was not pursued.

In the current study, selected TPS genes were targeted from different cultivars to assess the influence of SVs on a cultivar's ability to produce terpenoids. It is, therefore, hypothesised that SVs of TPS-encoding genes contribute to the typicity of different cultivars.

5.2 Materials and Methods

5.2.1 Isolation of putative TPS-encoding genes

Total RNA was isolated from grapevine flowers at E-L18 for the *V. vinifera* cultivars Muscat D'Alexandrie (MA), Pinot noir (PN), Sauvignon Blanc (SB) and Shiraz (SH) according to the method described by Reid *et al.* (2006). RNA was purified using the Bioline Isolate II Plant RNA kit (Celtic Molecular Diagnostics, RSA) followed by genomic DNA removal through on-column digestion with DNase I as described in the product manual of the aforementioned kit. The integrity of the RNA was assessed on an agarose gel followed by cDNA synthesis using the ImProm-II Reverse Transcription System (Promega, USA).

Putative TPS-encoding genes were isolated from cDNA pools using Phusion High Fidelity DNA polymerase (Thermo Scientific, RSA) followed by gel purification of PCR products with the Qiagen Gel Extraction kit (Qiagen, USA). Primers used for gene isolation (Table 5.1) were designed using the Genoscope 12X genome repository (http://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/). Gene targets were updated to the V1 genome IDs using the conversion tables available via CRIBI (http://genomes.cribi.unipd.it/grape/). For ease of reference numeric codes were assigned to the respective gene targets (Table 5.1).

5.2.2 Cloning of and sub-cloning putative TPS genes into yeast expression vectors

Purified PCR products were A-tailed by incubation with ExTaq (Seperations, RSA) and ligated into the pGEM-T Easy vector (Promega, USA) according to the supplier's protocol. Ligated vectors were transformed into competent *Escherichia coli* DH5α cells and plated onto antibiotic selection plates. Positive transformants were confirmed through PCR, using T7 and SP6 primers described in the Promega pGEM-T Easy product manual, followed by plasmid isolations with the GenElute Plasmid Miniprep kit (Sigma-Aldrich, RSA). Putative TPS-encoding genes underwent bi-directional sequencing (Central Analytical Facility, Stellenbosch University, South Africa) using T7/SP6 primers (Promega, USA) in combination with custom designed walking primers.

Isolated plasmids from confirmed positive transformants were digested with the appropriate enzymes (Table 5.1) for ligation into an inducible pBluescript (pBS) yeast expression vector: pBS::GAL::Cyct1::URA3::. The pBS vector was prepared for ligation with T4 DNA ligase (Promega, Madison, USA) through restriction digests with appropriate enzyme compatible with the genes of interest. Competent *E. coli* cells were transformed and verified through colony PCR. Plasmids were isolated according to the alkaline lysis method (Engebrecht *et al.*, 1991). Recombinant

Table 5.1 – Primer sequences for amplification of TPS genes. Primer modifications are underlined in lower font with the melting temperature (Tm) excluding this modification. The TPS code listed here is used throughout the study. The initial sequence reference (Gensoscope Acc. ID) used to design the primers and the ID conversion to that of CRIBI V1 is indicated.

Code	Primer Position	Primer Sequence (5'-3')	RE Mo	Phusion Tm (w/o RE mod)	Genoscope Acc. ID	V1 CRIBI ID
TPS01	3'	gtcgacTCATATTGGCACAGGGTCTA	SalI	61.0	GSVIVT01036361001	VIT_18s0001g05240
	5'	ggatccATGGCCTTAATTCTCGCA	BamHI	61.3		
TPS02	3'	gtcgacTCATATTATGGGGTCAATGAGC	SalI	62.8	GSVIVT01036312001	VIT_18s0001g04080
	5'	tctagaATGTCTACTCAAGTCTCAGAATGTCCT	XbaI	63.7		
TPS03	3'	gtcgacTTATATTACAGCGTTGATCAAGACA	SalI	61.7	GSVIVT01036322001	VIT_18s0001g04280
	5'	ggatccATGTCGGGTCAAGTCTTAGC	BamHI	61.2		
TPS04	3'	gtcgacCTCTCCACTTACATTGGCAC	SalI	60.3	GSVIVT01014174001	VIT_19s0014g01060
	5'	ggatcCATGTCTGTCTCACTTCCCG	BamHI	60.9		
TPS05	3'	gtcgacTCATATTGGCACAGGGTCTAT	SalI	61.3	GSVIVT01036344001	VIT_18s0001g04780
	5'	ggatccATGGCCTTAATTCTCGCTAC	BamHI	60.2		
TPS06	3'	actagtTCAGGCTGCTCTCTCAAAGAGGA	SpeI	69.0	GSVIVT01001155001	VIT_07s0151g01040
	5'	ggatccATGTCTTCTCAATCCACTATTCTCT	BamHI	60.5		
TPS07	3'	gtcgacTTATATTATAGGGTTAATGAAGAAAGCC	SalI	61.0	GSVIVT01036366001	VIT_18s0001g05290
	5'	ggatccATGTCTACTCAAGTCTCAGCATG	BamHI	60.5		
TPS08	3'	gtcgacTCATATTGGCACAGAATCTATAAG	SalI	61.9	GSVIVT01014566001	VIT_19s0014g04900
	5'	ggatccATGTCTGTTCAGTCTTCAGTGG	BamHI	61.3		
TPS09	3'	gtcgacTCATATTGGCACARGGTCAA	SalI	71.7	GSVIVT01014558001	VIT_19s0014g04810
	5'	ggatccATGTCTCTTCCACTCTCAGTTACA	BamHI	71.6		
TPS10	3'	gtcgacTCATGCGATAGGGTGAATG	SalI	62.6	GSVIVT01036348001	VIT_18s0001g04870
	5'	ggatccATGTCGGGTCAAGTCTTAGC	BamHI	61.2		
TPS11	3'	gtcgacTCATACTATGGGGTCAATGAGC	SalI	63.0	GSVIVT01036308001	VIT_18s0001g04050
	5'	ggatccATGTCTATTCAAGTCTCAACGTGT	BamHI	61.0		
TPS12	3'	gtcgacTCATATTGGCACAGGATCAA	SalI	62.0	GSVIVT01014569001	VIT_19s0014g04930
	5'	ggatccATGTCTGTTCAGTCTTCAGTGG	BamHI	61.3		

plasmids were linearised with the *Apa*I restriction enzyme and transformed to a modified GT051 *Saccharomyces cerevisiae* strain according to the TRAFO method (Gietz & Woods, 2002). The strain was modified from the W303a strain described by Thomas and Rothstein (1989) to increase the metabolic flux to the FDP terpene precursor by over-expression of *HMG1* (truncated) and *IDI1* genes (Bezuidenhoudt, I., IWBT, Stellenbosch University). Yeast transformants were plated on modified TRAFO synthetic drop-out plates that contained galactose as carbon source with the amino acids adenine, leucine and uracil omitted to maintain selective pressure (Gietz & Woods, 2002). Putative yeast transformants were verified by colony PCR screening.

5.2.3 Computational analysis of isolated TPS-encoding sequences

Full-length gene sequences were manually curated and analysed using the CLC Main Workbench software (CLC Bio-Qiagen, Aarhus, Denmark). The exon-intron organisation was determined with the NCBI-SPIDEY alignment tool (http://www.ncbi.nlm.nih.gov/spidey/) by comparing the isolated coding sequences to the reference genome sequence (V1 CRIBI ID in Table 5.1). The exon-intron gene structures were visualised using FancyGene online tool (http://bio.ieo.eu/fancygene/). The location of functional domains was determined using the Pfam protein family database (Finn *et al.*, 2014). The DDXXD, RRX_8W and NSE/DTE motifs characteristic of TPS proteins were identified using the FIMO tool of the MEME suite (Bailey *et al.*, 2009; Martin *et al.*, 2010; Grant *et al.*, 2011). Putative gene identification was performed using standard nucleotide and protein BLAST analysis through NCBI (http://www.ncbi.nlm.nih.gov/) and CRIBI (http://genomes.cribi.unipd.it/grape/).

Multiple sequence alignments of predicted protein sequences were performed against the V1 CRIBI predicted protein sequences using the CLC Sequence Viewer (CLC Bio-Qiagen, Aarhus, Denmark). Multiple sequence alignment parameters were as follow: gap open cost at two, gap extension cost at one with the end-gap cost as any other. A phylogenetic tree was constructed using the UPGMA method (Michener & Sokal, 1957) with Jukes-Cantor as the distance measure and 100 bootstrapping replicates. The percentage of amino acid residues that match that of the predicted protein sequence was determined through pairwise sequence alignment using the T-Coffee alignment tool (http://www.ebi.ac.uk/Tools/msa/tcoffee/).

5.2.4 Volatile terpene analysis from yeast cultures

Synthetic complete drop-out (SC) media (Gietz & Woods, 2002) was supplemented with $MgSO_4$ to a final Mg^{2+} concentration of 5 mM and buffered to a pH of 6.0 using citrate buffer. Pre-cultures of the respective yeast transformants were prepared in SC media with glucose (2 % w/v) as a carbon source. Cells were harvested through centrifugation (13,000 x g) and washed with sterile water. TPS-expression was induced in sealed 20 mL GC-vials containing 5 mL SC media with galactose as carbon source. Assays were performed in triplicate (three positive transformants). The starting optical density (OD) was 0.7 at 600 nm. After 16 h of induction at 30ºC with shaking, vials were

placed at 4ºC for 1 h before analysis. A 1 mL mixture of natamycin (Delvocid, 2 mg/mL) and anisole-D8 (internal standard, 50 µg/L) was added to each vial by piercing the vial septa using a sterile syringe. Delvocid was prepared in 0.1 M NaOH and anisole-D8 was prepared in acetonitrile.

Solid phase micro-extraction (SPME) of the vial head space (HS) was done using a 50/30 µm grey divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) fibre (Supelco, Bellefonte, PA) that underwent pre-conditioning at 270°C for 60 min in the GC injection port according to the manufacturer specifications.

Sample vials were pre-incubated for 5 min at 35°C in the autosampler heating chamber. The heating chamber was maintained at 35°C and agitated at 250 rpm to allow for equilibration of compounds between the sample and headspace. The fibre was inserted through the septa and exposed to the analytes in the headspace for 20 min, while maintaining the agitation speed and temperature at 250 rpm and 35°C, respectively. Desorbtion of the analytes took place in the GC injection port where after the fibre was maintained for 20 min in order to prevent any carryovers.

An Agilent 6890N gas chromatograph (Agilent, USA) system coupled to a CTC CombiPal Analytics auto-sampler and an Agilent 5975B inert XL EI/CI MSD mass spectrometer detector through a transfer line was used for the analyses. A Zebron 7HG-G009-11 ZB-FFAP capillary column (30 m x 250 ID µm, 0.25 µm film thickness) (Phenomenex, USA) was used. Desorption temperature for the analytes was 250°C for 5 min with a 10:1 split. Helium served as carrier gas with an initial flow rate of 1 mL/min. Initial oven temperature was maintained for 2 min at 40°C, followed by a linear increase of 10°C/min to a final temperature of 240°C which was held for an additional 2 min. The total run time was 24 min and the transfer line temperature was 250°C.

5.3 Results

5.3.1 Putative TPS-encoding genes isolated from grapevine flowers

Initial PCR screening of cDNA pools resulted in 32 positive amplifications from four cultivar specific cDNA pools, of which 23 were subsequently cloned. Five putative TPS-encoding genes were isolated from four cultivars, resulting in 23 cultivar-specific gene variants. The sequence characteristics for 21 of these predicted TPS-encoding genes are shown in Table 5.2. TPS06 was putatively characterised as a diterpene synthase but was not included for further analysis since the heterologous system is not optimised for diterpene characterisation. Nucleotide sequence length was indicated as well as the position of the premature stop codon if the ORF was non-functional. Pre-mature stop codons were present in six of the cultivar variants.

TPS-encoding genes that were predicted as functional were translated to their corresponding amino acid sequences. The N-terminal RRx_8W and C-terminal DDxxD and NSE/DTE motif sequences characteristic of TPSs are shown in Figure 5.1

Table 5.2 – Nucleotide and protein sequence analysis TPS-encoding genes isolated from Shiraz (SH), Sauvignon Blanc (SB), Pinot noir (PN) and Muscat D'Alexandrie (MA) cDNA pools. Nucleotide length and the presence of an open reading frame (ORF) is shown. If no ORF was found the nucleotide position of the premature stop codon is shown. Protein sequence length, functional domain/region, motif positions as well as motif sequences are shown.

		Nucleotide Sequence		Protein Sequence									
							RR(x)8W motif		DDxxD motif		NSE/DTE motif		
Target	Cultivar	Length	Functional ORF	Length	N-Terminal Region	C-Terminal Domain	Seq.	Position	Seq.	position	Seq.	Position	
TPS02	SH	1668	yes	556	29 - 203	234 - 499	RPRAKFHPNIW	19 -29	DDIYD	309-313	FMDDMTSHKFE	451-461	
	SB	1668	yes	556	29 - 203	234 - 499	RPRAKFHPNIW	19 -29	DDIYD	309-313	FMDDMTSHKFE	451-461	
	PN	1668	stop (774)	-	-	-	-	-	-	-	-	-	
	MA	1668	stop (355)	-	-	-	-	-	-	-	-	-	
TPS05	SH	2037	alternative splicing	-	-	-	-	-	-	-	-	-	
	SB	1682	stop (577)	-	-	-	-	-	-	-	-	-	
	PN	1662	yes	554	33 - 200	231 - 495	RQTANYQPSIW	23 -33	DDIFD	306 - 310	LMNDMATHKFE	447 -457	
	MA	1683	yes	561	33-208	239 - 503	RQTANYQPSIW	23 -33	DDIFD	314 - 318	LMNDMATHKFE	454 -464	
TPS08	SH	1674	yes	558	32 - 204	235 - 501	RRCANFHPSIW	22 -32	DDIYD	310 - 314	LMDDMVSHKFE	452 - 462	
	SB	1674	yes	558	32 - 204	235 - 501	RRCANFHPSIW	22 -32	DDIYD	310 - 314	LMDDMVSHKFE	452 - 462	
	PN	1674	yes	558	32 - 204	235 - 501	RRCANYHPSIW	22 -32	DDIYD	310 - 314	LMDDMVSHKFE	452 - 462	
	MA	1674	yes	558	32 - 204	235 - 501	RRCANYHPSIW	22 -32	DDIYD	310 - 314	LMDDMVSHKFE	452 - 462	
TPS11	SH	1668	yes	556	29 -203	234 -499	RPVAEFXPSIW	22 -32	DDIYD	309 - 313	LMDDITTHKFE	451 -461	
	SB	1668	yes	556	29 -203	234 -499	RPMAEFHPSIW	22 -32	DDIYD	309 - 313	LMDDITTHKFE	451 -461	
	PN	1668	yes	556	29 -203	234 -497	RPMAEFHPSIW	22 -32	DDIYD	309 - 313	LMDDITTHKFE	451 -461	
	MA	1668	yes	556	29 -203	234 -499	RPVAEFHPSIW	22 -32	DDIYD	309 - 313	LMDDITTHKFE	451 -461	
TPS12	SH	1673	stop (673)	-	-	-	-	-	-	-	-	-	
	SB-1	1673	stop (673)	-	-	-	-	-	-	-	-	-	
	SB-2	1169	stop (418)	-	-	-	-	-	-	-	-	-	
	PN	1674	yes	558	32 -204	235 -501	RRCANFHPSIW	22 - 32	DDIYD	310 -314	LMDDMVSRKFE	452 - 462	
	MA	1677	yes	559	32 - 205	236 - 502	RRCANFHPSIW	22 - 32	DDIYD	311 - 315	LMDDMVSHKFE	453 - 463	

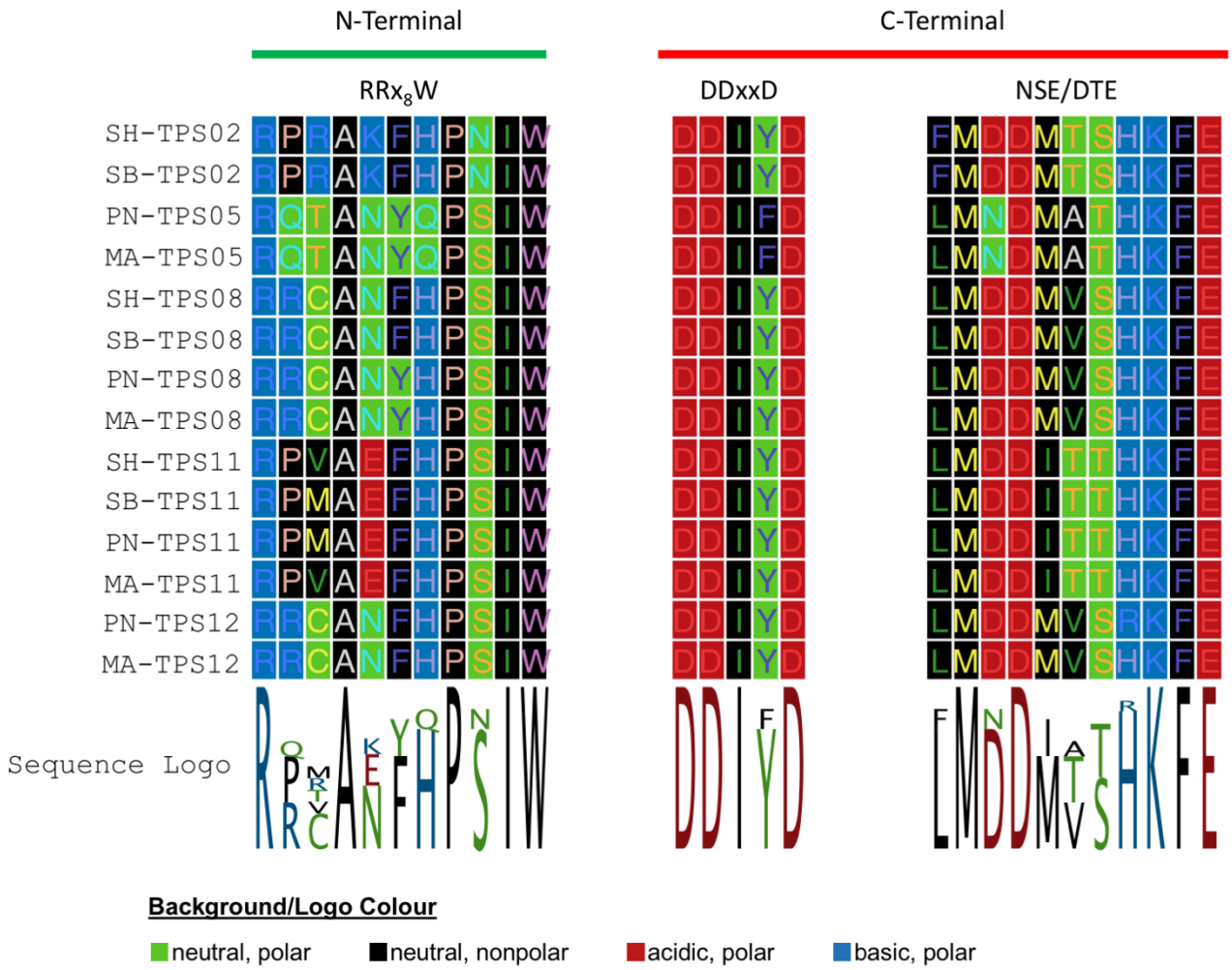


Figure 5.1 – The functional motifs of the N- and C-terminal domains for TPS-encoding genes predicted to result in functional proteins are shown. The polarity and pH at the isoelectric point for the amino acid residue that form these motifs are indicated by the background colours with the respective amino acids (single letter codes) coloured according to the RasMol scheme. The sequence logo reflects the conservation of residues.

5.3.2 Predicted gene structure and sequence analysis

Exon-intron structures of the TPS-encoding genes (Figure 5.2) shows a consensus structure for TPS02, -08, -11 and -12 cultivar variants. Two isoforms were isolated for SB-TPS12, with the second showing partial sequence similarity to the predicted genome sequence but are lacking a large internal part (exons three and four) having incomplete splicing between exons five and six, compared to the consensus structure of the other cultivar variants. This suggests a partial gene duplication for TPS12 in SB. TPS05 shows significant differences between cultivars in terms of sequence length resulting in splice variations. The PN and MA TPS05 gene structures are similar and mapped to the same genome position. Exon one for these cultivars differs at the terminal end with MA being 24 nucleotides longer. SH and SB variants for TPS05 map to the same genomic region but incomplete splicing of the SH variant rendered it non-functional.

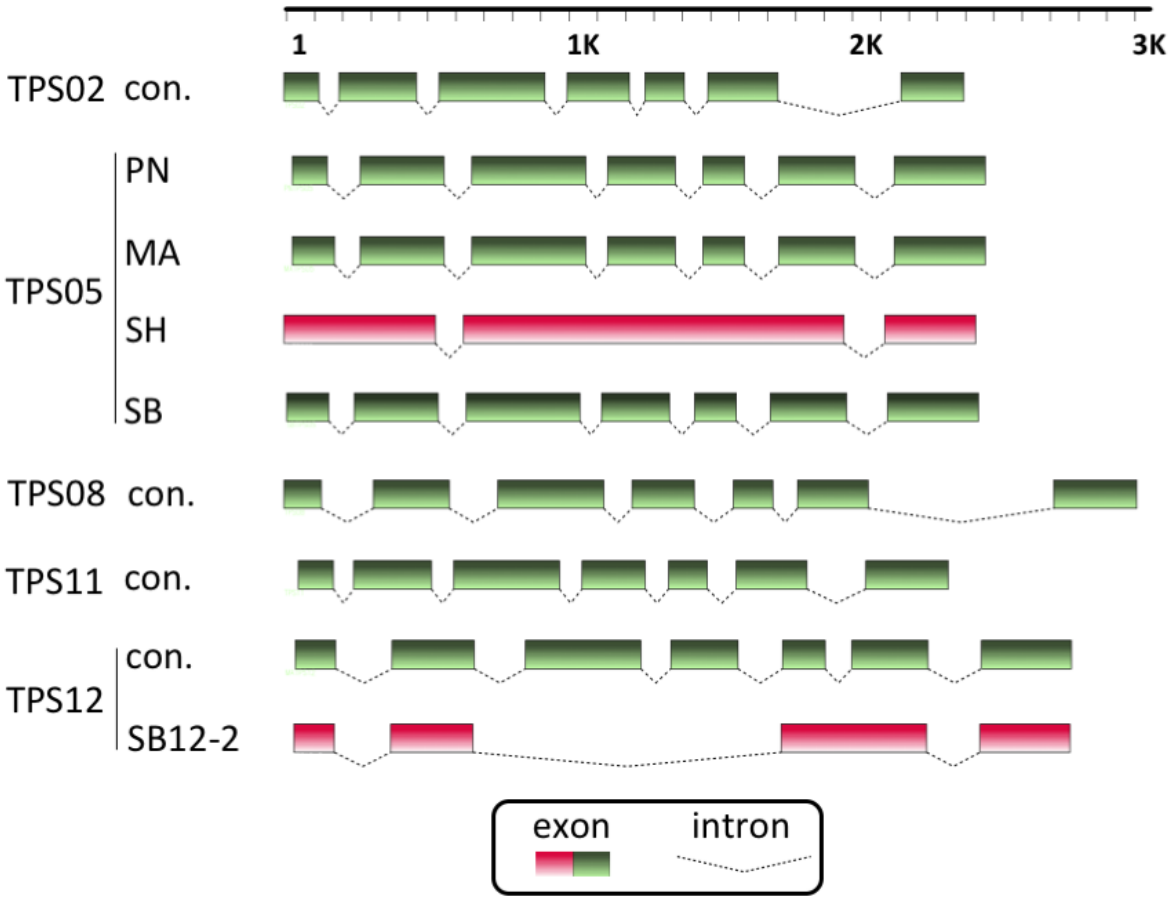


Figure 5.2 – The exon-intron structure of isolated TPS-encoding genes were predicted by alignment to the V1 predicted genomic sequence. Exons in red squares show aberrant splice variants (SH-TPS05 and SB-TPS12-2) with the green exons indicating putatively functional gene structures. The cultivar variants for TPS02, 08, 11 and 12 were identical in terms of gene structure with a consensus structure shown.

A phylogenetic tree comparing the isolated gene sequences to those of the predicted V1 CRIBI protein sequences is shown in Figure 5.3A. As a general trend, the PN and MA TPS variants group together (with the exception of TPS11) while those of SH and SB group, suggesting that these cultivars have similar TPS-encoding genes in flowers. The protein prediction of TPS02 is 72 amino acids longer than the isolated genes, explaining why it branches from the cultivar variants. The same discrepancy was seen for TPS08 where the predicted protein is 47 residues longer than the isolated genes. The T-Coffee percentage identity reflects pairwise similarity to the predicted protein sequence (shown in Figure 5.3B) and enforces the groupings seen in the phylogenetic tree (dendrogram). SH-TPS05 did not group with any of the other TPSs due to the presence of introns resulting in a low sequence identity with the predicted protein or any of the isolated variants.

Sequences that were predicted to encode for a functional protein were aligned and annotated with characteristics domains and motifs with amino acid differences of cultivar variants which are highlighted (Additional File, Figure 7).

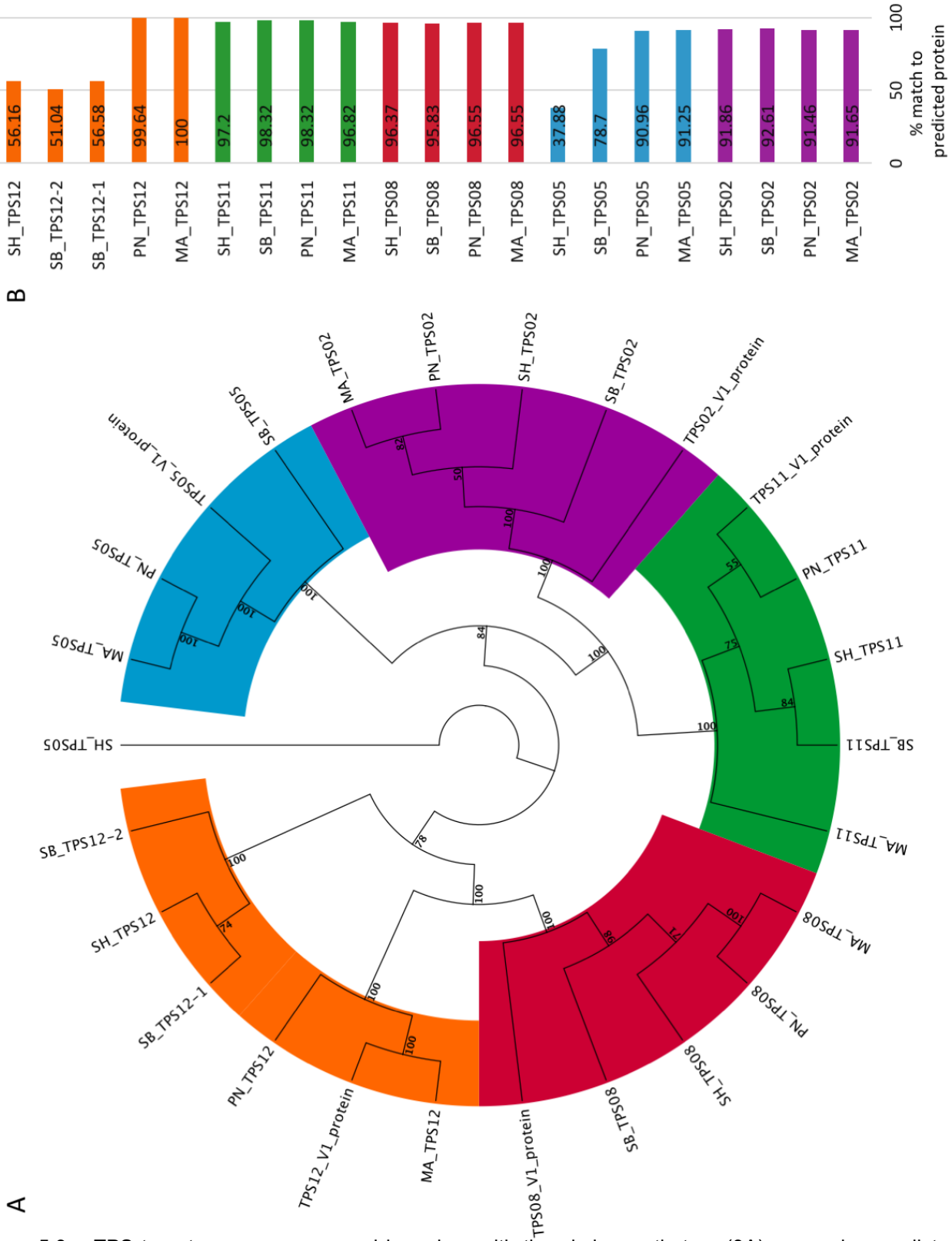


Figure 5.3 – TPS target genes are grouped by colour with the phylogenetic tree (3A) comparing predicted protein sequences (CRIBI V1) to TPS-encoding cultivar variants. The suffix "V1_protein" refers to the predicted protein with the respective isolated cultivar variants indicated be the prefixes. Numbers on the branches indicate the bootstrapping percentage of 100 replicates. The percentage pairwise sequence similarity between predicted proteins and the isolated TPS proteins was determined using the T-Coffee alignment tool and is shown in 3B.

5.3.3 Functional gene expression in a heterologous system

All 14 TPS-encoding genes predicted to be functional (namely, full length ORF in Table 5.2) were analysed for volatile terpene production through heterologous expression in yeast. Volatiles accurately identified and quantified using a reference standard were marked with an asterisk and the major volatile for each TPS indicated in bold text (listed in Table 5.3). Percentage contributions relative to the total terpenoids is illustrated in Figure 5.4.

It was found that TPS02 is functional for SB with β-caryophyllene as the only product. TPS08 was found to be functional in all cultivars except for SB with the major volatiles detected being β-caryophyllene (~65 percent) and β-selinene (~25 percent). TPS05 was functional as a single product enzyme in MA with *E*-β-farnesene as the only product. TPS11 was functional in SB and MA with α-selinene (~40 percent and ~60 percent, respectively) as the major product for both variants. Gamma- (~ 8 percent) and α-muurolene (~4 percent) were present as minor products in SB samples and were not detected for MA. Germacrene D (~55 percent) was the major volatile for MA TPS12 with no terpenic volatiles being detected in PN.

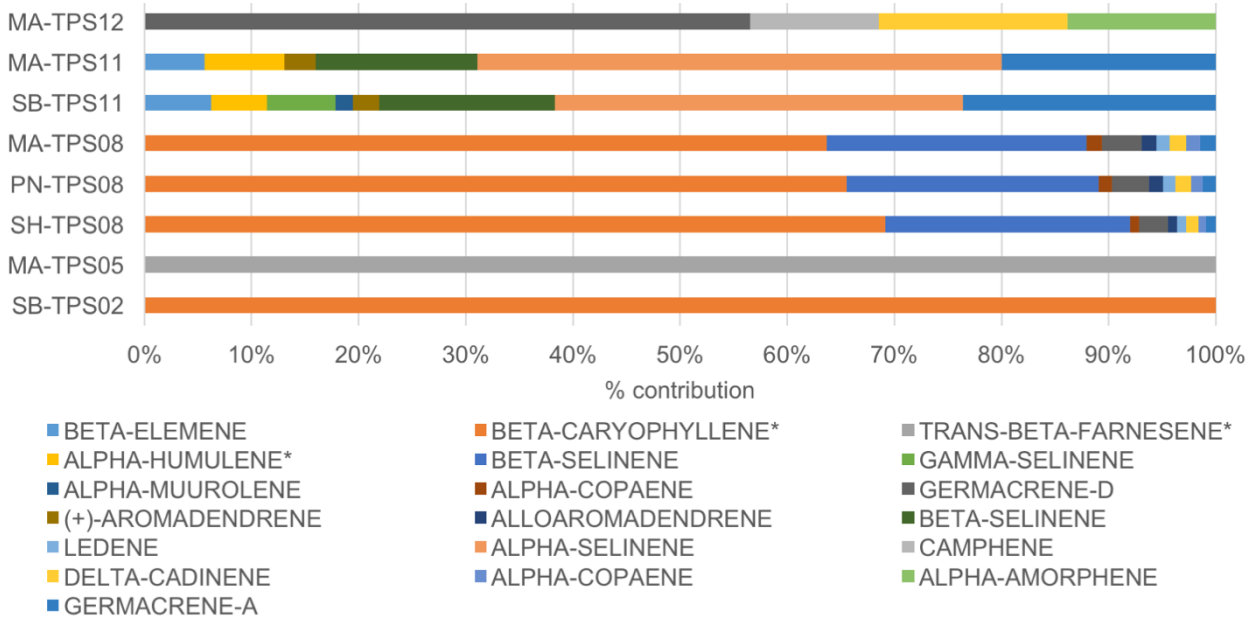


Figure 5.4 – Percentage contribution (ratio) of sesquiterpene volatiles produced through heterologous yeast expression of *VviTPSs* isolated from flowers.

5.3.4 Sequence similarity to characterised TPS genes

All genes shown to be functional in the heterologous *in vivo* yeast expression system were further computationally analysed (Table 5.4). High sequence similarity to known VviTPSs was found for three of the TPS-encoding genes targeted. TPS08 shows enough differences to be classified as a new β-caryophyllene synthase. MA-TPS05 shows no significant similarities to any characterised VviTPS and is likely to be a novel gene.

Table 5.3 – Volatile terpenoid products for TPS-encoding genes isolated from grapevine flowers and expressed in a heterologous yeast system. Data represents an average of one yeast transformant cultured in triplicate. Volatiles accurately identified with an authentic reference standard are indicated by the asterisk. The major volatile for the respective TPSs are shown in bold font.

	TPS02 SB	TPS05 MA	TPS08 SH	TPS08 PN	TPS08 MA	TPS11 SB	TPS11 MA	TPS12 MA
beta-elemene	-	-	-	-	-	0.60 (± 0.04)	0.32 (± 0.04)	-
beta-caryophyllene*	**3.52 (± 1.03)**	-	**51.52 (± 10.51)**	**29.81 (± 2.13)**	**23.99 (± 1.43)**	-	-	-
trans-beta-farnesene*	-	**4.73 (± 0.2)**	-	-	-	-	-	-
alpha-humulene*	-	-	-	-	-	0.51 (± 0.04)	0.43 (± 0.03)	-
beta-selinene	-	-	17.06 (± 2.36)	10.71 (± 0.67)	9.13 (± 0.49)	-	-	-
gamma-selinene	-	-	-	-	-	0.61 (± 0.05)	-	-
alpha-muurolene	-	-	-	-	-	0.16 (± 0.02)	-	-
alpha-copaene	-	-	0.63 (± 0.06)	0.57 (± 0.02)	0.54 (± 0.02)	-	-	-
germacrene-d	-	-	2.00 (± 0.3)	1.57 (± 0.06)	1.40 (± 0.06)	-	-	**2.15 (± 0.04)**
(+)-aromadendrene	-	-	-	-	-	0.24 (± 0.03)	0.17 (± 0.01)	-
alloaromadendrene	-	-	0.66 (± 0.04)	0.57 (± 0.02)	0.53 (± 0.01)	-	-	-
beta-selinene	-	-	-	-	-	1.59 (± 0.1)	0.88 (± 0.03)	-
ledene	-	-	0.64 (± 0.03)	0.55 (± 0)	0.44 (± 0)	-	-	-
alpha-selinene	-	-	-	-	-	**3.68 (± 0.4)**	**2.85 (± 0.07)**	-
camphene	-	-	-	-	-	-	-	0.46 (± 0)
delta-cadinene	-	-	0.83 (± 0.07)	0.64 (± 0.02)	0.60 (± 0.02)	-	-	0.67 (± 0.01)
alpha-copaene	-	-	0.55 (± 0.02)	0.51 (± 0.01)	0.49 (± 0.02)	-	-	-
alpha-amorphene	-	-	-	-	-	-	-	0.53 (± 0)
germacrene-a	-	-	0.64 (± 0.03)	0.54 (± 0.01)	0.53 (± 0.01)	2.28 (± -0.11)	1.16 (± 0.07)	-

Table 5.4 – Functional gene isolates were analysed for similarity to known *VvTPSs* using BLAST. Percentage identity to known *VviTPS* is shown along with the Genbank Acc. and reported major volatile.

Gene Isolate	NCBI ID	Identity	Major Volatile
SB-TPS02	VvGwECar3	541/555 (97.48 %)	β-caryophyllene
MA-TPS05	-	-	
TPS08 variants	AEP17005.1	507/557 (91.02 %)	β-caryophyllene
MA-TPS11	ADR66821.1	540/555 (97.30 %)	Germacrene A
SB-TPS11		548/555 (97.74 %)	
MA-TPS12	NP_001268213.1	548/558 (98.21 %)	Germacrene D

5.4 Discussion

TPS-encoding genes from flowers are conserved between cultivars but show differential functionality in the volatiles produced

Screening and subsequent isolation of TPS-encoding genes from the flowers of four commercial wine cultivars resulted in a complex cultivar-specific functional gene pattern. The four cultivar specific cDNA pools enriched for transcripts present in flowers, resulted in five target genes that were functional in at least one cultivar. Of the 21 putative genes isolated in this study, 14 were predicted to be functional while only eight of the expressed genes lead to the production of volatiles *in vivo*, as summarised in Figure 5.5. MA-TPS05 is a single product enzyme producing *E*-β-farnesene with most other isolated genes encoding for genes that produce multiple sesquiterpenes. Based on the major volatile for the respective TPSs we characterised TPS02 and TPS08 as β-caryophyllene synthases (but with different minor volatiles being formed between the respective cultivars), TPS11 as an α-selinene synthase and TPS12 as a germacene D synthase. In the case of TPS08 and TPS11 we isolated functional cultivar variants but no quantifiable difference in terms of their volatile products was observed. Primers for genes of interest were designed on the Pinot noir PN40024 genome sequence (Jaillon *et al.*, 2007) and, therefore, target genes present in the different cultivars that share homology with the reference genome sequence (namely, conserved across cultivars) and are likely core genes of the *V. vinifera* pan-genome. An attempt was made to target new VviTPS-encoding genes but based on sequence homology it was found that SB-TPS02, cultivar variants of TPS11 and MA-TPS12 were similar to known VviTPS-encoding genes (Table 5.4).

To our knowledge, MA-TPS05 is the first grapevine sesquiterpene synthase that produces *E*-β-farnesene as a single product. TPSs resulting in *E*-β-farnesene are reported for *Zea mays* (Schnee *et al.*, 2002), peppermint (Crock *et al.*, 1997), *Citrus junos* (Maruyama *et al.*, 2001), *Artimesia annua* (Picaud *et al.*, 2005) and maritime pine (Salin *et al.*, 1995). The synthases of *A.*

annua (GenBank Accession No. AY835398) and *C. juno* (GenBank Accession No. AF374462) are the only TPSs reported to produce *E*-β-farnesene as a single product. Comparison of MA-TPS05 to these two sequences show a 47.3 percent and 47.2 percent protein sequence similarity, respectively. These differences emphasise that vastly different TPSs can fulfil a similar function in different species. Furthermore, we found that TPS08 variants have enough sequence differences (~91 percent) with the closest BLAST match to be classified as a new sesquiterpene synthase.

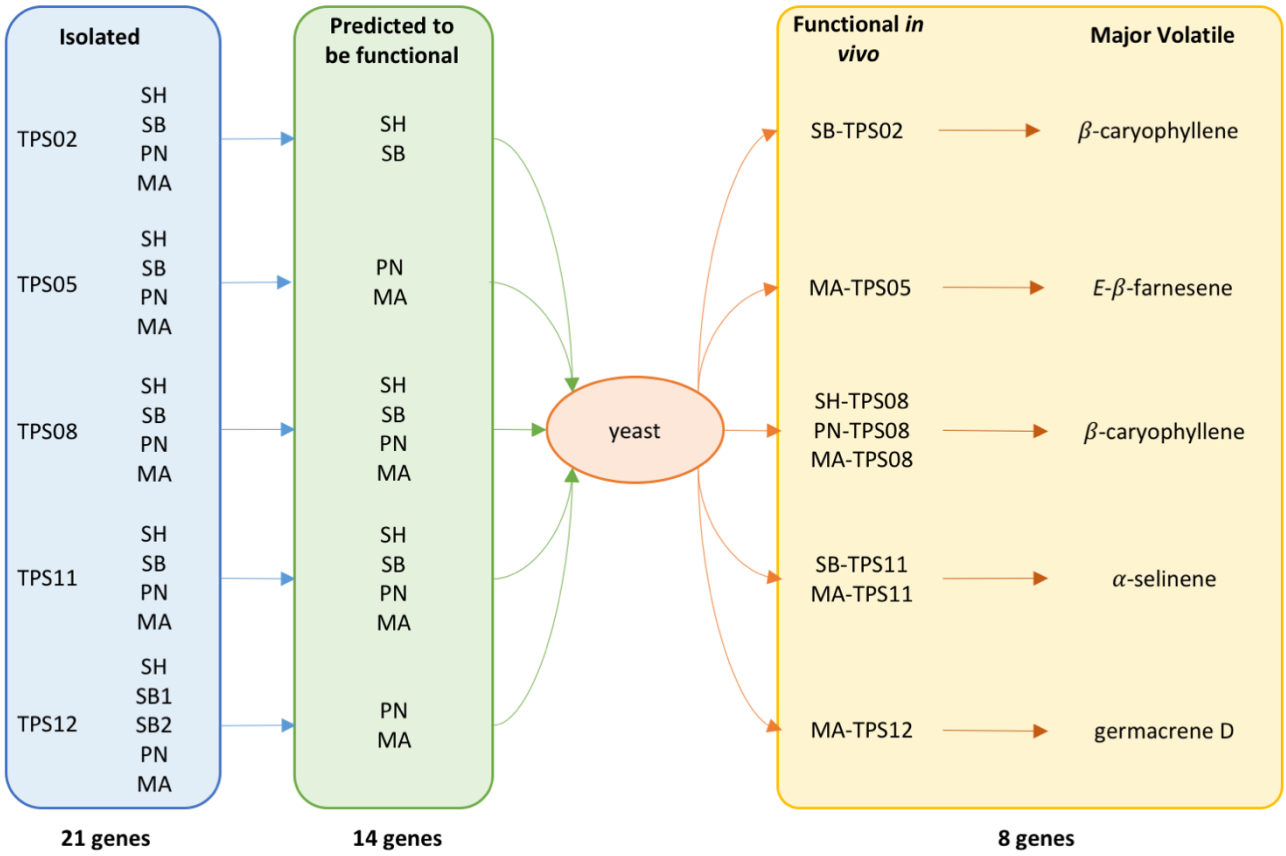


Figure 5.5 – Summary of the isolated cultivar variants, functional predictions and genes that resulted in volatile terpenoids when expressed in a heterologous yeast expression system.

TPS05 presents a unique exon-intron structure for each cultivar variant. The PN variant was predicted to be functional (an uninterrupted coding sequence) but no volatile products were found *in vivo* in the heterologous expression system. We observed SV between the functional MA and non-functional (*in vivo*) PN variant. The 24 nucleotide difference at the terminal end of PN exon one is likely a deletion that altered the N-terminal in such a way that protein folding is non-conducive to facilitating catalytic activity; it is not yet clear whether this is due to alternative splicing or a gene duplication event where part of the gene was lost. The *Vvi*TPS gene family show remnants of transposable elements (Martin *et al.*, 2010) and it is possible that a transposition event resulted in the deletion.

The sequence analysis demonstrates cultivar-specific differences that potentially affect the plant's ability to produce certain types of sesquiterpenes. Amino acid differences due to nucleotide

substitutions resulted in a loss of function due to a premature stop codon for six of the cultivar variants isolated in this study. Preliminary evidence for TPS-encoding genes subject to cultivar-specific splicing was found for SH-TPS05 and SB2-TPS12, resulting in non-functional splice variants. Simpson *et al.* (2008) suggested that plants may undergo more alternative splicing events than previously thought. Alternative splicing plays an important role in producing multiple proteins (isoforms) from a single gene, but when splicing, or genetic mutations, result in a non-functional transcript we see the potential for adverse effects on mRNA stability. These effects can contribute to the accumulation of transcripts that act as substrates for nonsense-mediated decay (NMD), a mechanism in which eukaryotes remove erroneous RNA (RNA that does not encode for a functional protein, ncRNA) (Lewis *et al.*, 2003). Pseudogenes are believed to result in these ncRNAs, with some studies showing that ncRNA acts as a decoy for RNA degrading machinery (namely, NMD), thereby increasing the transcriptional life of functional protein encoding RNA. Pseudogenes are copies of functional genes, made by means of gene duplications or transposition events, and have lost their ability to encode for a functional protein due to mutations or InDels (Muro *et al.*, 2011).

The results presented suggests that evolutionary differences have resulted in cultivar-specific pseudogenes. The aberrant changes of PN- and MA-TPS05 and SB-TPS12-2 suggest that they are pseudogenes that have residual transcriptional activity and potentially act as targets for NMD. This suggests that the highly duplicated VviTPS gene family has a yet to be identified regulatory mechanism by means of transcribed pseudogenes. Evidence of these types of interactions have been found in mammals where pseudogenes help regulate the mRNA stability of their parent genes (Hirotsune *et al.*, 2003). In *Arabidopsis* it has been shown that NMD is involved in transcriptional regulation where mutant plants (*upf1-5*, *upf3-1* and *smg7-1*), with elevated NMD-target transcripts, show increased resistance to pathogens compared to mutants that lack NMD (Rayson *et al.*, 2012). NMD has also recently been implicated in a defence response against plant viral infections (Garcia *et al.*, 2014). Whether TPS pseudogenes can act as targets for NMD is yet to be proven and this potential link to pathogen resistance warrants further investigation.

Amino acid changes seemingly have the greatest influence on TPS functionality. The TPS08 cultivar variants are near identical, with only four amino acid differences being observed (Additional File, Figure 7). All four variants are putatively functional with PN, MA and SH variants producing volatiles *in vivo*. The SB sequence show four amino acid differences with PN and MA and three differences with SH. The Ala_{407} -> Val mutation in the SB protein is the only difference that could explain the lack of terpene products *in vivo*. The position of this substitution suggests that the C-terminal active site scaffold of SB is altered. Amino acid substitutions are thought to contribute to TPS promiscuity with specific residues affecting the active site geometry (that is, how the catalytic site interacts with the prenyl diphosphate substrate), consequently altering both the enzyme activity and volatile profile. This is demonstrated in γ-humulene synthase from *Abies grandis* subjected to

site-directed mutagenesis that target plastic residues (namely, amino acids prone to substitution events), resulting in a systematic change of the enzyme's product profile (Yoshikuni *et al.*, 2006). Greenhagen *et al.* (2006) further demonstrated that subtle amino acid changes, which are beyond the identified active site, can alter terpene synthase functionality. A comparison of TPS10 sesquiterpene synthase orthologs from different maize species demonstrated, through site-directed mutagenesis, that a single amino acid in the catalytic site controlled the ratio of volatiles produced (Köllner *et al.*, 2009). These examples illustrate the importance of the active site to mediate the catalytic interactions with the terpenoid substrates. It appears that a subtle change, like those observed in TPS08 cultivar variants, potentially affects TPS activity. It is thus likely that the greater number of amino acid substitution of TPS02, -05 and 11, had a similar effect on the functionality of their respective cultivar variants. For PN-TPS12, a single amino acid deletion at position 133 likely caused a knock-on effect in terms of the protein folding, mitigating enzyme function.

Although heterologous systems are commonly used to functionally characterise genes, they have limitations. Various studies show that the engineered flux of terpene precursors influence the production capacity of a heterologous yeast (Herrero *et al.*, 2008; Albertsen *et al.*, 2011; Farhi *et al.*, 2011). Poor enzyme kinetics, as a result of amino acid changes combined with an inefficient expression system, could be an alternative reason for the lack of volatiles from genes predicted to be functional. Fischer *et al*. (2013) compares heterologous expression using different plant models (*Agrobacterium* transformation of grapevine calli, *Arabidopsis* floral dip and *Agrobacterium* infiltrations of *Nicotiana benthamiana*), *E. coli* and *S. cerevisiae.* It shows that the cellular background of the heterologous system influences the formation of minor volatiles but that the major volatile is constant between the systems. Translation efficiency of the heterologous genes can be improved through codon optimisation for *S. cerevisiae*, possibly improving volatile production (Lanza *et al.*, 2014). It is, thus, worthwhile to explore alternative expression systems for future studies to provide evidentiary support for the proposed effect of alternative splicing, pseudogenes and/or NMD.

Martinez-Zapater *et al.* (2010) suggest that grapevine could have a SNP frequency of 10 to 16 nucleotides per Kb, which is two to three times higher than that of *Arabidopsis*. This estimation was found to be fairly accurate for some of the gene variants compared in this study. The *Arabidopsis* pan-genome showed that nearly 10 percent of genes could be PAVs when comparing 80 genome accessions to that of the reference genome (Tan *et al.*, 2012). Soybean showed even greater variation in the genome, with up to 20 percent of the genes thought to be variable between wild relatives and cultivated soybean (Li *et al.*, 2014). Soybean also show that SNPs have an impact on 10 percent of the core genome genes from 17 wild and 14 cultivated plants (Lam *et al.*, 2010). Recently, a *de novo* transcriptome for *V. vinifera* L cv. Corvina showed the presence SVs that result in the varietal diversity associated with this cultivar. They identified 646 982 polymorphisms (21 % InDels and 78 % SNPs) compared to the reference genome, with 42 % of

these polymorpisms occurring in coding regions. The study also identified 180 PAVs that were not found in the reference genome with numerous non-annotated genes that potentially have novel functions (Venturini *et al.*, 2013). These results indicate that the grapevine pan-genome could be rather extensive in terms of dispensable gene variants. As more genome sequencing data become available, we will undoubtedly identify SVs contributing to cultivar differences.

5.5 Conclusion

Terpenoid metabolism is a complex process influenced by numerous factors. This discussion presents evidence of cultivar differences in the flower volatilome that could affect the plant's capacity to produce certain volatile sesquiterpenes. SNPs were prevalent in all of the variants, with subtle amino acid changes resulting in non-functional transcripts due to either a premature stop or an aberrant splice site variation effect on the protein. The TPS enzymes analysed appear to be sensitive to amino acid changes in or close to the catalytic site with cultivar variants providing preliminary information on which residues are critical for functionality. These variants can provide valuable insights for future studies on protein folding and amino acid plasticity by using a functional variant as a template to identify residues crucial for functional enzymes. Furthermore, we found candidate genes that are potentially affected by aberrant or alternative splicing which suggests that we might have greater TPS transcriptional diversity than previously thought (compared to the reference genome). The genes targeted in this study show that cultivar-specific evolution of the TPS gene family directly influences the plant's capacity to synthesise terpenoids.

We are, however, still limited by the use of a single genome to study the extent of species diversity. Our view of terpene synthases is currently restricted to those that share sequence similarity with the reference genome and form part of the core genes for grapevine resulting in predominantly SNPs being responsible for the observed differences in functionality. The duplications seen in the TPS-encoding gene family suggest that PAVs could act in a compensatory manner (namely, the role of a deleted/non-functional gene is fulfilled by a related/duplicated gene). The TPS gene family could act in a dosage dependent manner where multiple genes produce the same product. This is likely the case for the sesquiterpene β-caryophyllene. Literature suggests that there are four different genes involved in its production, with this study identifying a fifth gene.

The minor sequence differences observed in this study result in a profound impact on the functionality of the encoded enzymes. Cultivar-specific TPS-encoding variants are important contributors to grapevine terpenoid diversity and each cultivar will have its own repertoire of functional enzymes that ultimately result in its unique volatiles contributing to the varietal typicity of a cultivar.

5.6 References

Albertsen L, Chen Y, Bach LS, Rattleff S, Maury J, Brix S, et al. Diversion of flux toward sesquiterpene production in *Saccharomyces cerevisiae* by fusion of host and heterologous enzymes. Appl Environ Microbiol. 2011;77(3):1033–40.

Bailey TL, Boden M, Buske F a., Frith M, Grant CE, Clementi L, et al. MEME Suite: Tools for motif discovery and searching. Nucleic Acids Res. 2009;37(SUPPL. 2):202–8.

Crock J, Wildung M, Croteau R. Isolation and bacterial expression of a sesquiterpene synthase cDNA clone from peppermint (*Mentha x piperita* , L .) that produces the aphid alarm pheromone (*E*)-β-farnesene. 1997;94:12833–8.

Farhi M, Marhevka E, Masci T, Marcos E, Eyal Y, Ovadis M, et al. Harnessing yeast subcellular compartments for the production of plant terpenoids. Metab Eng. Elsevier; 2011;13(5):474–81.

Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2014;42(D1):D222–30.

Fischer MJC, Meyer S, Claudel P, Perrin M, Ginglinger JF, Gertz C, et al. Specificity of *Ocimum basilicum* geraniol synthase modified by its expression in different heterologous systems. J Biotechnol. 2013;163(1):24–9.

Garcia D, Garcia S, Voinnet O. Nonsense-Mediated Decay Serves as a General Viral Restriction Mechanism in Plants. Cell Host Microbe. 2014;16(3):391–402.

Gietz RD, Woods R a. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 2002;350(2001):87–96.

Golicz AA, Batley J, Edwards D. Towards plant pangenomics. Plant Biotechnol J. 2015;1–7.

Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. Bioinformatics. 2011;27(7):1017–8.

Greenhagen BT, O'Maille PE, Noel JP, Chappell J. Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. Proc Natl Acad Sci U S A. 2006;103(26):9826–31.

Herrero O, Ramón D, Orejas M. Engineering the *Saccharomyces cerevisiae* isoprenoid pathway for de novo production of aromatic monoterpenes in wine. Metab Eng. 2008;10(2):78–86.

Hirotsune S, Yoshida N, Chen A, Garret L, Sugiyama F, Takahashi S, et al. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature. 2003;423:91–6.

Hirsch CN, Buell CR. Tapping the Promise of Genomics in Species with Complex, Non-model Genomes. Annu Rev Plant Biol. 2013

Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449(7161):463–7.

Keller M. Botany and Anatomy. Sci Grapevines Anat Physiol. Elsevier; 2010. p. 1–47.

Köllner TG, Gershenzon J, Degenhardt J. Molecular and biochemical evolution of maize terpene synthase 10, an enzyme of indirect defense. Phytochemistry. 2009;70(9):1139–45.

Lam H-M, Xu X, Liu X, Chen W, Yang G, Wong F-L, et al. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat Genet. 2010;42(12):1053–9.

Lanza AM, Curran KA, Rey LG, Alper HS. A condition-specific codon optimization approach for improved heterologous gene expression in *Saccharomyces cerevisiae*. BMC Syst Biol. 2014;8(1):33.

Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci U S A. 2003;100(1):189–92.

Li Y, Zhou G, Ma J, Jiang W, Jin L, Zhang Z, et al. *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nat Biotechnol. 2014;32(10):1045–52.

Martin DM, Aubourg S, Schouwey MB, Daviet L, Schalk M, Toub O, et al. Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. BMC Plant Biol. 2010;10:226.

Martinez-Zapater JM, Carmona MJ, Diaz-Riquelme J, Fernandez L, Lijavetzky D. Grapevine genetics after the genome sequence: challenges and limitations. Aust J Grape Wine Res. 2010;16:33–46.

Maruyama T, Ito M, Honda G. Molecular cloning, functional expression and characterization of (*E*)-beta farnesene synthase from *Citrus junos*. Biol Pharm Bull. 2001;24(10):1171–5.

Michener CD, Sokal RR. A quantitative approach to a problem in classification. Evolution (NY). 1957;11:130–62.

Muro EM, Mah N, Andrade-Navarro MA. Functional evidence of post-transcriptional regulation by pseudogenes. Biochimie. 2011;93(11):1916–21.

Myles S. Improving fruit and wine: what does genomics have to offer? Trends Genet. 2013 Apr;29(4):190–6.

Pelsy F. Molecular and cellular mechanisms of diversity within grapevine varieties. Heredity (Edinb). 2010;104(4):331–40.

Picaud S, Olofsson L, Brodelius M, Brodelius PE. Expression, purification, and characterization of recombinant amorpha-4,11-diene synthase from *Artemisia annua* L. Arch Biochem Biophys. 2005;436:215–26.

Rayson S, Arciga-Reyes L, Wootton L, De Torres Zabala M, Truman W, Graham N, et al. A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in *Arabidopsis thaliana* NMD mutants. PLoS One. 2012;7(2):e31917.

Salin F, Pauly G, Charon J, Gleizes M. Purification and characterization of trans-β-farnesene synthase from Maritime Pine (*Pinus pinaster* Ait.) needles. J Plant Physiol. 1995;146(3):203–9.

Saxena RK, Edwards D, Varshney RK. Structural variations in plant genomes. Brief Funct Genomics. 2014;13(4):296–307.

Schnee C, Ko TG, Gershenzon J, Degenhardt J. The maize gene terpene synthase 1 encodes a sesquiterpene synthase catalyzing the formation of (*E*)-*β*-farnesene, (*E*)-nerolidol, and (*E,E*)-farnesol after herbivore damage. Plant Physiol. 2002;130:2049–60.

Simpson CG, Lewandowska D, Fuller J, Maronova M, Kalyna M, Davidson D, et al. Alternative splicing in plants. Biochem Soc Trans. 2008;36(Pt 3):508–10.

Tan S, Zhong Y, Hou H, Yang S, Tian D. Variation of presence/absence genes among *Arabidopsis populations*. BMC Evol Biol. 2012;12(1):86.

Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial 'pan-genome'. Proc Natl Acad Sci U S A. 2005;102(39):13950–5.

Thomas BJ, Rothstein R. Elevated recombination rates in transcriptionally active DNA. Cell. 1989 Feb 24;56:619–30.

Velasco R, Zharkikh A, Troggio M, Cartwright D a, Cestaro A, Pruss D, et al. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS One. 2007;(12):e1326.

Venturini L, Ferrarini A, Zenoni S, Tornielli GB, Fasoli M, Dal Santo S, et al. *De novo* transcriptome characterization of *Vitis vinifera* cv. Corvina unveils varietal diversity. BMC Genomics; 2013;14(1):41.

Yoshikuni Y, Ferrin TE, Keasling JD. Designed divergent evolution of enzyme function. Nature. 2006;440(7087):1078–82.

Chapter 6

General discussion and conclusion

Chapter 6 – General discussion and conclusion

Terpenoids are chemically and structurally diverse compounds deployed by plants to increase fitness by allowing for complex adaptations to both biotic and abiotic stresses. The structural diversity and activity of terpenoids is a result of promiscuous enzymes that result in multiple products. Diversity is further amplified by both enzymatic and non-enzymatic modifications that allow for expanded bioactivity (Werck-Reichhart & Feyereisen, 2000; Degenhardt *et al.*, 2009; Bönisch *et al.*, 2014). The numerous ecological functions reported for terpenoids (Tholl, 2006; Gershenzon & Dudareva, 2007) suggest an important *in planta* role for grapevine that has largely been overlooked. Studies have shown cultivar-specific diversity in grape berries that subsequently affect the wine Grape berry terpenoids are well studied and show compartmentalised regulation of *VviTPSs* that have expression patterns correlating with berry ripening (Skinkis *et al.*, 2008; Martin *et al.*, 2012). Most studies show that berries and wine are dominated by monoterpenoids with low levels of sesquiterpenoids reported (Kennedy, 2002; Siebert *et al.*, 2008; Wood *et al.*, 2008; Martin *et al.*, 2012). A few non-berry related studies have indicated that terpenoids are produced in grapevine flowers, although the specific volatilome and their biological functions are still unknown (Martin *et al.*, 2009; Matarese *et al.*, 2014).

The aim of this study was therefore to contribute to our understanding of the composition of the flower volatilome and how/if specific VviTPS-encoding genes can be (functionally) linked to this. Since it is known that grapevine cultivars differ in their volatile aromas, the analysis was conducted on nine selected wine grape cultivars. The main findings of this study will be contextualised with the specific research questions and objectives as outlined in Chapter 1 of the thesis:

6.1 Flower-specific TPS-encoding genes in grapevine revealed a *VviTPS* expression difference between inflorescence development and flowering

In this study, the grapevine TPS-encoding gene family was characterised by means of *in silico* methods using the grapevine gene atlas. A flower developmental expression pattern was shown with sesqui-TPS correlating to the inflorescence stages while mono-TPSs expression was more prevalent in the flowering stage. This suggests a transcriptional switch between inflorescence development and flower bloom. Co-expression analysis showed numerous genes that could influence flower volatile metabolism. Transcription factors (TFs) that were strongly positively correlated (>0.95) with *TPS* expression were identified that could differentially regulate inflorescence or flowering volatile metabolism. A MYB and AUX/IAA TF family members were identified for flowering and inflorescence development, respectively. The reported roles of these TFs, and their high expression levels in the respective developmental stages, suggest that they are involved in regulating the transcriptional switch that coordinates volatile metabolism in flower organogenesis (Reeves *et al.*, 2012). In *Arabidopsis* it was shown that some CYPs have

specificity for terpenoids that result in modifications driving/directing terpenoid bioactivity towards a desired function (Ginglinger *et al.*, 2013). Candidate CYP-encoding genes that potentially fulfil such a role were identified based on co-expression analysis and hold promise for understanding grapevine terpenoid modifications. Enzymes that catalyse glycosylation and glycosidation reactions were identified. These enzymes are potential candidates involved in regulating the conjugation reactions of terpenoids. Conjugation (i.e. glycosylation) modifies terpenoids for increased solubility and altered bioactivity facilitating transport and storage. Glycosidases result in liberation of these conjugates as volatiles (Rivas *et al.*, 2013). These two mechanisms are important regulators in terms of how a plant responds to ecological stresses. Glycosidases are deployed by the plant or by insects during herbivory to release potent volatiles that act as toxins or deterrents against the invading organism (Wink, 2006; Morant *et al.*, 2008; Heil, 2009).

The differential expression patterns of *VviTPSs* during flower development and the co-expressing genes identified presents a first step towards understanding the metabolic networks involved in transcriptional and post-translational regulation of grapevine terpenoid diversity.

6.2 Grapevine flowers from commercial wine cultivars have unique volatile terpenoid profiles

Chemical characterisation of the flower volatilome from different wine cultivars showed a unique composition for the nine cultivars studied. Volatile terpenoid composition was affected by the abundance and/or presence/absence of specific terpenoids in each of the cultivars. The analytical method used in this study enriched for free volatiles and therefore reflected the volatiles that were potentially emitted at flowering and consequently biologically relevant. These emissions were dominated by sesquiterpenoids. It was found that four relatively diverse cultivars (Chenin Blanc, Sauvignon Blanc, Weisser Riesling and Shiraz) produced predominantly valencene (and its rearrangement product 7-epi-α-selinene) as major volatiles. These results were in support of the results reported for Cabernet Sauvignon flowers (Martin *et al.*, 2009). Interestingly, Muscat D'Alexandrie showed a unique volatilome with *E*-β-farnesene as major volatile (60% of the total). This cultivar also expresses a novel cultivar-specific TPS-encoding gene that produced *E*-β-farnesene as single product *in vivo*. Viognier's flower volatilome was also different with β-caryophyllene as the major volatile found (50%). These two cultivars are seen as aromatic (typically floral) varieties (Doligez *et al.*, 2006; Vilanova *et al.*, 2013) with their unique volatilomes suggesting differences in their regulation of terpenoid metabolism. Pinot noir was one of the lowest producers of sesquiterpenes (196 ng/mL versus 539.69 ng/mL in Muscat D'Alexandrie) and showed decreased volatile diversity compared to the other eight cultivars. It suggests that the reference genome (Jaillon *et al.*, 2007), derived from a near homozygous Pinot noir cultivar, under-represents the grapevine *TPS* gene family.

These results suggest that the cultivar differences reported for wine terpenoids (i.e the typicity, of

wine) are even greater in flowers. The absence of monoterpenoids can partially be explained by the *in silico* expression analysis and GC-MS method that was more suited to measure only free volatiles. Co-expression data suggested increased expression of genes regulating glycosylation reactions at flowering. It is thus possible that the mono- and sesquiterpenoids are deployed to fulfil separate functions with the former stored as glycosides to be deployed in defence responses while the latter serves as a volatile infochemcial for insect interactions (Junker *et al.*, 2011). Future studies measuring both free and bound volatiles of flowers would be important to clarify the presence/absence of monoterpenoids and to what extent grapevine produces conjugated terpenes. Methods have been developed to analyse conjugated terpenoids in wine and could be adapted to analyse such terpenoids in other plant organs (Skinkis *et al.*, 2008; Martin *et al.*, 2012).

6.3 Can the grapevine flower volatilome of the different cultivars be linked with specific TPS-encoding genes and enzyme activities?

Functional characterisation, by means of yeast heterologous expression, revealed that TPS function was influenced by cultivar-specific mutations. The Pinot noir grapevine genome showed an over-represented *TPS* gene family with the results in this study presenting evidence that cultivar diversity, with specific reference to *VviTPSs*, were presumably influenced by independent selective pressure(s). Domestication (and speciation) of grapevine likely resulted in the observed gene duplication and insertion/deletion mutations seen in the genome (Myles, 2013). Vegetative propagation often increases the frequency of SNPs, possibly explaining why certain cultivars showed an increased number of SNPs in their TPS-encoding genes (Martinez-Zapater *et al.*, 2010; Myles, 2013). The concept of a pan-genome takes into account these variations and highlights the limitations of using a single reference genome to study within species diversity (Saxena *et al.*, 2014; Golicz *et al.*, 2015). The results presented here provide new insights in terms of cultivar-specific volatiles and the underlying genes that contribute to cultivar diversity. The limited number of cultivars analysed for gene functionality showed profound differences. It is known that terpenoid metabolism is largely transcriptionally controlled in plants (Cheng *et al.*, 2007; Hedhili *et al.*, 2007; Rodríguez *et al.*, 2011) however, we provide evidence that the presence of a putative *TPS* transcript does not equate to a functional enzyme.

We provide new insights on the function of *VviTPS* transcripts. Our results show that between different cultivars the same locus can differ in its TPS functionality, affected by amino acid substitutions and/or frameshift mutations that alter the enzyme active site geometry (e.g. loss of catalytic activity or altered substrate specificity) (Yoshikuni *et al.*, 2006; Garms *et al.*, 2012). In many cases the presence of a premature stop codon resulted in a non-functional gene. The non-protein coding transcripts can be regarded as pseudogenes with their presence suggesting a more complex transcriptional role in terpenoid regulation than previously thought (Zheng & Gerstein, 2007). The grapevine genome shows a large number of *TPS* pseudogenes (63) (Martin *et al.*,

2010) with our study suggesting that their prevalence differs in a cultivar-specific manner, likely due to independent cultivar selection pressures and mutations. Pseudogenes are historically seen as "junk" genes, but studies on human pseudogenes suggest an important regulatory role that increases RNA stability and therefore gene transcription (Muro *et al.*, 2011). It is not yet clear to what extent (or even if) *VviTPS* pseudogenes are involved in transcriptional regulation. We suggest that the mutations that lead to a loss of function (in other words the generation of pseudogenes) are important factors in determining a cultivar's capacity to produce terpenoids. The subtle changes that influence functionality should be considered when conducting expression analysis as not to confuse the presence of *VviTPSs* transcripts with a functional TPS. It is, however, possible that heterologous expression systems (such as the yeast expression system used in this study) give an indication of the TPS functionality, but not necessarily a true reflection of the enzymes function *in planta* (due to for example, compartmentalisation or presence of co-factors). This aspect could be addressed by alternative expression systems that more closely represent the *in planta* conditions of VviTPSs (for example, cell cultures or transient expression in leaves), but each system will have its own advantages and disadvantages (Kapila *et al.*, 1997; Fischer *et al.*, 2013).

6.4 Conclusions and perspectives

The complexities of terpenoid metabolism are highlighted in this study. The presence of a large grapevine *TPS* gene family allows for great diversity in terms of potential metabolites with profound differences seen in the flowers of commercially important wine cultivars. The organ-specific and spatio-temporal regulation of *VviTPSs* were shown through *in silico* methods and suggest ecologically-relevant transcriptional patterns. The transcriptional switch between inflorescence development and flower bloom suggest a specialised function for mono- and sesquiterpenoids during flower organogenesis. By profiling the flower volatilome of different cultivars we provide new insights in terms of terpenoid diversity in grapevine. The development and organ-specific expression of *VviTPSs* were evident in the cultivar-specific volatilomes, with sesquiterpenoids dominating the floral emissions.

Co-expression analysis provided further insight in terms of the metabolic processes involved in terpenoid diversity. Recent advances in the field of CYP enzymatic modifications suggests that terpenoids serve as substrates for oxidation, epoxidation or hydroxylation (Ginglinger *et al.*, 2013). All of these modifications will allow for altered bioactivity and increase the repertoire of terpenoids available for important ecological functions like defence or attraction. The candidate CYP-encoding genes, together with genes encoding for glycosyltransferases and glycosidases, identified in this study will direct further investigations on the metabolic regulation of terpenoid diversity. Limited knowledge is available on the ecological function of grapevine terpenoids; this study provides new insights on terpenoid diversity in flowers that can be used to identify important

infochemcials. The reported roles of farnesene-type sesquiterpenes and β-caryophyllene in mediating insect interactions hold promise for the sesquiterpenoids identified in this study (Tasin *et al.*, 2007, 2011; Anfora *et al.*, 2009). The cultivar differences in the flower volatilomes suggests that valencene and its rearrangements are also important constituents of flower scent and warrant further investigations into the impact of sesquiterpenoid diversity (and cultivar-specific terpenoids) on plant-insect interactions.

In conclusion, this study provided new insights in terms of terpenoid gene regulation and functionality. The complexity of cultivar-specific terpenoids was shown to be influenced by compounding factors that include transcriptional regulation, mutations with aberrant effects on functionality and potential interactions with modifying enzymes. We provide an updated view of grapevine terpenoid metabolism for flower organs with nine unique volatilomes for commercially important wine cultivars. It is evident that the grapevine *TPS* gene family is complex and highly cultivar-specific, directly influencing the volatilome at a transcriptional level. These finding support the necessity for a pan-genomic view of the *TPS* gene family in order to characterise the genetic and metabolic mechanisms that lead to cultivar-specific volatilomes.

6.5 References

Anfora G, Tasin M, de Cristofaro A, Ioriatti C, Lucchi A. Synthetic grape volatiles attract mated *Lobesia botrana* females in laboratory and field bioassays. J Chem Ecol. 2009;35(9):1054–62.

Bönisch F, Frotscher J, Stanitzek S, Rühl E, Wüst M, Bitz O, et al. A UDP-glucose:monoterpenol glucosyltransferase adds to the chemical diversity of the grapevine metabolome. Plant Physiol. 2014;165(2):561–81.

Cheng A-X, Lou Y-G, Mao Y-B, Lu S, Wang L-J, Chen X-Y. Plant terpenoids: biosynthesis and ecological functions. 2007;49(2):179–86.

Degenhardt J, Köllner TG, Gershenzon J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry. 2009;70(15-16):1621–37.

Doligez a., Audiot E, Baumes R, This P. QTLs for muscat flavor and monoterpenic odorant content in grapevine (*Vitis vinifera* L.). Mol Breed. 2006;18(2):109–25.

Fischer MJC, Meyer S, Claudel P, Perrin M, Ginglinger JF, Gertz C, et al. Specificity of *Ocimum basilicum* geraniol synthase modified by its expression in different heterologous systems. J Biotechnol. 2013;163(1):24–9.

Garms S, Chen F, Boland W, Gershenzon J, Köllner TG. A single amino acid determines the site of deprotonation in the active center of sesquiterpene synthases *Sb*TPS1 and *Sb*TPS2 from *Sorghum bicolor*. Phytochemistry. 2012;75:6–13.

Gershenzon J, Dudareva N. The function of terpene natural products in the natural world. Nat Chem Biol. 2007;3(7):408–14.

Ginglinger J-F, Boachon B, Höfer R, Paetz C, Köllner TG, Miesch L, et al. Gene coexpression analysis reveals complex metabolism of the monoterpene alcohol linalool in *Arabidopsis* flowers. Plant Cell. 2013;25(11):4640–57.

Golicz AA, Batley J, Edwards D. Towards plant pangenomics. Plant Biotechnol J. 2015;1–7.

Hedhili S, Courdavault V, Giglioli-Guivarc'h N, Gantet P. Regulation of the terpene moiety biosynthesis of *Catharanthus roseus* terpene indole alkaloids. Phytochem Rev. 2007;6(2-3):341–51.

Heil M. Damaged-self recognition in plant herbivore defence. Trends Plant Sci. 2009;14(7):356–63.

Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449(7161):463–7.

Junker RR, Gershenzon J, Unsicker SB. Floral odor bouquet loses its ant repellent properties after inhibition of terpene biosynthesis. J Chem Ecol. 2011;37(12):1323–31.

Kapila J, De Rycke R, Van Montagu M, Angenon G. An *Agrobacterium*-mediated transient gene expression system for intact leaves. Plant Sci. 1997;122:101–8.

Kennedy J. Understanding grape berry development. Pract Winer Vineyard. 2002;1–5.

Martin DM, Aubourg S, Schouwey MB, Daviet L, Schalk M, Toub O, et al. Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. BMC Plant Biol. 2010;10:226.

Martin DM, Chiang A, Lund ST, Bohlmann J. Biosynthesis of wine aroma: transcript profiles of hydroxymethylbutenyl diphosphate reductase, geranyl diphosphate synthase, and linalool/nerolidol synthase parallel monoterpenol glycoside accumulation in Gewürztraminer grapes. Planta. 2012;236(3):919–29.

Martin DM, Toub O, Chiang A, Lo BC, Ohse S, Lund ST, et al. The bouquet of grapevine (*Vitis vinifera* L. cv. Cabernet Sauvignon) flowers arises from the biosynthesis of sesquiterpene volatiles in pollen grains. Proc Natl Acad Sci U S A. 2009;106(17):7245–50.

Martinez-Zapater JM, Carmona MJ, Diaz-Riquelme J, Fernandez L, Lijavetzky D. Grapevine genetics after the genome sequence: challenges and limitations. Aust J Grape Wine Res. 2010;16:33–46.

Matarese F, Cuzzola A, Scalabrelli G, D'Onofrio C. Expression of terpene synthase genes associated with the formation of volatiles in different organs of *Vitis vinifera*. Phytochemistry. 2014

Morant AV, Jørgensen K, Jørgensen C, Paquette SM, Sánchez-Pérez R, Møller BL, et al. Beta-glucosidases as detonators of plant chemical defense. Phytochemistry. 2008;69(9):1795–813.

Muro EM, Mah N, Andrade-Navarro MA. Functional evidence of post-transcriptional regulation by pseudogenes. Biochimie. 2011;93(11):1916–21.

Myles S. Improving fruit and wine: what does genomics have to offer? Trends Genet. 2013;29(4):190–6.

Reeves PH, Ellis CM, Ploense SE, Wu MF, Yadav V, Tholl D, et al. A regulatory network for coordinated flower maturation. PLoS Genet. 2012;8(2):e1002506.

Rivas F, Parra A, Martinez A, Garcia-Granados A. Enzymatic glycosylation of terpenoids. Phytochem Rev. 2013;12:327–39.

Rodríguez A, San Andrés V, Cervera M, Redondo A, Alquézar B, Shimada T, et al. Terpene down-regulation in orange reveals the role of fruit aromas in mediating interactions with insect herbivores and pathogens. Plant Physiol. 2011;156(2):793–802.

Saxena RK, Edwards D, Varshney RK. Structural variations in plant genomes. Brief Funct Genomics. 2014;13(4):296–307.

Siebert TE, Wood C, Elsey GM, Pollnitz AP. Determination of rotundone, the pepper aroma impact compound, in grapes and wine. J Agric Food Chem. 2008;56(10):3745–8.

Skinkis PA, Bordelon BP, Wood K V. Comparison of monoterpene constituents in Traminette, Gewürztraminer, and Riesling winegrapes. Am J Enol Vitic. 2008;59(4):440–5.

Tasin M, Bäckman A-C, Coracini M, Casado D, Ioriatti C, Witzgall P. Synergism and redundancy in a plant volatile blend attracting grapevine moth females. Phytochemistry. 2007;68(2):203–9.

Tasin M, Betta E, Carlin S, Gasperi F, Mattivi F, Pertot I. Volatiles that encode host-plant quality in the grapevine moth. Phytochemistry. 2011;72(16):1999–2005.

Tholl D. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Curr Opin Plant Biol. 2006;9:297–304.

Vilanova M, Genisheva Z, Graña M, Oliveira JM. Determination of odorants in varietal wines from international grape cultivars (*Vitis vinifera*) grown in NW Spain. South African J Enol Vitic. 2013;34(2):212–22.

Werck-Reichhart D, Feyereisen R. Cytochromes P450: a success story. Genome Biol. 2000;1(6):REVIEWS3003.

Wink M. Importance of plant secondary metabolites for protection against insects and microbial infections. In: Rai M, Carpinella MC, editors. Nat Occur Bioact Compd. 2006. p. 251–68.

Wood C, Siebert TE, Parker M, Capone DL, Elsey GM, Pollnitz AP, et al. From wine to pepper: Rotundone, an obscure sesquiterpene, is a potent spicy aroma compound. J Agric Food Chem. 2008;56(10):3738–44.

Yoshikuni Y, Ferrin TE, Keasling JD. Designed divergent evolution of enzyme function. Nature. 2006;440(7087):1078–82.

Zheng D, Gerstein MB. The ambiguous boundary between genes and pseudogenes: the dead rise up, or do they? Trends Genet. 2007;23(5):219–24.

Additional File

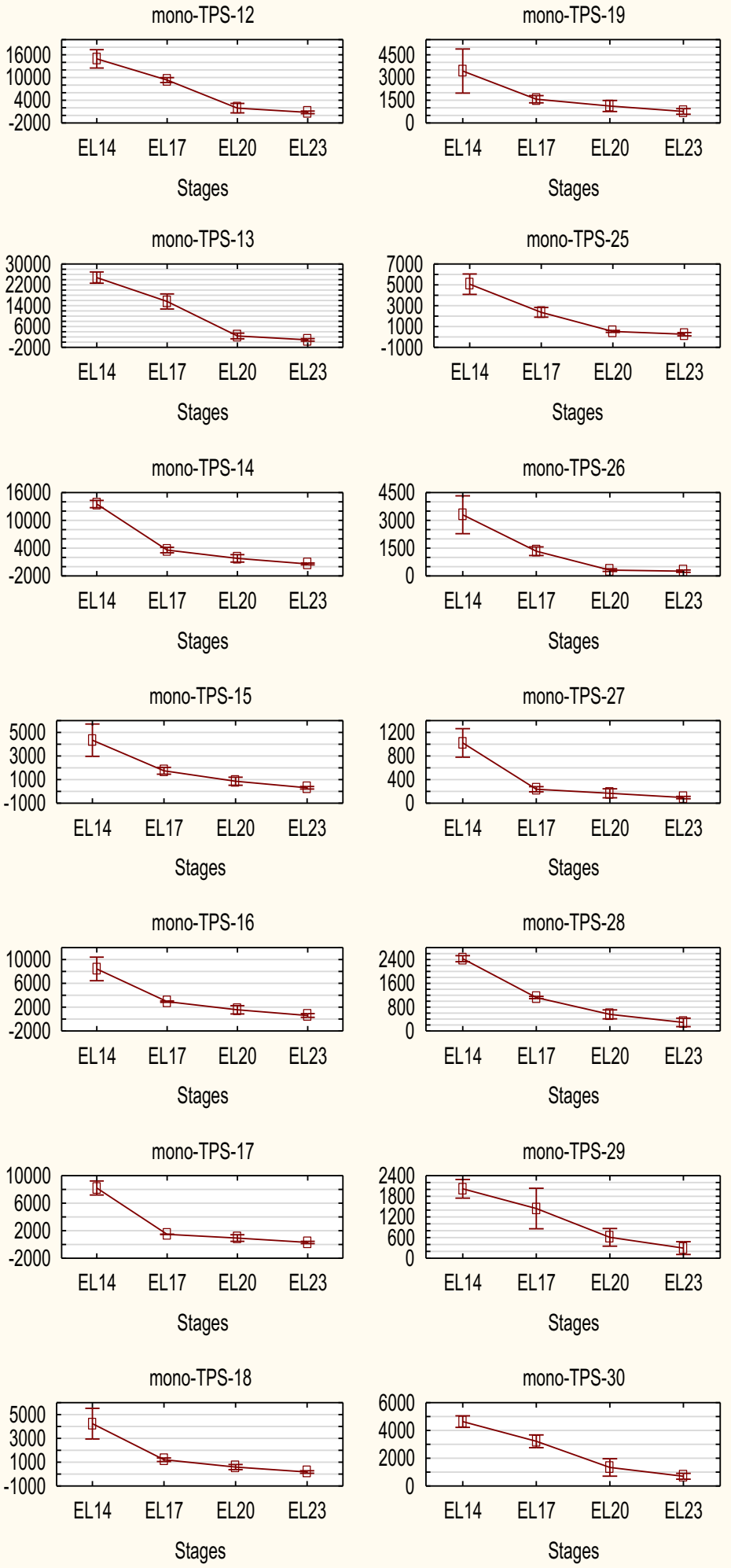


Figure 1 – Gene expression patterns of TPSs that show increased expression in the inflorescence stages (E-L 14 and 17)

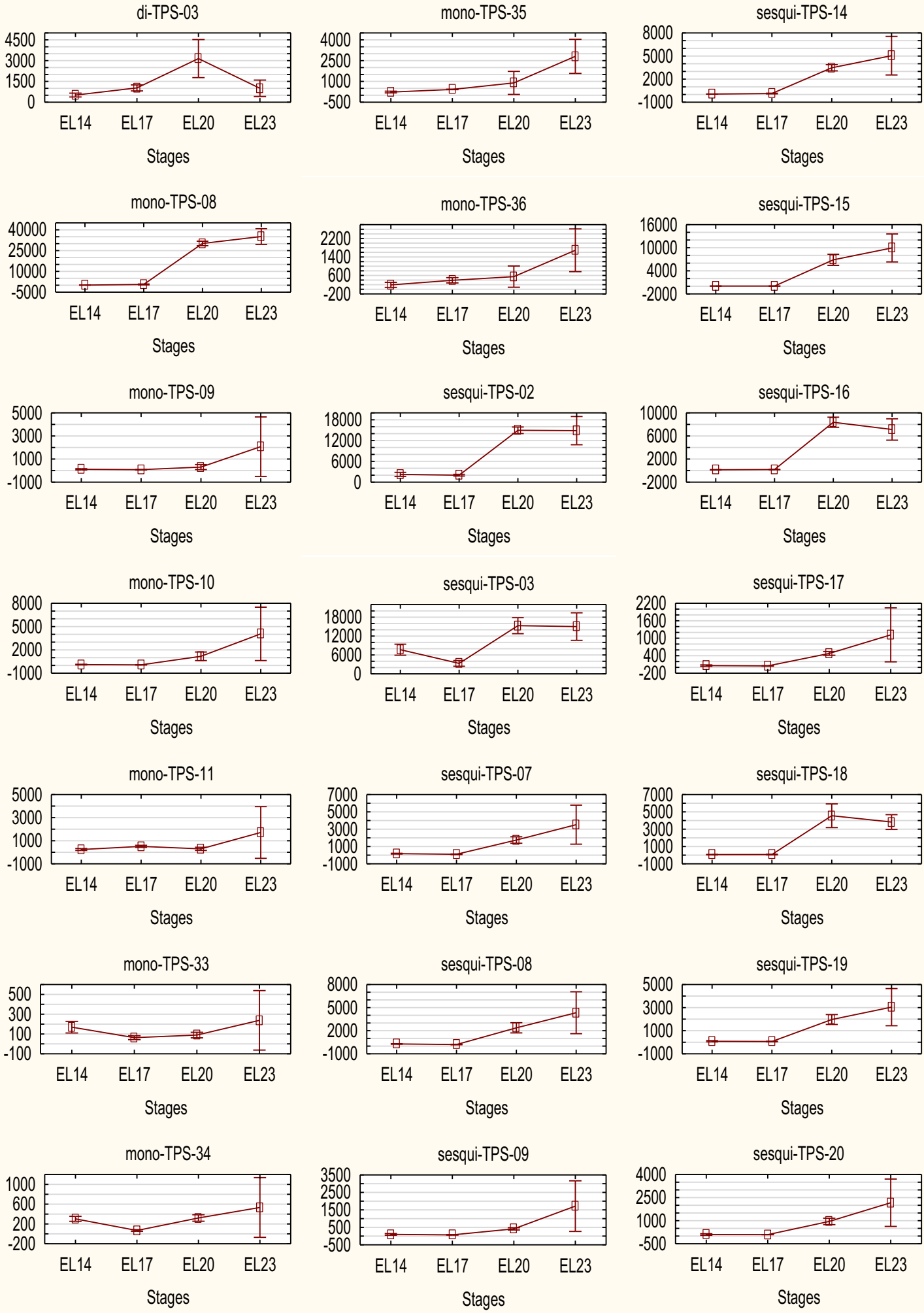


Figure 2 – Gene expression patterns of TPSs that show increased expression in the flowering stages (E-L 20 and 23)

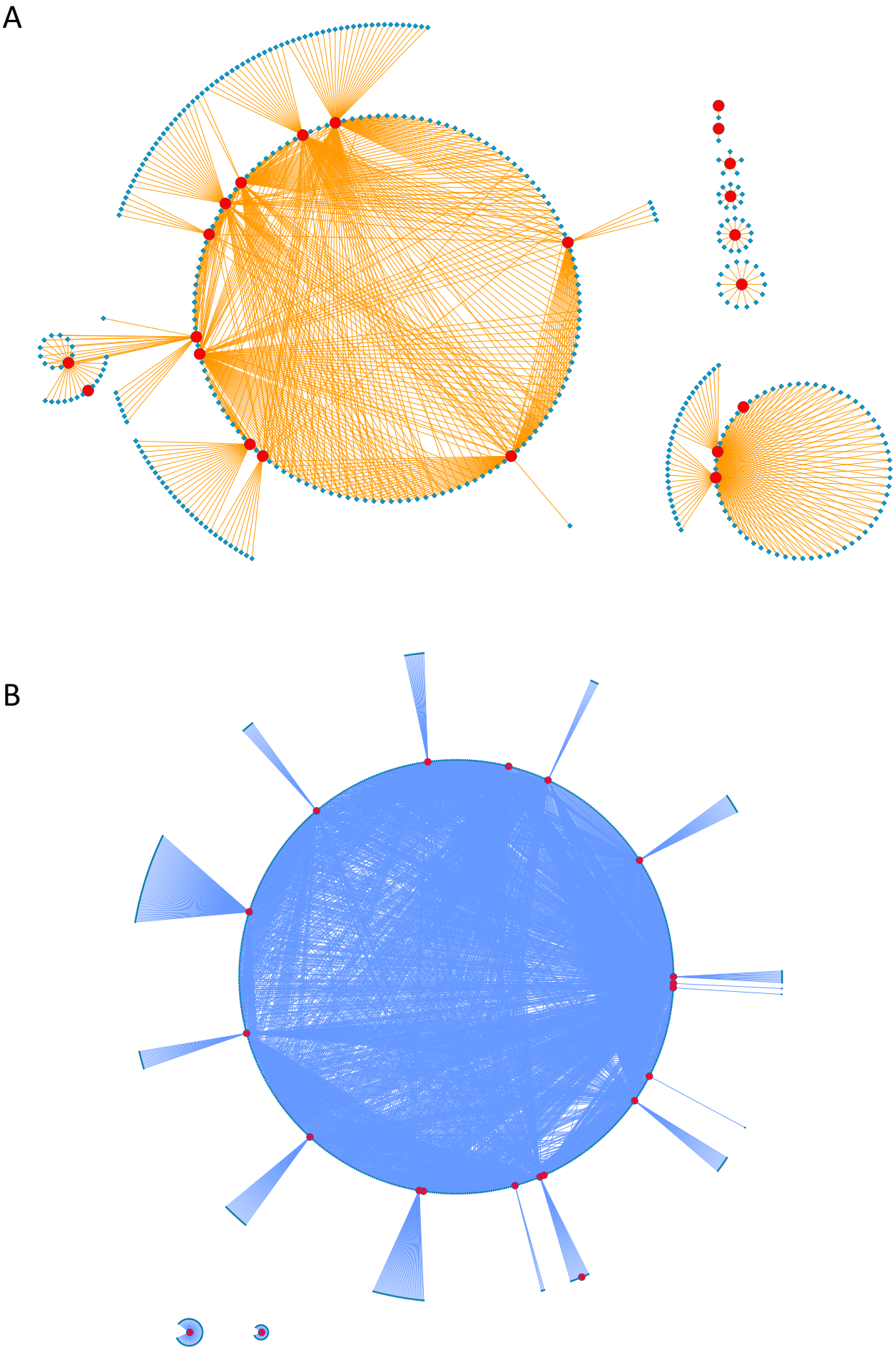


Figure 3 – GGCN-4 (**A**) and GGCN-5 (**B**) represent co-expression network enrichments for GEC-4 and -5 respectively, therefore flowering and inflorescence specific enrichments. Gene baits obtained from GEC-4 and -5 are shown as red circles. These networks served as parent networks for the construction of GGCN-6

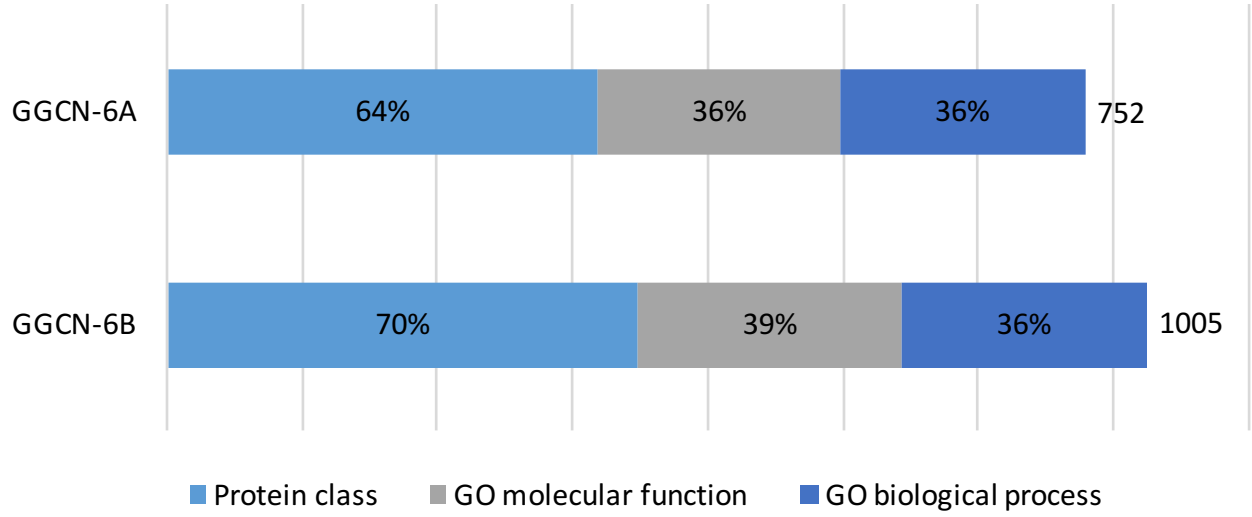


Figure 4 – Percentage of genes relative to the number of nodes per network that showed no classification for the three different analyses

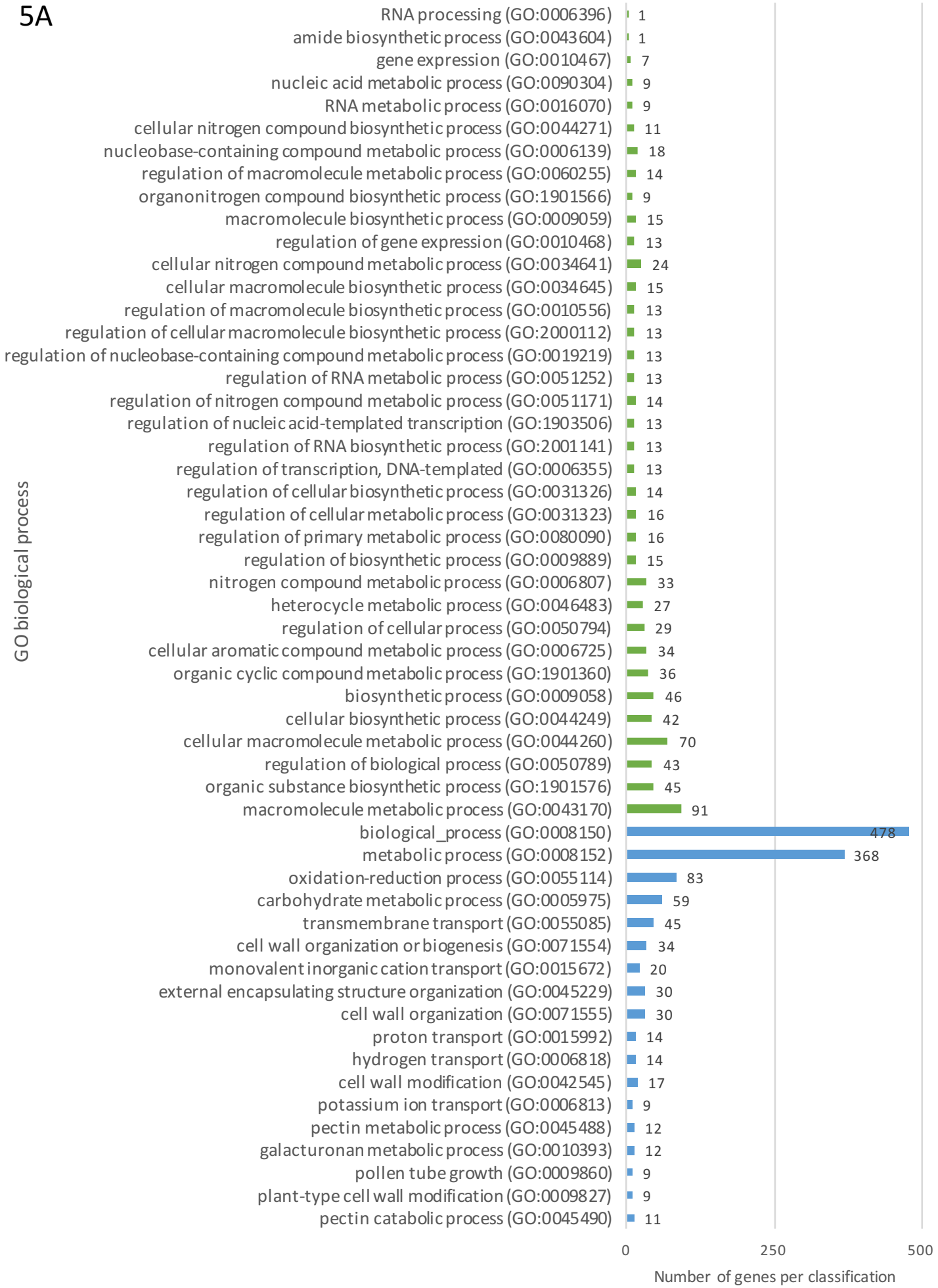
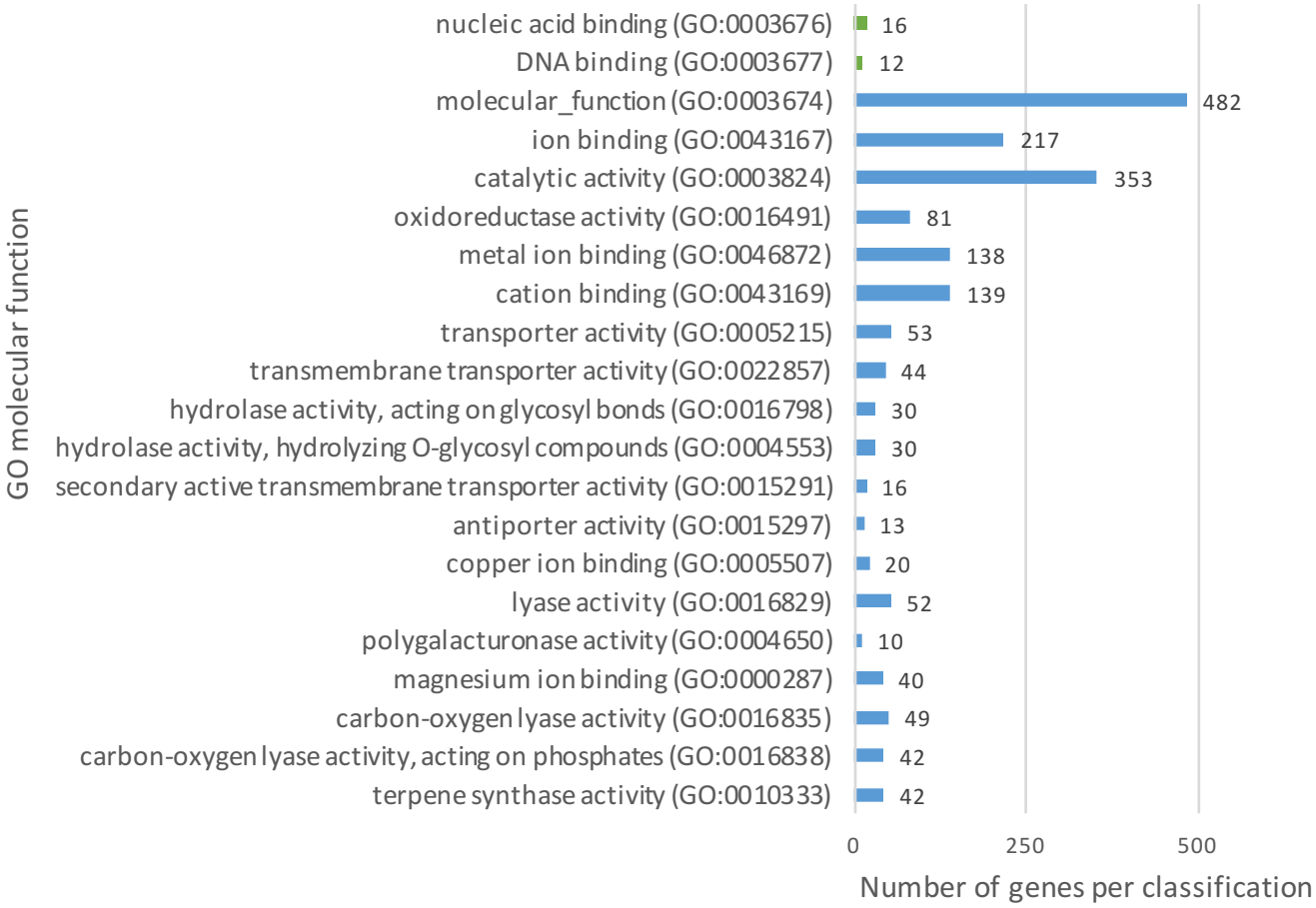
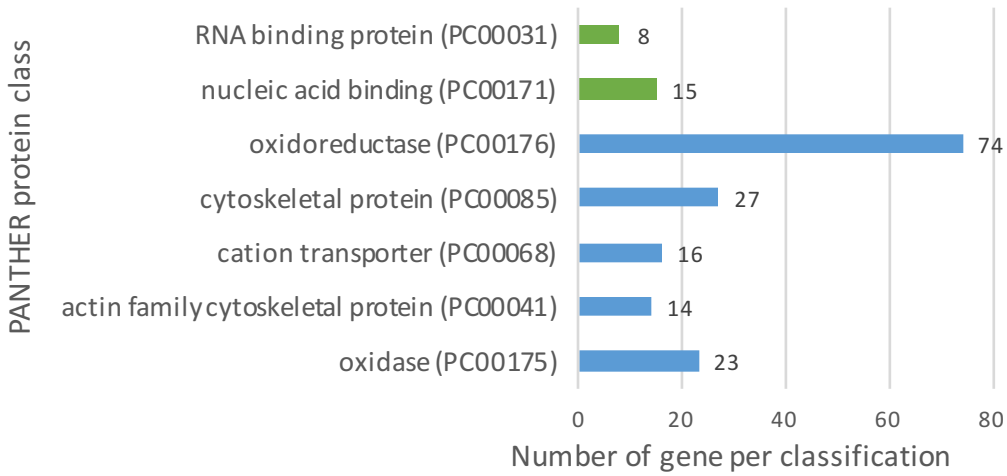


Figure 5 – Gene Ontology (GO) classifications of co-expressing genes enriched for flowering and flower organs (GGCN-6A). Classifications relating to GO biological processes (**A**) and molecular functions (**B**) are shown as well as the protein classifications according to the PANTHER database (**C**).

5B


GO molecular function
nucleic acid binding (GO:0003676) 16
DNA binding (GO:0003677) 12
molecular_function (GO:0003674) 482
ion binding (GO:0043167) 217
catalytic activity (GO:0003824) 353
oxidoreductase activity (GO:0016491) 81
metal ion binding (GO:0046872) 138
cation binding (GO:0043169) 139
transporter activity (GO:0005215) 53
transmembrane transporter activity (GO:0022857) 44
hydrolase activity, acting on glycosyl bonds (GO:0016798) 30
hydrolase activity, hydrolyzing O-glycosyl compounds (GO:0004553) 30
secondary active transmembrane transporter activity (GO:0015291) 16
antiporter activity (GO:0015297) 13
copper ion binding (GO:0005507) 20
lyase activity (GO:0016829) 52
polygalacturonase activity (GO:0004650) 10
magnesium ion binding (GO:0000287) 40
carbon-oxygen lyase activity (GO:0016835) 49
carbon-oxygen lyase activity, acting on phosphates (GO:0016838) 42
terpene synthase activity (GO:0010333) 42
0
250
500
Number of genes per classification

5C


PANTHER protein class
RNA binding protein (PC00031) 8
nucleic acid binding (PC00171) 15
oxidoreductase (PC00176) 74
cytoskeletal protein (PC00085) 27
cation transporter (PC00068) 16
actin family cytoskeletal protein (PC00041) 14
oxidase (PC00175) 23
0
20
40
60
80
Number of gene per classification

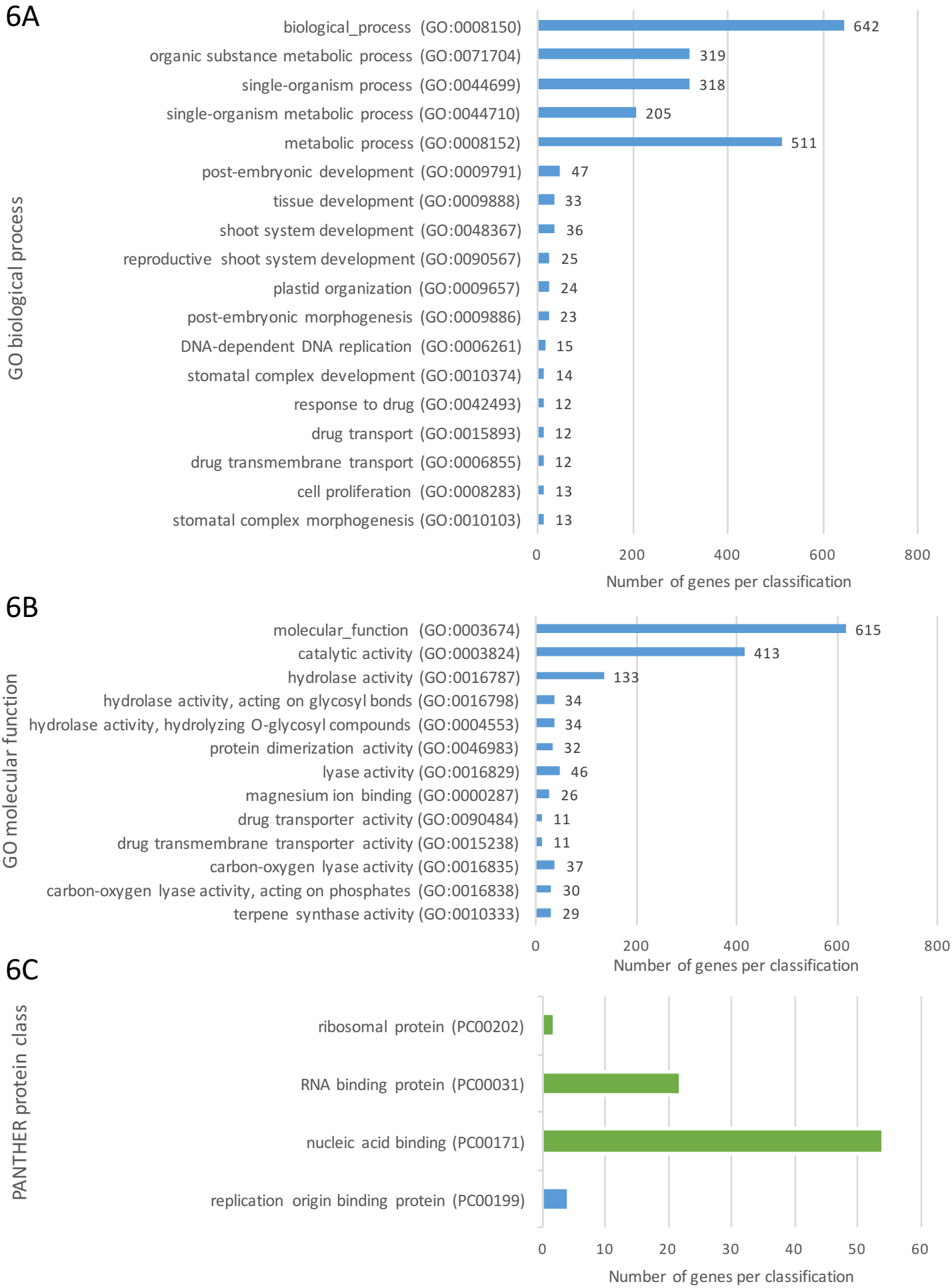
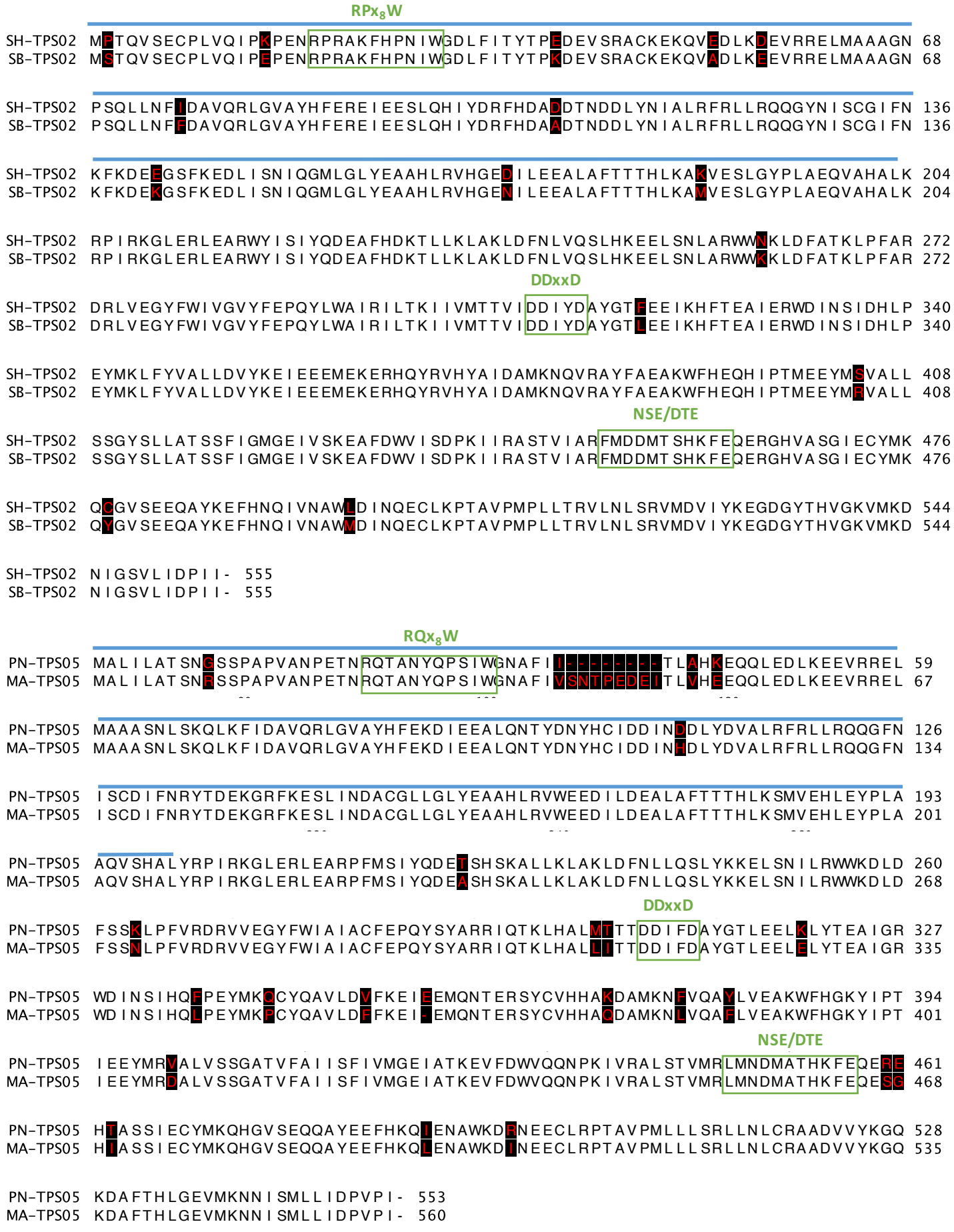
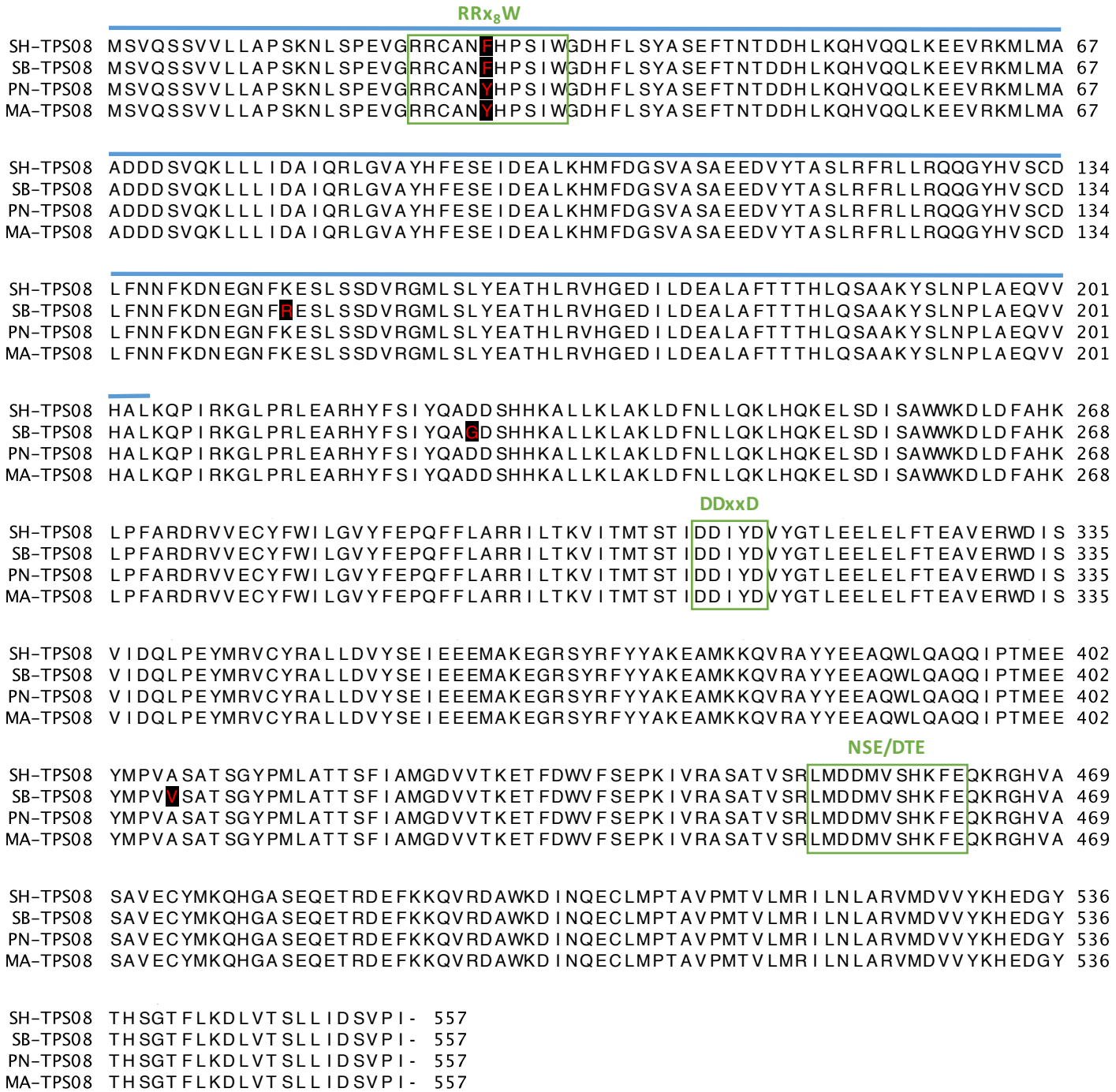
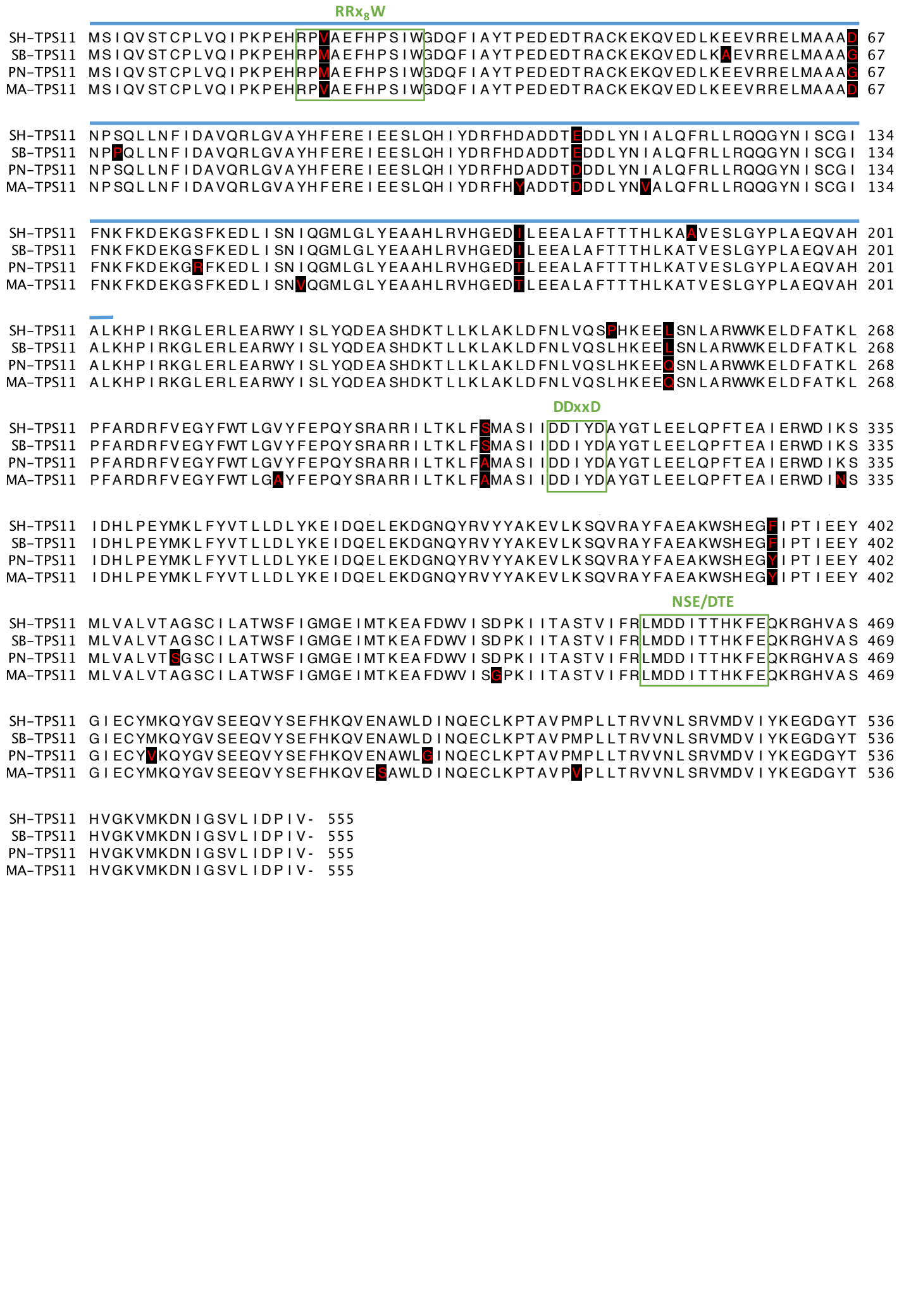


Figure 6 – Gene Ontology (GO) classifications of co-expressing genes enriched for inflorescence development (GGCN-6B). Classifications relating to GO biological processes (**A**) and molecular functions (**B**) are shown as well as the protein classifications according to the PANTHER database (**C**).

Figure 7 – Multiple sequence alignments of *VviTPSs* isolated from grapevine flowers that were predicted to be functional. Position of TPS motifs are indicated by green squares with the N-terminal region marked by a blue line. Amino acid differences are highlighted. The specific TPS target gene is shown on the left of each alignment along with the cultivar the gene was isolated from. Cultivars were abbreviated as follow: Shiraz (SH), Sauvignon Blanc (SB), Pinot noir (PN) and Muscat D'Alexandrie (MA). Note that the sequence alignments cover multiple pages.




RRx8W
DDxxD
NSE/DTE


RRx8W
DDxxD
NSE/DTE


RRx8W
PN-TPS12 MSVQSSVVLLAPSKNLSPEVGRRCANFHPSIWGDHFLSYASEFTNTDDHLKQHVQQLKEEVRKMLMA 67
MA-TPS12 MSVQSSVVLLAPSKNLSPEVGRRCANFHPSIWGDHFLSYASEFTNTDDHLKQHVQQLKEEVRKMLMA 67
PN-TPS12 ADDDSAQKLLLIDAIQRLGVAYHFESEIDEVLKHMFDGSVVSAEEDVYTASLRFRLLRQQGYHVSC- 133
MA-TPS12 ADDDSAQKLLLIDAIQRLGVAYHFESEIDEVLKHMFDGSVVSAEEDVYTASLRFRLLRQQGYHVSCA 134
PN-TPS12 DLFNNFKDNEGNFKESLSSDVRGMLSLYEATHFRVHGEDILDEALAFTTTHLQSATKHSSNPLAEQV 200
MA-TPS12 DLFNNFKDNEGNFKESLSSDVRGMLSLYEATHFRVHGEDILDEALAFTTTHLQSATKHSSNPLAEQV 201
PN-TPS12 VHALKQPIRKGLPRLEARHYFSVYQADDSHNKALLKLAKLDFNLLQKLHQKELSDISAWWKDLDFAH 267
MA-TPS12 VHALKQPIRKGLPRLEARHYFSVYQADDSHNKALLKLAKLDFNLLQKLHQKELSDISAWWKDLDFAH 268
DDxxD
PN-TPS12 KLPFARDRVVECYFWILGVYFEPQFFFARRILTKVIAMTSIIDDIYDVYGTLEELELFTEAVERWDI 334
MA-TPS12 KLPFARDRVVECYFWILGVYFEPQFFFARRILTKVIAMTSIIDDIYDVYGTLEELELFTEAVERWDI 335
PN-TPS12 SAIDQLPEYMRVCYQALLYVYSEIEEEMAKEGRSYRLYYAKEAMKNQVRAYYEEAKWLQVQQIPTME 401
MA-TPS12 SAIDQLPEYMRVCYQALLYVYSEIEEEMAKEGRSYRLYYAKEAMKNQVRAYYEEAKWLQVQQIPTME 402
NSE/DTE
PN-TPS12 EYMPVALVTSAYSMLATTSFVGMGDAVTKESFDWIFSKPKIVRASAIVCRLMDDMVSRKFEQKRGHV 468
MA-TPS12 EYMPVALVTSAYSMLATTSFVGMGDAVTKESFDWIFSKPKIVRASAIVCRLMDDMVSHKFEQKRGHV 469
PN-TPS12 ASAVECYMKQHGASEQETHNEFHKQVRDAWKDINEECLIPTAVPMPILMRVLNLARVIDVIYKNEDG 535
MA-TPS12 ASAVECYMKQHGASEQETHNEFHKQVRDAWKDINEECLIPTAVPMPILMRVLNLARVIDVIYKNEDG 536
PN-TPS12 YTHSGTVLKDFVTSMLSDPVPI- 557
MA-TPS12 YTHSGTVLKDFVTSMLIDPVPI- 558