UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI AGRONOMIA ANIMALI ALIMENTI RISORSE NATURALI ED AMBIENTE

CORSO DI DOTTORATO DI RICERCA IN CROP SCIENCE

CICLO XXXV

THE REGULATION OF GENE EXPRESSION IN GRAPEVINE FLOWERING PROCESS

Coordinatore: Ch.mo Prof. Claudio Bonghi

Supervisore: Ch.ma Prof. Margherita Lucchin

Co-Supervisore: Ch.mo Prof. Alessandro Vannozzi

Dottorando: Gabriele Magon

ANNO ACCADEMICO 2022 - 2023

ABSTRACT

Climate change is rapidly altering the delicate balance between organisms, environment and ecosystem. Global warming, ecosystems destruction and catastrophic extreme events have been constantly increasing because of the uncontrolled anthropic activity and nowadays agriculture represents one of the sectors majorly affected by these phenomena. It is now widely accepted that the next challenge for agriculture will be the implementation of productive systems and genotypes characterized by low inputs, maximization of process efficiency and high resilience.

The present project was aimed at shading light on the molecular mechanisms governing the flowering process in grapevine (*Vitis vinifera* L.). Although this crop is generally propagated asexually, there are many reasons justifying an interest the flowering process in this species. In contrast to the historical and cultural legacies that see viticulture as an extremely conservative discipline, the importance of conventional breeding has been increasingly affirming with several breeding programmes aimed at the constitutions of grapevine varieties characterized by resistance to biotic stress, resilience to climate changes and improved quality traits. In this sense, the understanding of the mechanisms related to floral development and the ontogenetic determination of the main players in the processes of sexual reproduction in plants, pollen, and egg cells, is a step of crucial importance. Moreover, quantitative, and qualitative traits of the production are directly influenced by flowering, which represents the first constituted step of the reproductive phase leading to the development of the main product, the berry. The flower structure and the inflorescence architecture affect the organization of the bunch, with many consequences at the production level. To improve this kind of knowledge we developed a grapevine floral expression atlas in different organs and different stages of *V. vinifera* cv Pinot noir. The calyx, calyptra, filament, anther, stigma, ovary, and embryo of P. noir flowers were dissected and collected in two time points representing pre- and post-anthesis phases and their transcriptome was analysed by means of RNA-Seq approach, in order to identify hub genes that unequivocally distinguish the different tissues. Several analytical techniques such as the weighted gene co-expression network analysis (WGCNA) and tau analysis were used to isolate clusters of genes showing similar expression pattern and high/absolute specificity for each whorl considered in the study. Amongst whorls-specific genes, several transcription factors (TFs) were identified and 17 of them were selected to be analysed by a novel technique never used before in grapevine, the DNA Affinity Purification sequencing (DAP-Seq). This assay allows the identification of the binding sites within a genome for a given TF through the hybridization of

the *in vitro* expressed TF-protein with the fragments of gDNA. After identifying the entire cistrome of the 17 TFs, we were able to define each TF gene targets based on their proximity to TF-binding regions. These data were crossed with those obtained by the flower atlas, WGCNA and tau analyses, providing a list of genes specifically expressed in a given whorl and controlled by a given TF. These genes were defined as high confidence targets (HCTs) and represent a sort of TF- and tissue-specific regulative network, constituted of genetic determinants strictly interconnected and responsible for the peculiar developmental processes of the related floral organ. Subsequently, we focus our efforts on the anther-specific TF among *VviMYB108A*, and on its paralogue *VviMYB108B.* These genes were selected for functional characterization in order to ascertain their putative involvement in male fertility and *Botrytis cinerea* susceptibility. It was investigated the subcellular localization through transient overexpression in *Nicotiana benthamiana*. Both *VviMYB108A* and *VviMYB108B* were used to produce tomato (var. "Micro Tom") and *Arabidopsis thaliana* stable overexpressing lines and are currently under analysis. Finally, two *Arabidopsis* mutant lines were complemented.

Summary

Structure of the thesis

The main purpose of this study was to investigate the role of some transcriptional regulators belonging to R2R3-MYB, WRKY and NAC families putatively involved in the identity, development and functioning of the main floral organs in grapevine (*Vitis vinifera* L.), with the final aim to contribute to the description of the main regulative networks behind the flowering mechanism.

Chapter I provides a general introduction to the entire work, giving a framework in which this work fits, and describing the background knowledge on the flowering process in grapevine, its transcriptional regulation and the main TFs families involved.

Chapter II reports a paper published during the Ph.D. By means of RNA-Seq approach, a grapevine transcriptomic atlas was produced, considering a number of different floral whorl at two different time-points representing the pre- and post-anthesis phases. Several bioinformatic analysis were performed to identify groups of genes showing the same expression profile and specificity for a given floral organ. Among them, several transcription factors were identified.

Chapter III reports the optimization of a DAP-Seq protocol for grapevine and the description of the whole cistrome for 17 transcription factors specifically expressed in different floral whorls and putatively involved in their functionality. The results were thus crossed with the data retrieved with WGCNA and tau analysis described in Chapter II in order to obtain tissue-specific TF-centred regulative network maps. Results were then integrated with a *de novo* motif discovery analysis.

Chapter IV reports a second paper published under the supervision of Dr Josè Tomàs Matus (I2SYSBIO – CSIC Universitat de Valencia) during the Ph.D. that illustrates the application of DAP-Seq to discover the cistrome of the two TFs, namely MYB14 and MYB15. DAP-Seq results were combined with aggregate gene co-expression networks (GCNs) built from more than 1400 transcriptomic datasets from leaves, fruits, and flowers to narrow down bound genes to a set of high confidence targets. The analysis showed that in addition to the few previously known stilbene synthase (STS) targets, these regulators bind to 30 of 47 STS family genes. Overall, Subgroup 2 R2R3-MYBs appear to play a key role in binding and directly regulating several primary and secondary metabolic steps leading to an increased flux towards stilbenoid production.

Chapter V. Functional characterization of *VviMYB108A* and *VviMYB108B*. For both, it was investigated the subcellular localization through the heterologous transient overexpression in

Nicotiana benthamiana leaves. Moreover, stable overexpressing lines of tomato (var. “Micro Tom”) and *Arabidopsis thaliana* were constituted. Finally, two SALK registered mutant line of *Arabidopsis thaliana* were complemented.

Chapter VI reports a third published paper that reviews the main approaches used to protect grapevine from fungal and oomycete diseases, starting from conventional breeding, which allowed the establishment of new resistant varieties, followed by biotechnological methods, such as transgenesis, cisgenesis, intragenesis and genome editing, and ending with more recent perspectives concerning the application of new products based on RNA interference (RNAi) technology. Evidence of their effectiveness, as well as potential risks and limitations based on the current legislative situation, are critically discussed.

CHAPTER I: Introduction

1.1 Biotechnology and climate changes

In the last decade there was an increasing interest of community on climate change. It is only in the very last years that climate change became a hot topic, despite already in 1896, the Swedish scientist Svante Arrhenius being the first to sustain that burning fossil fuels can lead to increased global warming (Arrhenius, 1896). Korres et al. (2016) defined climate change as identifiable long-term alterations in the state of the climate, such as increased temperatures, atmospheric CO_2 levels, precipitations, etc. Over geologic eras, many climate change phenomena occurred. However, the current term refers to variations in climatic patterns that took place over the last century and caused by anthropogenic activities that release greenhouse gasses (De Ollas et al., 2019). The Intergovernmental Panel on Climate Change (IPCC) reports a current temperature increase of 0,2 °C *per* decade due to the cumulative effect of past and present emissions. Furthermore, the predictions state that the level of 700 ppm of CO_2 (IPCC, 2018) will be reached by the end of the century, with a consequent temperature increase of 4°C (ICPP, 2018). Omitting the catastrophic consequences for mankind that these changes would bring and which we sincerely hope will be addressed in the appropriate locations from public decision makers, here the attention will be focused on the direct effect on agriculture. Warming is related to rainfall decrease, to a rise of evaporative demand, drought stress, and to extended growing season with a reduction in the duration of quiescence period. All these effects, together with many other, will result in the collapse of production, from both quantitative and qualitative point of view (Zandalinas et al., 2017; Medlyn, 2011; Gordo and Sanz, 2010). Grapevine-agroecosystem interactions are complex and featured by many variables with nested relationships affecting final product quantity and quality. Synthetically, all the effects that climate, soil and agronomical practises imprint in grapevine berry composition are called with the term "*terroir"* (Perin et al., 2020). This word is used to link a wine typicity with the geographical production area (Foroni et al., 2017). From a molecular point of view, *terroir* is the phenotypical manifestation of the peculiar interaction of a genotype with the surrounding environment (Van Leeuven and Seguin, 2006) whereas phenotypic plasticity is defined as the capacity of a single genotype to give rise to different phenotypes as a function of environment (Sultan et al., 2000). This makes it possible to produce wines with different organoleptic properties in different geographical areas starting from the same grapevine variety (Van Leeuven and Seguin, 2066; Dai et al., 2011). In grapevine phenotypic plasticity was observed also among plants of the same vineyard, bunches of the same plant and even among berries within the same cluster (Keller, 2010). The effects of climate

on grapevine have now been clarified and climate change is scientific evidence, thus the next challenge is to cope with these changes to preserve productions, their typicity, and their sustainability. Since an adaptation of agronomical practices is not enough for a long-term strategy (Delrot et al., 2020), it is undeniable that nowadays is essential to invest on biotechnology in order to rise it up to play a first order role in the implementation of an integrated management system, that could effectively respond to the problems put in place by the current challenges in agriculture.

1.2 The flowering process in grapevine

Grapevine (*V. vinifera* L.) is one of the most worldwide cultivated crops with a global production of 75 million tons of berries spread on a 7,5 million of hectares global surface. The main derivative product is wine, with a dominance of 68% of the viticulture, followed by table grape (30%) and raisins (1,8%), with a 0,2% leftover occupied by minor products, such as juices, jellies, ethanol, vinegar, grape seed oil, tartaric acid, and fertilizers (OIV, 2016). Despite the huge economic importance of the grapevine fruit, very few is known about the flowering process and the molecular mechanisms behind it, notwithstanding it directly affects winemaking production. From a botanical point of view, the reproductive system of grapevine American and Asian wild relatives (*non-vinifera* species) is characterized by dioecy, with distinct male and female flowers. On the contrary, the vast majority of the productive cultivars belonging to *V. vinifera* species have hermaphroditic flowers (Battilana et al., 2013). In this case, the flower is composed by four concentric whorls, respectively from outside to inside: sepals, petals, stamens and carpels (Fig. 1.1). The five sepals are fused together to compose calyx, an abiding structure that represents the support base on which the other whorls engage. All the inner organs of the flower are enclosed by the so-called calyptra, composed by the five petals fused to generate this sort of protective shell. Flower's heart is represented by the five stamens and pistil, composing androecium and gynoecium respectively. Each stamen is composed by a filament that sustains the anther, a bilocular structure containing four pollen sacs and organized in three distinguished layers: epidermis, endothecium and tapetum. Concerning the gynoecium, it is basically formed by the pistil (or carpel), a unique organ that is featured by the ovary divided into two locules containing in turn two ovules each one. The ovary is a swelling basal structure connected with the style, a spindly tubular channel ending with the stigma, deputy to pollen reception (Palumbo et al., 2019; Vasconcelos et al., 2009; Carmona et al., 2008). Since grapevine is a biennial woody plant, flowers are generated during the second season of a shot (Vasconcelos et al., 2009). During the budbreak period, the first whorl to differentiate is calyx, followed immediately by

calyptra, then the stigma appears, and the stamens develop, finally the complete development of pistil occurs (Gerrath, 1993; Pratt, 1971). Grapevine inflorescence is generally defined as a conical panicle with several branches: secondary branches form on central rachis and tertiary ones in turn on the seconds, thus making articulated cluster architecture and resulting in the most peripheral unity: the dichasium, a triplet of flowers (May, 2004). The differentiation rhythm of the whorls depends on the position of the flower inside the panicle and across the branches (Vasconcelos et al., 2008). Microgametogenesis consists in primary mother cell meiosis resulting in four haploid microspores. These last are contained in anther locules and maintained by a nutrient fluid formed by the degeneration of tapetum until the moment of dehiscence. At this point, endothecium detaches from the inner wall to the centre of the anther (Oberle, 1938). On the other hand, macrosporogenesis consists of the maturation of the four ovules in ovary locules, that basically act as protective barrier against mechanical damages and drying. Every single ovule is equipped by an embryonic sac characterized by diploid polar nuclei. Chronologically, egg development is immediately subsequent pollen maturation (Bernard and Charliès, 1987). When both pollen and ovules are mature, corresponding in a few days before fully bloom, it forms an abscission layer at the base of one of the petals, involving progressively the near one till involving all the grafting line of calyptra in calyx (Kozma, 2003). This process, together with anthers elongation that makes stamens exert pressure on the inner top wall, carries out the so-called capfall. Anther burst and pollen release attaining fertilization could take place before, during or after this process depending on grapevine variety. For this reason, it is not always possible to speak of cleistogamy *stricto sensu*, but more correctly of self-pollination (Meneghetti et al., 2006; Heazlewood and Wilson, 2004; Saudt, 1999; Pratt, 1971).

1.2.1 Molecular mechanisms of floral organs identity: the ABCDE-model

The ABCDE model represents the first approach trying to explain, from a molecular point of view, the complex genetic relationships occurring between transcription factors acting as activators and repressors towards the different whorls (Vasconcelos et al., 2009; Coen and Meyerowitz, 1991). In a first instance, this model involved three classes of genes, defined as “A”, “B” and “C”. A-class genes are responsible for sepal identity specification when expressed alone, while, when expressed together with several B-class genes, they specify petals identity (Jack, 2004). Moreover, in these two whorls, A-class genes ensure the repression of C-class genes (Coito et al., 2018). C-class genes are responsible for carpel identity when expressed alone, whereas they determine stamen identity when expressed in combination with B-class genes. Moreover, C-class repress A- genes in these two whorls, defining in this way a mutual repressive relation

between A- and C- class genes (Jack, 2004; Coen and Meyerowitz, 1991). The first version of this model included only the three aforementioned classes. Over the time, interest for this argument rose up in scientific environment, and the intensification of studies in this issue brought to discover two more gene classes affecting floral organ identity. More in detail, D-class genes were described to cooperate with some C-class genes to specify ovule identity inside the pistil locules (Vasconcelos et al., 2009; Carmona et al., 2008; Favaro et al., 2003). Finally, the last gene class to be added to the model was E-class. This class caused an entire revision of the model because is the only one expressed in all the whorls. Its function is not yet fully clear and understood, but this class genes seem to form complexes with the proteins of the other classes of the model, in a redundant functional way (Castillejo et al., 2005; Vandenbussche et al., 2033). Since this model was formulated and firstly described in *Arabidopsis thaliana* and its implementation over the time was achieved firstly in other model species like *Antirrhinum majus* and *Petunia hybrida* (Jack, 2004), but also in some economic importance crops (Fornara et al., 2003), it is now pivotal to gain the same status of knowledge also for grapevine, indeed still very little is known about sexual organ identity if compared with other species, and few genes were already characterized and associated to certain developmental stages of reproductive structure genesis (Carmona et al., 2008). A paper published by our laboratory provided a comprehensive screenshot of the transcriptional behaviour of 18 representative grapevine ABCDE genes encoding MADS-box transcription factors in a developmental kinetic process, from pre-anthesis to the postfertilization stage and in different flower organs, namely, the calyx, calyptra, anthers, filaments, ovary, and embryos (Fig. 1.1). The transcript levels found were compared with the proposed model for *Arabidopsis* to evaluate their biological consistency. With a few exceptions, the results confirmed the expression pattern expected based on the *Arabidopsis* data (Palumbo et al., 2019).

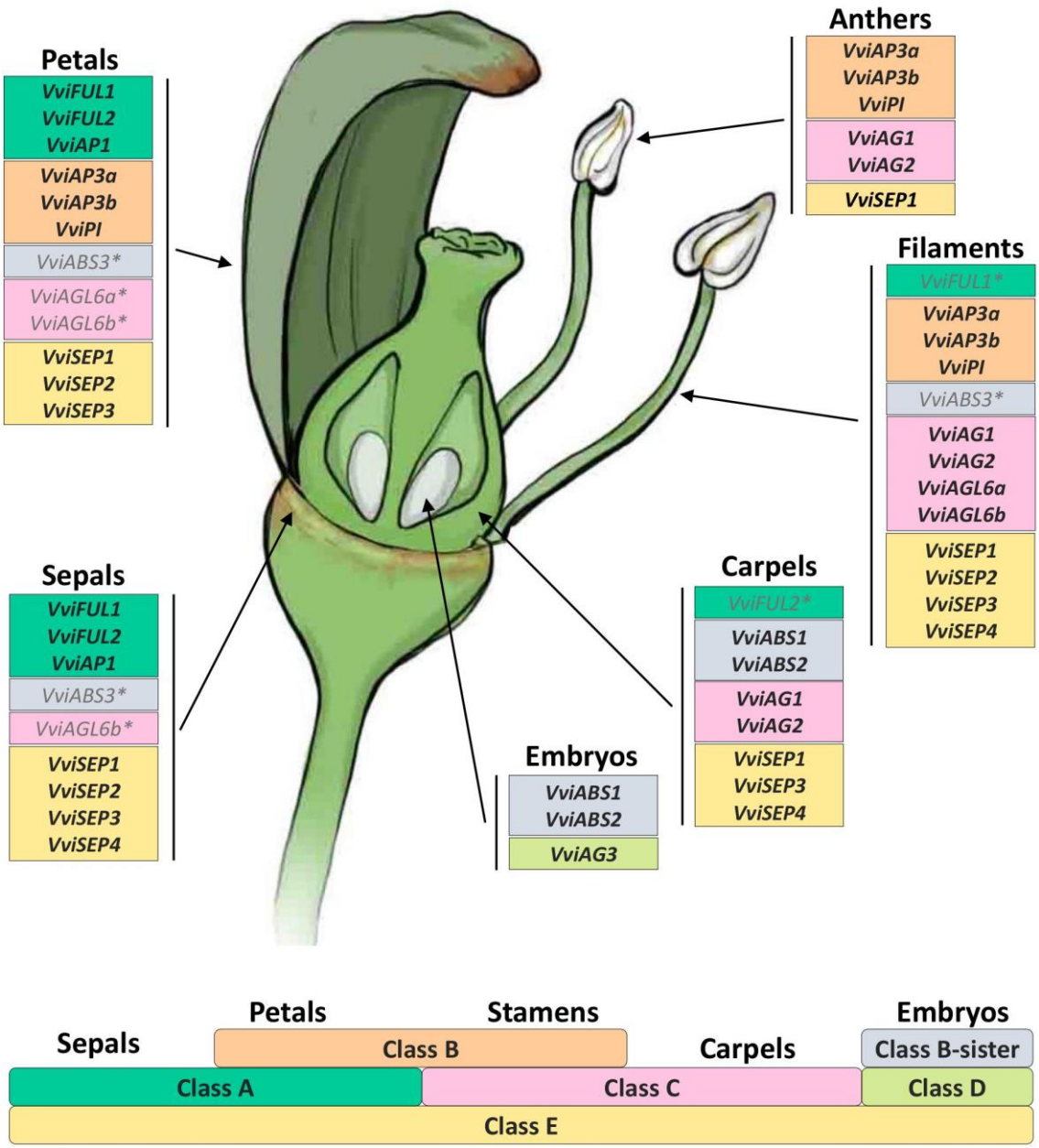


Figure 1.1. The floral model and the underlying ABCDE model of organ identity determination in *V. vinifera* L. The bottom part of the figure shows the genetic ABCDE model. According to this model, organ identity during flower development in the model organism A. thaliana is controlled by five classes of floral homeotic genes providing overlapping floral homeotic functions: A, B, C, D, and E. A+E genes specify sepals, A+B+E specify petals, B+C+E specify stamens, C+E specify carpels, and C+D+E specify ovules. In the upper part of the figure, we reported all genes belonging to different classes according to their expression in different flower organs analysed in another paper from our lab (Palumbo et al., 2019). Asterisks indicate genes whose expression was not expected in a given whorl based on the floral ABCDE model (picture taken from Palumbo et al., 2019)

1.3 Regulation of gene expression

All the cells constituting a complex organism contain exactly the same genetic information, although their final transcriptional, proteomic and biochemical profile can deeply change as a

function of differentiation, which underlies the development of the organisms (Scott, 2000). The product is affected by several factors that could exert influence at different levels, making the final protein quantitatively and qualitatively different (Li et al., 2009). These different levels in eukaryotic organisms are known to be physically compartmentalized and controlled by several determinants (Bilas et al., 2016). Firstly, the functional initiation complex is assembled on the DNA strand. The initiation site is then localized and bound by the RNA-polymerase II, which is responsible for the synthesis of the *pre-m*RNA. After the maturation steps involving introns removal, exons link and attachment of 5' cap and 3' polyA tail to the transcript ends, mature *m*RNA is therefore transferred to the cytoplasm where translation occurs (Phillips, 2008; Klug and Cummings, 2003). The linear aminoacidic chain is formed starting from the nucleotides sequence and follows the formation of the *m*RNA-ribosome complex. The efficiency of translation initiation is function of the region flanking the start codon, called Translation Initiation Context (TIC), and of its distance from the 5' end. The most effective TIC in dicots seems to be TAAACAATGG (Koul et al., 2012).

1.3.1 Promoter structure in eukaryotes

In animals, the regulatory sequences determining the correct spatiotemporal expression of a gene can extend over tens kilobases (Kb). In contrast, regulatory sequences in plants usually span much shorter DNA intervals, often less than 1 or 2 kb (Bonifer, 2000). Based on their structure regulators are classified in different families. In particular, *cis*-elements are non-coding DNA sequences that interact with *trans*-elements, proteins forming complexes and affecting gene expression at different levels (Vaughn et al., 2012; Venter and Botha, 2010). Despite the structural differences between animal and plant *cis*-regulatory and promoter regions, regulation of gene expression is often the result of multiple inputs, reflecting, or taking advantage of, the combinatorial nature of the mechanisms of eukaryotic transcription. Multiple stimuli can converge through different *cis*-acting elements on a promoter to co-ordinately regulate the expression of the corresponding gene (Arnone and Davidson, 1997; Yuh et al., 1998). *Cis*-acting elements are generally organized as modules: both in animals and plants, the regulatory region of a gene can be partitioned into discrete sub-elements, each one containing one or several TF binding sites and performing a determined regulatory function (Benfey and Chua, 1990; Arnone and Davidson, 1997). *Cis*-elements are localized in the promoter sequence, a region upstream the coding part of a gene. TFs bind RNA polymerase II, which in turn links its own binding sites, and *cis*-elements, giving rise to transcription initiation (Porto et al., 2014; Russell, 1996). From a structural point of view, promoter can be divided in two main regions that contain *cis*-elements

and interact with *trans*-acting factors: the core and the distal region, these last characterized by enhancing and/or silencing activity (Porto et al., 2014; Klug and Cummings, 2003; Fig. 2.1). The core region is deputed to the correct regulation of the transcription initiation by the RNA polymerase II, and it is composed of some subunits not universally conserved, like the TATA-box, the CAAT-box and the initiator region (Bilas et al., 2016; Porto et al., 2014; Juven-Gershon and Kadonaga, 2010). The TATA-box is an 8 bp sequence located 25-30 bp upstream the transcription start site (TSS). This sequence is composed exclusively by A and T nucleotides with the consensus sequence TATA(A/T)A(A/T) and it is flanked by a GC-rich region (Twyman, 2003; Lewin, 2001). TFs and RNA polymerase II bind to the TATA-box in a determined order. First, the transcription factor TFIID is bound, forming a complex with TFIIA. TFIIB attaches to this complex via TATA-binding protein (TBP), which is a part of TFIID, and as a result of direct interaction with the DNA strand. RNA polymerase II is bound in the next step and TFIIF attaches to it. Thereupon TFIIE and TFIIH join the complex and the full transcription apparatus is formed (Kwak and Lis, 2013). The initiator region (Inr) usually overlaps with TSS or however it is very close to it and shows a consensus sequence YYCARR (Twyman, 2003). This region cooperates with TATA-box in transcription initiation and is recognized by TFIID (Kadonaga, 2012). The following *cis*-element is the CAAT-box which usually locates 80 bp upstream the TSS. Because of the high susceptibility to the mutations featuring the sequence of this CRE, it is considered one of the most affective elements for the effectivity of the transcription (Bilas et al., 2016). It seems that in plant TATA-box is substituted by the AGGA-box element (Porto et al., 2014). Not all these elements are necessarily present in a promoter core at the same time. Sometimes, it could lack one, some or all of them and in those cases the initiation of transcription may occur in several places and is not strictly prescriptive (Bilas et al., 2016). For example, in a TATA-box lacking core, the downstream promoter element (DPE), which usually is 28-32 bp downstream from the Inr adenine, could be the responsible for TFIIB protein binding (Bilas et al., 2016). TATA-box, CAAT/AGGA-box and Inr were mentioned here since they are the most common *cis*-elements in the core promoter, but they are not the unique ones, indeed motifs interacting with *trans* proteins like those are several and their function must be yet clarified. Unlike the previously described elements, enhancer and silencer *cis*-elements can be located (i) further from the TSS, (ii) in non-coding intron sequences, (iii) near their target core promoter or even (iv) within the transcribed region of a gene body, and can act under specific environmental conditions or in definite tissues (Spitz and Furlong, 2012; Ong and Corces, 2011; Peremarti et al., 2010). Interestingly, *cis*-elements can have a combinatorial effect in plants, and this was already proven in *Arabidopsis thaliana* light responsive elements (LREs), found in the promoter of photosynthesis-related genes (Riechmann, 2002). The functioning of these genes is the results

of a coordinated action among LREs that enable to respond to a wide spectrum of light through the phytochrome signal transduction pathways, to cope with the chloroplast developmental state, and to confer a photosynthetic-cell specific expression pattern, therefore satisfying the strict definition of light-inducible (Chattopadhyay et al., 1998b; Puente et al., 1996). Surprisingly, LREs cannot singularly induce a satisfactory light response (Riechmann, 2002). Moreover, the promoter of the gene LEAFY, which is a responsible for the flower meristem identity, was observed to act as information processing system by integrating external stimulus of daylength and the endogenous signal of gibberellins in order to manage the flowering time in *Arabidopsis thaliana* (Blàzquez and Weigel, 2000).

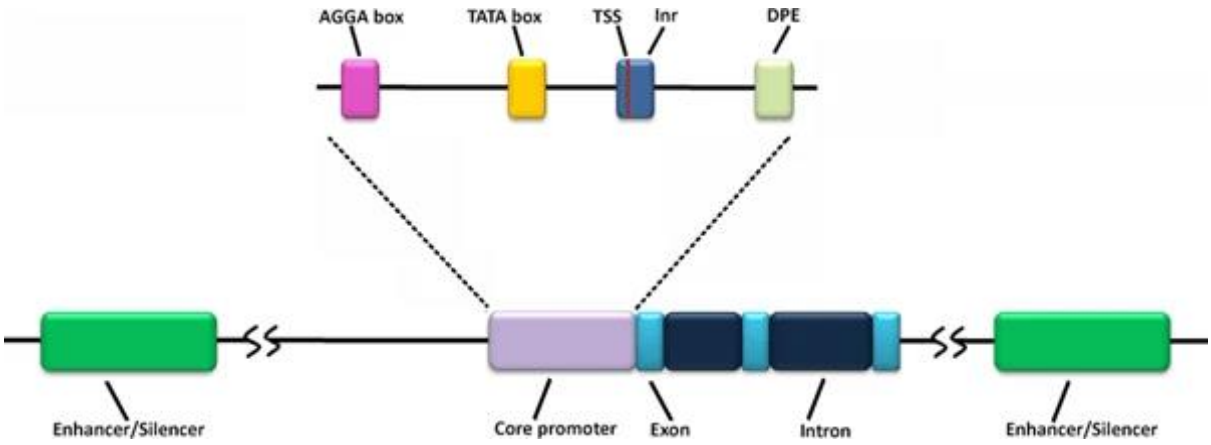


Figure 2.1. General organization of gene structure in plants (picture taken from Klug and Cummings, 2003)

1.4 Transcription factors

The transcription factors (TFs) discussed in chapter 1.3.1 belong to the basic transcription apparatus which in eukaryotic organisms forms protein complexes enabling the correct transcriptional initiation. Now the discussion will focus on the classical type of TFs, or rather the ones able to directly bind DNA. These are proteins with distinct and functionally separate domains that can modulate gene expression by activating or repressing transcription and having tissue, cell-type, temporal, or stimulus-dependent specific activity (Riechmann, 2002). Depending on their binding domain, TFs can be classified in several families (Luscombe et al., 2000). The abundance of TFs, meant as the regulatory network that they form, was proposed as indicator of the biological complexity of an organism and for plants, it could be considered as an index of the articulation of secondary metabolism and of the interactions occurring between genotype and environment (Szathmáry et al., 2001). The next chapters will be focused on the MYB, WRKY and NAC transcription factors families, for this reason here below is provided a general overview on these three families in plants, and then specifically in grapevine.

1.4.1 MYB transcription factors

A common feature of all MYB proteins is the DNA-binding domain at N-terminus, which is conserved in all eukaryotic organisms (Martin and Paz-Ares, 1997) and that was already proven to bind DNA in a sequence specific manner (Howe et al., 1990; Sakura et al., 1989). The first MYB to be discovered was *v-Myb* in avian myeloblastosis virus (Weston and Bishop, 1989). Subsequently many other MYB TFs have been identified in animals, fungi, and plants (Cao et al., 2016; Stacke et al., 2001; Rosinski and Atchley, 1998). The MYB family members typically contain one, two, three, or four imperfect repeats (Cao et al., 2016; Stacke et al., 2001). Each repeat is constituted by 50-53 nucleotides and encodes for three α-helices and the second and third of those form a helix-turn-helix (HTH) structure which allows to combine with target DNA and regulate gene expression (Zhou et al., 2020; Dubos et al., 2010; Martin and Paz-Ares, 1997; Lipsik, 1996). Based on the number of repeats in the structure of the DNA-binding domain, MYBs are classified into four subfamilies called 1R-MYB, 2R-MYB, 3R-MYB, and 4R-MYB (Cao et al., 2020) (Fig. 3.1). The 1R-MYB subfamily is featured by one MYB domain and plays an essential role in regulating plant transcription and maintaining the structure of chromosomes. 1R-MYBs are in turn clustered into several subgroups, including I-box-binding-like, TBP-like, TRF-like, CPC-like, CCA1-like, and other MYB-related proteins (Chen et al., 2006). MybSt1 is the first identified MYB-related protein in plants and can be used as a transcriptional activator in potato (Baranowskij et al., 1994). Since it is the second-largest subfamily of MYBs, 1R-MYB is widely

distributed in plants: there are 64 members in model plants like *Oryza sativa* and 68 members in *A. thaliana* (Dubos et al., 2010). 1R-MYB TFs are involved in flower and fruit development, response to phosphate starvation, circadian clock control, and phenylpropanoid-derived compounds (Feller et al., 2011). The second subfamily is the 2R-MYB which is characterized by the presence of two MYB domains. This subfamily is the most widely existed among the four and was probably originated by a duplication event in a 1R-MYB or by a loss of a R1 domain by a 3R-MYB (Jiang et al., 2004; Rosinski et al., 1998). There are more than 90 members in *O. sativa* and 120 in *A. thaliana* (Dubos et al., 2010). Basing on the conserved amino acid sequence motifs present in their most C-terminal MYB domain, 2R-MYBs can be subclustered into 28 subgroups (Shelton et al., 2012; Dubos et al., 2010; Stacke et al., 2001). Due to their wide diffusion, 2R-MYBs could play an important role in plant-specific mechanisms, indeed it was already described the involvement in determination of cell fate and identity, regulation of development, hormone signal transduction, primary and secondary metabolism – with pivotal mention of phenylpropanoids and flavonoids pathways - and response to abiotic and biotic stresses (Cao et al., 2020; He et al., 2019; Gonzalez et al., 2008; Lepiniec et al., 2006; Borevitz et al., 2000). The next subfamily is the 3R-MYB one, characterized by a three-repeats binding domain. It is quite small, indeed only five members were individuated both in *O. sativa* and *A. thaliana* (Dubos et al., 2010). 4R-MYB is the last subfamily, and the members show four R1/R2 domain. These kinds of MYBs are the less numerous and in *O. sativa* and *A. thaliana* were individuated only one and two, respectively (Cao et al., 2020; Katyiar et al., 2012). 4R-MYBs still need to be described in their biological function. Overall, in MYBs can be recognized four different mechanisms of regulation: the first two are the regulation of proteins interaction mechanisms and the regulation of transcriptase mechanisms (Cao et al., 2020). The third is peculiar of 2R-MYBs and it is the regulation of redox reactions: the second domain contains a conserved amino acid residue composed of four cysteines, two of which are oxidized to form a disulphide bond (S-S) within the protein molecule, in order to avoid the DNA binding domain to be oxidated (Cao et al., 2016; Morse et al., 2009). The last mechanism is the regulation of ubiquitin, which in stress conditions plants can be performed both in an ABA dependent or independent way (Cui et al., 2013; Fujita et al., 2011; Reyes et al., 2007). In grapevine, compared to *Arabidopsis*, it was observed a strong R2R3-MYBs expansion due to dispersed mechanisms of transposable elements activity related to grapevine genomic variability (37%) and to the tandem duplication (21.3%; Wong et al., 2016; Benjak et al., 2008). In particular, these last affects MYBA (9 out of 14 genes) and MYBPA (5 out of 15 genes) TFs, involved in anthocyanin and protoanthocyanidin biosynthesis respectively (Wong et al., 2016). This could be explicable because of the anthropic selection of grapevine related to a higher concentration of flavonoids in berry and wine (Matus

et al., 2008). Even if more moderately, a similar expansion due to the tandem duplication events was observed in C2 repressors clade (Wong et al., 2016). Also in this case, the reason of this situation could be breeding, indeed also these MYBs were described related to the metabolic pathways of flavonoids and phenylpropanoids (Cavallini et al., 2015; Huang et al., 2014). A further confirmation could be that this type of transcriptional sub-specialization after a duplication event was also observed in other plant species where a high flavonoids content is a wanted trait (Albert, 2015). Moreover, gene set enrichment analysis (GSEA) conducted on top 100 co-expressed genes of all grapevine MYBs gene co-expression network (GCN) highlighted that – among secondary metabolic pathways – the most enriched terms are related to phenylpropanoid and terpenes derivatives (Wong et al., 2016). As general observation, extensive research has confirmed that the expression of anthocyanin biosynthesis genes is controlled by the MBW protein complex, with R2R3-MYB transcription factors playing a central role in the coordinated activation of genes specific to anthocyanin pathways (Costantini et al., 2015; Albert et al., 2014).

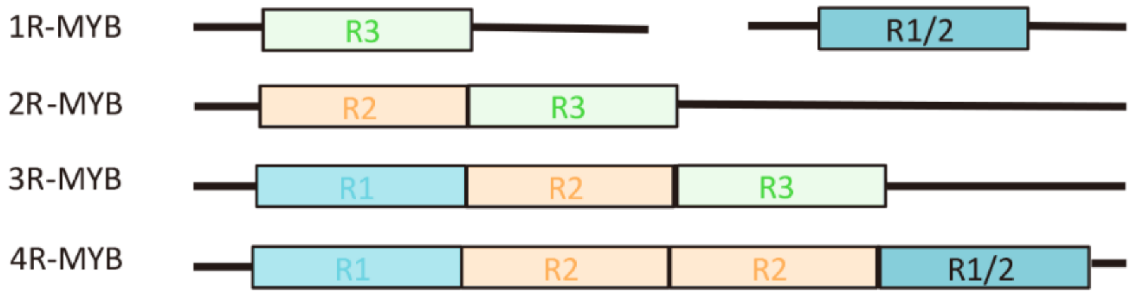


Figure 3.1. MYB sub-families structure based upon the repeated domains (picture taken from Cao et al., 2020)

1.4.2 WRKY transcription factors

The first WRKY was described in sweet potato as a "DNA binding proteins that played potential roles in the regulation of gene expression by sucrose" (Ishiguro and Nakamura, 1994). Since that moment, a huge amount of WRKY proteins have been discovered, described and characterized, not only in model plants, but also in crop species featured by an economical relevance (Rushton et al., 2010). The DNA binding domain is the defining feature of WRKY transcription factors: it contains the WRKY amino acid signature – constituted by about 60 residues - near the N-terminus and a novel zinc-finger structure at the C-terminus (Eulgem et al., 2000; Rushton et al., 1996). The WRKY domain consists of a four-stranded b-sheet, with the zinc coordinating Cys/His residues forming a zinc-binding pocket. The WRKYGQK residues correspond to the most N-

terminal β-strand, which partly protrude from one surface of the protein, thereby enabling access to the major DNA groove and contacts with the DNA. It was proposed that the β -strand containing the WRKYGQK motif contacts an approximately 6-bp region, being consistent with the length of the W box (Yamasaki et al., 2005). This last is the minimal consensus sequence (TTGACC/T) required for specific DNA binding and goes in parallel with the conservation of the WRKY domain (Ciolkowski et al., 2008; Eulgem et al., 2000; Rushton et al., 1996). At the beginning, WRKYs were divided into three groups (Group I, Group II and Group III) basing on the repeats of WRKY domain and the zinc finger structure (Eulgem et al., 2000). Subsequently, Group II were subdivided into five subgroups based on the primary aminoacidic sequence: IIa, IIb, IIc, IId and IIe (Fig. 4.1). On the other hand, in superior plants, following a pure phylogenetic approach, is more accurate to individuate a total amount of five subgroups, or rather Group I, Group IIa + Iib, Group Iic, Group Iid + Iie and Group III, since Group II seems not to be monophyletic (Rushton et al., 2008; Zhang and Wang, 2005). For which concerns their biological function, WRKYs are central components of many aspects of the innate immune system of the plant, including microbe- or pathogen-associated molecular pattern triggered immunity, effector-triggered immunity, basal defence and systemic acquired resistance in a very wide range of species (Eulgem and Somssich, 2007). Apart from plant defence systems, WRKY TFs are related to a wide range of abiotic stress responses, like heat and drought tolerance (Wu et al., 2009), salt tolerance (Jiang and Deyholos, 2009; Qiu and Yu, 2009), and cold and mannitol tolerance (Eulgem, 2006). In addition, this TFs family is involved also in several developmental processes which regulate some biological cycle phases of plants, like seed dormancy, germination and development, and senescence (Jiang and Deyholos, 2009; Jing et al., 2009; Ulker et al., 2007; Xie et al., 2006; Luo et al., 2005; Robatzek and Somssich, 2002). The interesting point of WRKYs is the fact that are TFs almost exclusively distinguishing plants since in incredibly few non-vegetal species some members of this family have been individuated (Rushton et al., 2010). The reason why of this fact remains unclear, but a hypothesis could be that the WRKY family may represent a lineage-specific expansion following the emergence of an ancestral WRKY gene from a transposon, and it seems more likely that a few non-plant species may have gained WRKY genes rather than the wholesale loss of WRKY genes in multiple lineages and the sporadic distribution of the WRKY family in phylogenetically distant eukaryotes is suggestive of their possible spread through intra-eukaryotic lateral transfers (Babu et al., 2006). As well as for what generally observed, also in grapevine WRKYs play a key role in several biological processes, very different from each other. If we consider *VviWRKY13* for example, its ectopic expression in *Arabidopsis* showed several diversified phenotypes like delayed seed germination, smaller stomatal aperture size and promotion of ABA and Ethylene synthesis (Hao et al., 2017; Ma et al.,

2015). Regarding the fruit development, *VviWRKY26* was observed to be highly expressed in the inner integument of the seed coat during the early stage of berry development, suggesting an involvement in protoanthocyanidin related genes regulation, by also affecting seed development and plant fertility (Amato et al., 2017; Dilkes et al., 2008). Another WRKY, namely *VviWRKY22*, seems to be involved in berry molecular mechanisms too, in this case by affecting the sugar metabolism (Huang et al., 2021), and on the other hand, *VviWRKY2* directly controls lignin biosynthesis and xylem development (Guillaumie et al., 2010). Moreover, the transcription of most WRKY protein members can be altered by at least one stress treatment, which suggests that they are widely involved in plant response to biotic or abiotic stress (Wu et al., 2022). In particular, *WRKY* genes were often seen to be induced by the typical disease-related hormones (i.e., salicylic acid, jasmonic acid, ethylene and abscisic acid) and also involved in regulating downstream responses mediated by these hormones (Wu et al., 2022). Even if the functions of some grapevine WRKY members under biotic stress have been studied and functionally characterized, most of their molecular mechanisms leading to the functions are still poorly understood. Noteworthy to point out, is the fact that many WRKYs are highly correlated to the stilbene synthases (secondary metabolites playing an important role in protecting plants from pests, pathogens and abiotic stresses) expression in microarray and RNA-Seq data, and some of them were proven to even bind the promoter of these genes (Qu et al., 2021; Wang et al., 2020; Vannozzi et al., 2018; Schnee et al., 2008; Langcake and Pryce, 1976).

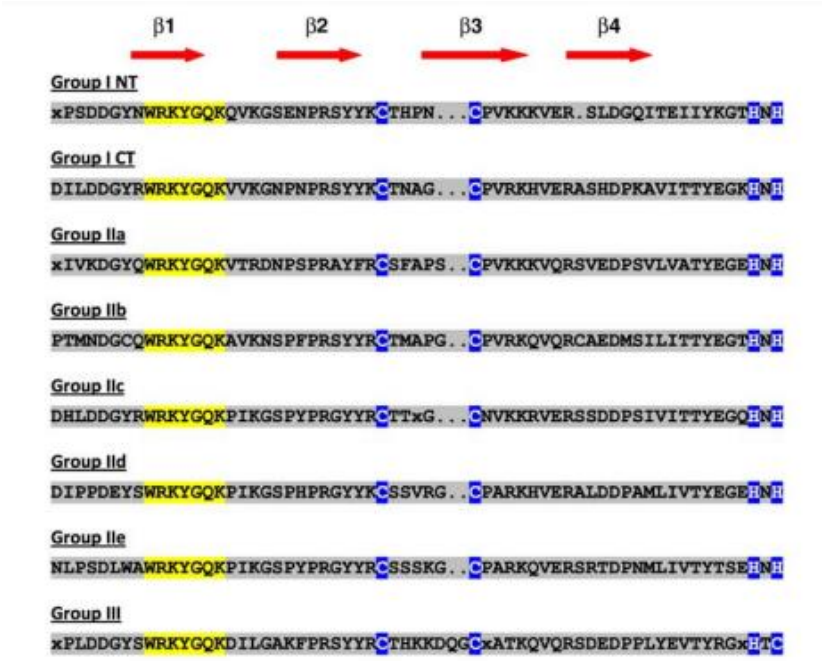


Figure 4.1. The WRKY domain consensus for each WRKY subfamily in higher plants. The WRKY motif is highlighted in yellow and the cysteines and histidines that form the zinc finger are shown

in blue. The four β-strands are shown in red. I CT and I NT denote the N-terminal and C-terminal WRKY domains from Group I WRKY proteins (picture taken from Rushton et al., 2010)

1.4.3 NAC transcription factors

NACs constitute one of the largest plants families of transcription factors. Extensive investigation identified several NAC TFs in a wide range of vegetal species, 117 in *Arabidopsis*, 151 in rice, 79 in grape, 26 in citrus, 163 in poplar, and 152 both in soybean and tobacco (Le et al., 2011; Nuruzzaman et al., 2012a, 2010; Hu et al., 2010; Rushton et al., 2008). NAC acronym is derived from three genes that were initially discovered to contain a particular domain (the NAC domain): NAM (for no apical meristem), ATAF1 and −2, and CUC2 (for cup-shaped cotyledon. Aida et al., 1997; Souer et al., 1996). This domain is localized at N-terminus of the protein and consists in a 150-160 aminoacidic residues divided into five subunits, called from A to E (Ooka et al., 2003). If on one side the NAC domain is conserved among different proteins of this family and defines the DNA-binding properties with the common motif (CATGTG) conserved in the target genes, (Nuruzzuman et al., 2013), on the other hand, C-terminus of NAC TFs is extremely divergent (Ooka et al., 2003) and reason why were observed regulatory differences between the transcriptional activation activity of NAC proteins (Jensen et al., 2010; Yamaguchi et al., 2008; Xie et al., 2000). Moreover, since C-terminus shows protein binding activity by operating as a functional domain, can act as a transcriptional activator or repressor (Nuruzzuman et al., 2013; Kim et al., 2007b; Hu et al., 2006; Tran et al., 2004). From a functional point of view, several studies show that many NACs play roles in the transcriptional reprogramming associated with plant immune responses, indeed there are many examples in which the overexpression or knockdown has effects on plant defense (Jensen et al., 2010, 2008, 2007; He et al., 2005; Delessert et al., 2005). Moreover, NAC TFs were observed in many plant species to modulate their expression in response to attack of bacteria, fungi, and viruses (Xia et al., 2010a; Wang et al., 2009a, 2009b; Jeong et al., 2008). Special mention deserves the NACs involvement in reactive oxygen species (ROS) production. The production of these compounds is induced by pathogens attack and mediate the tolerance against several biotic stresses (Davletova et al., 2005). Some experiments suggest a direct NACs role in both these mechanisms of response and tolerance ROS-mediated (Nakashima et al., 2007). Since ROS are closely related to programmed cell death (PCD) – process massively occurring as infection response but also in leaf senescence – and for the observed induction of some NACs in senescence process, it is likely conceivable that these TFs may function at the node of convergence between the pathogen defense and senescence signaling pathways (Nuruzzaman et al., 2013; Lee et al., 2012; de Zelicourt et al., 2012; Distelfeld

et al., 2012; Guo and Gan, 2006; Guo et al., 2004; Gepstein et al., 2003). Moreover, NACs involvement was extensively observed in a wide range of abiotic responses in different plant species, like drought, salinity, cold, heat, boron-toxicity, waterlogging, wounding, and oxidative stress (Zhou et al., 2013; You et al., 2013; Liu et al., 2011; Ochiai et al., 2011; Christianson et al., 2009; Sindhu et al., 2008), but also in other biological processes, *e.g.,* phytohormones signaling ABA-, JA- and SA-mediated (Jensen et al., 2010; Bu et al., 2009, 2008; Greve et al., 2003) that anyway seem to make converging abiotic and biotic stresses through the action of ATAF subfamily of NAC TFs (Jensen et al., 2007; Delessert et al., 2005). Wang et al. (2013) identified 74 different grapevine NACs predicted on 12x assembled *V. vinifera* PN40024 genome. 70 out of 74 were tested in qRT-PCR to verify the expression in all the grapevine tissues and the results shown that several *NACs* were expressed specifically in one tissue (except for root, where no *NAC* were observed to be specifically expressed). Moreover, no single *NAC* shown a constitutive expression profile in all tissues under exam. On the other hand, all tissues expressed at least one *NAC* gene, suggesting that the *NAC* gene family plays roles during the majority of grapevine developmental stages (Wang et al., 2013). Also for NAC family, there are several studies assessing the involvement of these TFs in different grapevine biological processes, like drought, salt and cold stresses, JA and ABA signaling, *Botrytis cinerea* resistance, berry size variation and leaf senescence (Zhu et al., 2019; Fang et al., 2016; Tello et al., 2015; Le Henanff et al., 2013).

1.4.4 Methods for transcription factor binding sites identification

Over time, to investigate within a genome the localization of the transcription factor binding sites (TFBS), different assays were implemented, basing on several approaches and technologies. Here is provided an overview on the main ones applied to study the complete set of the *cis*-element of a given TF, namely the cistrome. The first method which allowed to create high-resolution genome-wide maps of the in vivo interactions between DNA-associated proteins and DNA was defined by the union of chromatin immunoprecipitation (ChIP) and whole-genome DNA microarrays, the so-called ChIP-chip. For this procedure, cells are grown under the desired experimental condition and then fixed with formaldehyde which crosslinks proteins. After crosslinking, the extract is sonicated to shear the DNA fragments to the desired size, usually 1 kb or smaller. DNA fragments crosslinked to the protein of interest are then precipitated, the formaldehyde crosslinks are then reversed and the DNA is purified. Enriched DNA is then labeled with a fluorescent molecule and a microarray analysis is performed (Buck and Lieb, 2003; Iyer et al., 2001; Lieb et al., 2001; Ren et al., 2000). With the advent of next generation sequencing (NGS), it was possible to directly sequence the DNA-fragments of interest instead of being

hybridized on an array: arose in this way the Chromatin Immunoprecipitation Sequencing (ChIP–seq), one of the first NGS applications (Barski et al., 2007; Johnson et al., 2007; Mikkelsen et al., 2007; Robertson et al., 2007). This upgrade enabled many adventages. First, the maximum resolution is in the order of magnitude of a single nucleotide. The coverage is limited only by alignability of reads to the genome and increases with read length: many repetitive regions can be covered, contrary to the ChIP-chip where repetitive regions were usually masked out. Sensible reduction of the ChIP-chip biases due to the cross-hybridization between probes and nonspecific targets. The required amount of ChIP DNA drastically decreased, passing from a few micrograms to 10-50 nanograms. In ChIP-Seq, the dynamic range of signal detection is limitless, contrary to ChIP-chip, in which there is a saturation in case of high signal. In ChIP-chip, the amplification is strictly necessary, while in ChIP-Seq is not, indeed single-molecule sequencing without amplification is available, and finally the multiplexing is achievable (Park, 2009). Certainly, ChIP-Seq is a very widespread method for TFBS identification but is not the only one. Mapping DNase I hypersensitive (HS) sites has historically been a valuable tool for identifying all different types of regulatory elements, including promoters, enhancers, silencers, insulators, and locus control regions. This method utilizes DNase I to selectively digest nucleosome-depleted DNA (presumably by transcription factors), whereas DNA regions tightly wrapped in nucleosome and higher-order structures are more resistant. The procedure was then refined until becoming a high-throughput method (DNase-Seq) that identifies DNase I HS sites across the whole genome by capturing DNase-digested fragments and sequencing them by high-throughput NGS (Song and Crawford, 2010). The DNase-Seq assay was then integrated to histone modifications occupancy analysis, obtaining a method for the detection of open chromatin regions and active binding sites (Gusmao et al., 2014). This assay, called HINT, was then combined to ATAC-Seq to generate HINT-ATAC (Li et al., 2019). In this regard, ATAC-seq (transposase-accessible chromatin using sequencing) captures open chromatin sites and reveals the interplay between genomic locations of open chromatin, DNA-binding proteins, individual nucleosomes and chromatin compaction at nucleotide resolution (Buenrostro et al., 2013). HINT-ATAC extends HINT by the proposal of a generalized framework for cleavage event counting and bias correction. This new framework allows cleavage events to be displaced from the read start and is based on a probability distribution assuming dependency between nucleotide positions. This allows to consider bias spanning larger genomic areas. HINT-ATAC also extends HINT by the use of strand-specific and fragment-size decomposition of cleavage signals as input (Li et al., 2019). Finally, the last TFBS identification method mentioned is the DAP-Seq. Since this assay is the fulcrum of the present study, a detailed discussion is referred to the dedicated chapter (Chapter III).

BIBLIOGRAPHIC REFERENCES

Aida M., Ishida T., Fukaki H., Fujisawa H. and Tasaka M. (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. *Plant Cell* 9, 841–857. Doi: 10.1105/tpc.9.6.841

Albert N. W. (2015). Subspecialization of R2R3-MYB repressors for anthocyanin and proanthocyanidin regulation in forage legumes,Front. Plant Sci., 6, 1–13.

Albert N. W., Davies K. M., Lewis D. H., Zhang H., Montefiori M., Brendolise C., et al. (2014). A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. Plant Cell. 26, 962–980. doi: 10.1105/tpc.113.122069

Amato A., Cavallini E., Zenoni S., Finezzo L., Begheldo M., et al. (2017). A grapevine TTG2-like WRKY transcription factor is involved in regulating vacuolar transport and flavonoid biosynthesis. Frontiers in plant science 7:1979

Arabidopsis Genome Initiative Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature. 2000;4081(1):796–815.

Arnone M. I., Davidson E. H. The hardwiring of development: organization and function of genomic regulatory systems. Development. 1997;1241(1):1851–1864.

Arrhenius S. (1896). On the influence of carbonic acid in the air upon the temperature of the ground. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 41(251), 237–275. (Fifth series). https://doi.org/10.1080/14786449608620846

Babu M. M., Iyer L. M., Balaji S. & Aravind L. (2006). The natural history of the WRKY–GCM1 zinc fingers and the relationship between transcription factors and transposons. *Nucleic acids research*, *34*(22), 6505-6520.

Baranowskij N., Frohberg C., Prat S., & Willmitzer L. (1994). A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. The EMBO journal, 13(22), 5383-5392.

Barski A., Cuddapah S., Cui K., Roh T. Y., Schones D. E., Wang Z., ... & Zhao K. (2007). High-resolution profiling of histone methylations in the human genome. Cell, 129(4), 823-837.

Battilana J., Lorenzi S., Moreira F. M., Moreno-Sanz P., Failla O., Emanuelli F., et al. (2013). Linkage mapping and molecular diversity at the flower sex locus in wild and cultivated grapevine reveal a prominent ssr haplotype in hermaphrodite plants. Mol. Biotechnol. 54, 1031–1037. Doi: 10.1007/s12033-013-9657-5

Benjak A., Forneck A., Casacuberta J. M. 2008, Genome-wide analysis of the 'cut-and-paste' transposons of grapevine,PLoS One, 3, e3107.

Benfey P., Chua N. H. The Cauliflower Mosaic Virus 35S promoter: combinatorial regulation of transcription in plants. Science. 1990;2501(1):959–966.

Bernard A. C. and C. Chaliès. 1987. Meiosis in *Vitis vinifera* cvs Grenach and Merlot. In Proceedings of the 3rd International Symposium on Grapevine Physiology. J. Bouard and R. Pouget (eds.), pp. 19-23. OIV, Paris.

Biłas R., Szafran K., Hnatuszko-Konka K., & Kononowicz A. K. (2016). Cis-regulatory elements used to control gene expression in plants. Plant Cell, Tissue and Organ Culture (PCTOC), 127(2), 269-287.

Blázquez M. A., Weigel D. Integration of floral inductive signals in Arabidopsis. Nature. 2000;4041(1):889–892.

Bonifer C. Developmental regulation of eukaryotic gene loci: which cis-regulatory information is required? Trends Genet. 2000;161(1):310–315

Borevitz J. O., Xia Y., Blount J., Dixon R. A., Lamb C. J. Activation tagging identifies a conserved myb regulator of phenylpropanoid biosynthesis. Plant Cell 2000, 12, 2383–2393.

Bu Q., Jian H., Li C. B., Zhai Q., Zhang J., Wu X., et al. (2008). Role of the *Arabidopsis thaliana* NAC transcription factors ANAC019 and ANAC055 in regulating jasmonic acid signaled defense responses. *Cell Res*. 18, 756–767. Doi: 10.1038/cr.2008.53

Bu Q., Li H., Zhao Q., Jiang H., Zhai Q., Zhang J., et al. (2009). The Arabidopsis RING finger E3 ligase RHA2a is a novel positive regulator of abscisic acid signaling during seed germination and early seedling development. *Plant Physiol*. 150, 463–481. Doi: 10.1104/pp.109.135269

Buck M. J. & Lieb J. D. (2004). ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. Genomics, 83(3), 349-360.

Buenrostro J., Giresi P., Zaba L. et al. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 10, 1213–1218. https://doi.org/10.1038/nmeth.2688

Cao Y., Han Y., Li D., Lin Y., Cai Y. J. Myb transcription factors in 30rabido pear (*Pyrus bretschneideri* rehd.): Genome-wide identification, classification, and expression profiling during fruit development. Front. Plant Sci. 2016, 7, 577.

Cao Y., Li K., Li Y., Zhao X. & Wang L. (2020). MYB transcription factors as regulators of secondary metabolism in plants. Biology, 9(3), 61.

Carmona M. J., Chaïb J., Martínez-Zapater J. M. and Thomas M. R. (2008). A molecular genetic perspective of reproductive development in grapevine. J. Exp. Bot. 59, 2579–2596. Doi: 10.1093/jxb/ern160

Castillejo C., Romera-Branchat M. and Pelaz S. (2005). A new role of the *Arabidopsis* SEPALLATA3 gene revealed by its constitutive expression. *Plant J.* 43, 586–596. Doi: 10.1111/j.1365-313X.2005.02476.x

Cavallini E., Matus J. T., Finezzo L. et al. (2015). The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine,Plant Physiol., 167, 1448–70.

Chattopadhyay S., Puente P., Deng X. W., Wei N. Combinatorial interaction of light-responsive elements plays a critical role in determining the response characteristics of light-regulated promoters in Arabidopsis. Plant J. 1998b;151(1):69–77.

Chen Y., Yang X., He K., Liu M., Li J., Gao Z., Lin Z., Zhang Y., Wang X., Qiu X. et al. The myb transcription factor superfamily of 30rabidopsis: Expression analysis and phylogenetic comparison with the rice myb family. Plant Mol. Biol. 2006, 60, 107–124.

Christianson J. A., Wilson I. W., Llewellyn D. J., and Dennis E. S. (2009). The lowoxygen induced NAC domain transcription factor ANAC102 affects viability of *Arabidopsis thaliana* seeds following low-oxygen treatment. *Plant Physiol*. 149, 1724–1738. Doi: 10.1104/pp.108.131912

Ciolkowski I., Wanke D., Birkenbihl R. P., & Somssich I. E. (2008). Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant molecular biology*, *68*(1), 81-92.

Coen E. S., and Meyerowitz E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353, 31–37. Doi: 10.1038/353031a0

Coito J. L., Silva H., Ramos M. J. N., Montez M., Cunha J., Amâncio S., et al. (2018). Vitis flower sex specification acts downstream and independently of the ABCDE model genes. Front. Plant Sci. 9:1029. Doi: 10.3389/fpls.2018.01029

Costantini L., Malacarne G., Lorenzi S., Troggio M., Mattivi F., Moser C., et al. (2015). New candidate genes for the fine regulation of the colour of grapes. J. Exp. Bot. 66, 4427–4440. doi: 10.1093/jxb/erv159

Cui F., Brosché M., Sipari N., Tang S., Overmyer K. J. Regulation of ABA dependent wound induced spreading cell death by myb 108. New Phytol. 2013, 200, 634–640.

Dai Z. W., Ollat N., Gomès E., Decroocq S., Tandonnet J.-P., Bordenave L., et al. (2011). Ecophysiological, Genetic, and Molecular Causes of Variation in Grape Berry Weight and Composition: A Review. Am. J. Enol. Vitic. 62, 413–425. Doi: 10.5344/ajev.2011.10116

Davletova S., Rizhsky L., Liang H. J., Zhong S. Q., Oliver D. J., Coutu J., et al. (2005). Cytosolic ascorbate peroxidase 1 is a central component of the reactive oxygen gene network of *Arabidopsis*. *Plant Cell* 17, 268–281. Doi: 10.1105/tpc.104.026971

de Zélicourt A., Diet A., Marion J., Laffont C., Ariel F., Moison M., et al. (2012). Dual involvement of a *Medicago truncatula* NAC transcription factor in root abiotic stress response and symbiotic nodule senescence. *Plant J*. 70, 220–230. Doi: 10.1111/j.1365-313X.2011.04859.x

De Ollas C., Morillon R., Fotopoulos V., Puértolas J., Ollitrault P., Gómez-Cadenas A., et al. (2019). Facing climate change: biotechnology of iconic Mediterranean woody crops. Front. Plant Sci. 10:427. Doi: 10.3389/fpls.2019.00427

Delessert C., Kazan K., Wilson I. W., Van Der Straeten D., Manners J., Dennis E. S., et al. (2005). The transcriptionfactor ATAF2 represses the expression of pathogenesis-related genes in Arabidopsis. *Plant J*. 43, 745–757. Doi: 10.1111/j.1365-313X.2005.02488.x

Delrot S., Grimplet J., Carbonell-Bejerano P., Schwandner A., Bert P.-F., Bavaresco L., Costa L.D., Di Gaspero G., Duchêne E., Hausmann L., Malnoy M., Morgante M., Ollat N., Pecile M., Vezzulli S. Genetic and Genomic Approaches for Adaptation of Grapevine to Climate Change. In Genomic Designing of Climate-Smart Fruit Crops. Springer International Publishing (2020), pp. 157-270

Dilkes B. P., Spielman M., Weizbauer R., Watson B., Burkart-Waco D., et al. (2008). The maternally expressed WRKY transcription factor TTG2 controls lethality in interploidy crosses of Arabidopsis. PLoS Biology 6:2707−20

Distelfeld A., Pearce S. P., Avni R., Scherer B., Uauy C., Piston F. et al. (2012). Divergent functions of orthologous NAC transcription factors in wheat and rice. *Plant Mol. Biol*. 78, 515–524. Doi: 10.1007/s11103-012-9881-6

Dubos C., Stracke R., Grotewold E., Weisshaar B., Martin C., Lepiniec L. J. Myb transcription factors in 31rabidopsis. Trends Plant Sci. 2010, 15, 573–581.

Eulgem T. (2006) Dissecting the WRKY web of plant defense regulators. PloS Pathog. 2, e126

Eulgem T., Rushton P. J., Robatzek S., & Somssich I. E. (2000). The WRKY superfamily of plant transcription factors. Trends in plant science, 5(5), 199-206.

Eulgem T. and Somssich I.E. (2007) Networks of WRKY transcription factors in defense signaling. Curr. Opin. Plant Biol. 10, 366–371

Fang L., Su L., Sun X., Li X., Sun M., Karungo S. K., Fang S., Chu J., Li S., Xin H. (2016). Expression of Vitis amurensis NAC26 in Arabidopsis enhances drought tolerance by modulating jasmonic acid synthesis. J. Exp. Bot. 67, 2829–2845.

Favaro R., Pinyopich A., Battaglia R., Kooiker M., Borghi L., Ditta G., et al. (2003). MADS-box protein complexes control carpel and ovule development in Arabidopsis. *Plant Cell* 15, 2603–2611. Doi: 10.1105/tpc.015123

Feller A., Machemer K., Braun E. L., Grotewold E. J. Evolutionary and comparative analysis of myb and bhlh plant transcription factors. Plant J. 2011, 66, 94–116.

Fornara F., Marziani G., Mizzi L., Kater M., and Colombo L. (2003). MADS-box genes controlling flower development in rice. *Plant Biol.* 5, 16–22. Doi: 10.1055/s-2003-37975

Foroni F., Vignando M., Aiello M., Parma V., Paoletti M. G., Squartini A., et al. (2017). The smell of terroir! Olfactory discrimination between wines of different grape variety and different terroir. Food Qual. Prefer. 58, 18–23. Doi: 10.1016/j.foodqual.2016.12.012

Fujita Y., Fujita M., Shinozaki K., Yamaguchi-Shinozaki K. J. ABA-mediated transcriptional regulation in response to osmotic stress in plants. J. Plant Res. 2011, 124, 509–525.

Gepstein S., Sabehi G., Carp M. J., Hajouj T., Nesher M. F., Yariv I., et al. (2003). Large-scale identification of leaf senescence-associated genes. *Plant J*. 36, 629–642. Doi: 10.1046/j.1365-313X.2003.01908.x

Gerrath J. M. 1993. Developmental morphology and anatomy of grape flowers. Hortic. Rev 13:315-337

Gonzalez A., Zhao M., Leavitt J. M., Lloyd A. M., J. Regulation of the anthocyanin biosynthetic pathway by the ttg1/bhlh/myb transcriptional complex in 32rabidopsis seedlings. Plant J. 2008, 53, 814–827.

Gordo O. and Sanz J. J. (2010). Impact of climate change on plant phenology in 32rabidopsis32n ecosystems. Glob. Chang. Biol. 16, 1082–1106. Doi: 10.1111/j.1365-2486.2009.02084.x

Guillaumie S., Mzid R., Méchin V., Léon C., Hichri I., et al. (2010). The grapevine transcription factor WRKY2 influences the lignin pathway and xylem development in tobacco. Plant Molecular Biology 72:215−34

Guo Y., Cai Z. and Gan S. (2004). Transcriptome of Arabidopsis leaf senescence. *Plant Cell Environ*. 27, 521–549. Doi: 10.1111/j.1365-3040.2003.01158.x

Guo Y. and Gan S. (2006). AtNAP, a NAC family transcription factor, has an important role in leaf senescence. *Plant J*. 46, 601–612. Doi: 10.1111/j.1365-313X.2006.02723.x

Gusmao E. G., Dieterich C., Zenke M., Costa I. G. (2014). Detection of active transcription factor binding sites with the combination of DNase hypersensitivity and histone modifications. Bioinformatics. 30(22):3143-51. doi: 10.1093/bioinformatics/btu519.

Greve K., La Cour T., Jensen M. K., Poulsen F. M. and Skriver K. (2003). Interactions between plant RING-H2 and plant specific NAC (NAM/ATAF1/2/CUC2) proteins: RING-H2 molecular specificity and cellular localization. *Biochem. J*. 371, 97–108. Doi: 10.1042/BJ20021123

Hao J., Ma Q., Hou L., Zhao F., Liu X. (2017). VvWRKY13 enhances ABA biosynthesis in Vitis vinifera. Acta Societatis Botanicorum Poloniae 86:3546

Heazlewood J. E., and S. Wilson. 2004. Anthesis, pollination and fruitset in Pinot noir. Vitis 43:65-68.

He C., da Silva J. A. T., Wang H., Si C., Zhang M., Zhang X., Li M., Tan J., Duan J. J. Mining myb transcription factors from the genomes of orchids (phalaenopsis and dendrobium) and characterization of an orchid r2r3-myb gene involved in water-soluble polysaccharide biosynthesis. Sci. Rep. 2019, 9, 1–19.

He X. J., Mu R. L., Cao W. H., Zhang Z. G., Zhang J. S., and Chen S. Y. (2005). AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. *Plant J*. 44, 903–916. Doi: 10.1111/j.1365-313X.2005.02575.x

Howe K. M., Reakers C. F. L. and Watson R. J. (1990) EMBO J. 9,161-169

Hu H., Dai M., Yao J., Xiao B., Li X., Zhang Q., et al. (2006). Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proc. Natl. Acad. Sci. U.S.A*. 103, 12987–12992. Doi: 10.1073/pnas.0604882103

Hu, R., Qi, G., Kong, Y., Kong, D., Gao, Q., and Zhou, G. (2010). Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. *BMC Plant Biol*. 10:145. Doi: 10.1186/1471-2229-10-145

Huang T., Yu D., Wang X. (2021). *VvWRKY22* transcription factor interacts with *VvSnRK1.1/VvSnRK1.2* and regulates sugar accumulation in grape. Biochemical and Biophysical Research Communications 554:193−98

Huang Y. F., Vialet S., Guiraud J. L. et al. (2014). A negative MYB regulator of proanthocyanidin accumulation, identified through expression quantitative locus mapping in the grape berry,New Phytol., 201, 795–809.

IPCC. 2018. Global Warming of 1.5 °C – an IPCC Special Report on the Impacts of Global Warming of 1.5 °C above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty, available at http://www.ipcc.ch/report/sr15/.

International Human Genome Sequencing Consortium Initial sequencing and analysis of the human genome. Nature. 2001;4091(1):860–921.

Ishiguro S. and Nakamura K. (1994) Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 50 upstream regions of genes coding for sporamin and beta-amylase from sweet potato. Mol. Gen. Genet. 244, 563–571.

Iyer V. R., Horak C. E., Scafe C. S., Botstein D., Snyder M. & Brown P. O. (2001). Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. Nature, 409(6819), 533-538.

Jack T. (2004). Molecular and genetic mechanisms of floral control. *Plant Cell* 16, S1–S17. Doi: 10.1105/tpc.017038.S2

Jensen M. K., Hagedorn P. H., de Torres-Zabala M., Grant M. R., Rung J. H., Collinge D. B., et al. (2008). Transcriptional regulationbyanNAC (NAMATAF1,2-CUC2) transcription factor attenuates ABA signaling for efficient basal defence towards *Blumeria graminis* f sp *hordei* in Arabidopsis. *Plant J*. 56, 867–880. Doi: 10.1111/j.1365-313X.2008.03646.x

Jensen M. K., Kjaersgaard T., Nielsen M. M., Galberg P., Petersen K., O'shea C. & Skriver K. (2010). The *Arabidopsis thaliana* NAC transcription factor family: structure–function relationships and determinants of ANAC019 stress signalling. *Biochemical Journal*, *426*(2), 183-196.

Jensen, M. K., Rung, J. H., Gregersen, P. L., Gjetting, T., Fuglsang, A. T., Hansen, M., et al. (2007). The HvNAC6 transcription factor: a positive regulator of penetration resistance in barley and Arabidopsis. *Plant Mol. Biol*. 65, 137–150. Doi: 10.1007/s11103-007-9204-5

Jeong R. D., Chandra-Shekara A. C., Kachroo A., Klessig D. F. and Kachroo P. (2008). HRT-mediated hypersensitive response and resistance to Turnip crinkle virus in Arabidopsis does not require the function of TIP, the presumed guardee protein. *Mol. Plant Microbe Interact*. 21, 1316–1324. Doi: 10.1094/MPMI-21-10-1316

Jiang C., Gu J., Chopra S., Gu X., Peterson T. J. Ordered origin of the typical two-and three-repeat myb genes. Gene 2004, 326, 13–22.

Jiang Y. and Deyholos M. (2009) Functional characterization of Arabidopsis NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. Plant Mol. Biol. 69, 91–105

Jing S., Zhou X., Song Y. & Yu D. (2009). Heterologous expression of OsWRKY23 gene enhances pathogen defense and dark-induced leaf senescence in Arabidopsis. *Plant Growth Regulation*, *58*(2), 181-190.

Johnson D. S., Mortazavi A., Myers R. M., & Wold B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science, 316(5830), 1497-1502.

Juven-Gershon T, Kadonaga J. T. (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol 339:225–229

Kadonaga J. T. (2012) Perspectives on the RNA polymerase II core promoter. Wiley Interdiscip Rev Dev Biol 1(1):40–51

Katiyar A., Smita S., Lenka S. K., Rajwanshi R., Chinnusamy V., Bansal K. C. J. Genome-wide classification and expression analysis of myb transcription factor families in rice and 34rabidopsis. BMC Genom. 2012, 13, 544.

Keller M. (2010). Managing grapevines to optimise fruit development in a challenging environment: a climate change primer for viticulturists. Aust. J. Grape Wine Res. 16, 56–69. Doi: 10.1111/j.1755-0238.2009.00077.x

Kim S. Y., Kim S. G., Kim Y. S., Seo P. J., Bae M., Yoon H. K., et al. (2007b). Exploring membrane-associated NAC transcription factors in Arabidopsis: implications for membrane biology in genome regulation. *Nucleic Acids Res*. 35, 203–213.

Klug W. S. and Cummings M. R. (2003) Concepts of genetics. Prentice Hall, Upper Saddle River

Korres N. E., Norsworthy J. K., Tehranchian P., Gitsopoulos T. K., Loka D. A., Oosterhuis D. M., et al. (2016). Cultivars to face climate change effects on crops and weeds: a review. Agron. Sustain. Dev. 36, 1–22. Doi: 10.1007/s13593-016-0350-5

Koul B., Yadav R., Sanyal I., Sawant S., Sharma V., Amla D. V. (2012) Cis-acting motifs in artificially synthesized expression 34rabidop leads to enhanced transgene expression in tomato (Solanum lycopersicum L.) Plant Physiol Biochem 61:131–141

Kozma P. (2003). Exploration of flower types in grapes. In Floral Biology, Pollination and Fertilisation in Temperate Zone Fruit Species and Grape. P. Kozma et al. (eds.), pp. 75-226. Akademiai Kiado, Budapest

Kwak H. and Lis J. T. (2013) Control of transcriptional elongation. Annu Rev Renet 47:483

Langcake P., Pryce RJ. (1976). The production of resveratrol by Vitis vinifera and other members of the Vitaceae as a response to infection or injury. Physiological Plant Pathology 9:77−86

Le D. T., Nishiyama R., Watanabe Y., Mochida K., Yamaguchi-Shinozaki K., Shinozaki K., et al. (2011). Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress. *DNA Res*. 18, 263–276. Doi: 10.1093/dnares/dsr015

Le Henanff G., Profizi C., Courteaux B., Rabenoelina F., Gerard C., Clement C., Baillieul F., Cordelier S., Dhondt-Cordelier S. (2013). Grapevine NAC1 transcription factor as a convergent node in developmental processes, abiotic stresses, and necrotrophic/biotrophic pathogen tolerance. J. Exp. Bot. 64, 4877–4893.

Lee S., Seo P. J., Lee H. J. and Park C. M. (2012). A NAC transcription factor NTL4 promotes reactive oxygen species production during drought-induced leaf senescence in Arabidopsis. *Plant J*. 70, 831–844. Doi: 10.1111/j.1365-313X.2012.04932.x

Lepiniec L., Debeaujon I., Routaboul J.-M., Baudry A., Pourcel L., Nesi N., Caboch, M. J. (2006). Genetics and biochemistry of seed flavonoids. Ann. Rev. Plant Biol. 57, 405–430.

Lewin B. (2001) Genes VII. Artmed Editora, Porto Alegre

Li Z., Schulz M. H., Look T. et al. (2019). Identification of transcription factor binding sites using ATAC-seq. Genome Biol 20, 45. https://doi.org/10.1186/s13059-019-1642-2

Lieb J. D., Liu X., Botstein D. & Brown P. O. (2001). Promoter-specific binding of Rap1 revealed by genome-wide maps of protein–DNA association. Nature genetics, 28(4), 327-334.

Lipsick J. S. J. (1996). One billion years of myb. Oncogene, 13, 223–235.

Liu Q. L., Xu K. D., Zhao L. J., Pan Y. Z., Jiang B. B., Zhang H. Q., & Liu G. L. (2011). Overexpression of a novel chrysanthemum NAC transcription factor gene enhances salt tolerance in tobacco. *Biotechnology letters*, *33*(10), 2073-2082.

Liu S. and Bao Y. (2009) Effects of copy number of an octopine synthase enhancer element and its distance from the TATA box on heterologous expression in transgenic tobacco. Acta Phys Plant 31:705–710

Luo M., Dennis E. S., Berger F., Peacock W. J., & Chaudhury A. (2005). MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in Arabidopsis. *Proceedings of the National Academy of Sciences*, *102*(48), 17531-17536.

Luscombe N. M., Austin S. E., Berman H. M., Thornton J. M. (2000). An overview of the structures of protein-DNA complexes. 1 Genome Biol. Reviews001.001-001.037.

Ma Q., Zhang G., Hou L., Wang W., Hao J., et al. (2015). *Vitis vinifera VvWRKY13* is an ethylene biosynthesis-related transcription factor. Plant Cell Reports 34:1593−603

Martin C. & Paz-Ares J. (1997). MYB transcription factors in plants. Trends in Genetics, 13(2), 67-73.

Martin D. M., Toub, O., Chiang A., Lo B.C., Ohse S., Lund S.T. and Bohlmann J. (2009) 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. USA, 106, 7245– 7250.

Matus J. T., Aquea F., Arce-Johnson P. (2008). Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes,BMC Plant Biol., 8, 83.

May P. (2004). Flowering and Fruitset in Grapevines. Lythrum Press, Adelaide

Medlyn, B. E. (2011). Comment on “Drought-induced reduction in global terrestrial net primary production from 2000 through 2009”. Science 333:1093. Doi: 10.1126/science.1199544

Meneghetti S., M. Gardiman and A. Calo. (2006). Flower biology of grapevine. A review. Adv. Hortic. Sci. 20:317-325.

Mikkelsen T. S., Ku M., Jaffe D. B., Issac B., Lieberman E., Giannoukos G., ... & Bernstein B. E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature, 448(7153), 553-560.

Morse A. M., Whetten R.W., Dubos C., Campbell M. M. J. (2009). Post-translational modification of an r2r3-myb transcription factor by a map kinase during xylem development. New Phytol. 183, 1001–1013.

Nakashima K., Tran L. S., VanNguyen D., Fujita M., Maruyama K., Todaka D., et al. (2007). Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. *Plant J*. 51, 617–630. Doi: 10.1111/j.1365-313X.2007.03168.x

Nuruzzaman M., Manimekalai R., Sharoni A. M., Satoh K., Kondoh H., Ooka H., et al. (2010). Genome-wide analysis of NAC transcription factor family in rice. *Gene* 465, 30–44. Doi: 10.1016/j.gene.2010.06.008

Nuruzzaman M., Sharoni A. M. & Kikuchi S. (2013). Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Frontiers in microbiology*, *4*, 248.

Nuruzzaman M., Sharoni A. M., Satoh K., Kondoh H., Hosaka A., and Kikuchi S. (2012a). “A genome-wide survey of the NAC transcription factor family in monocots and eudicots,” in *Introduction to Genetics – DNA Methylation, Histone Modification and Gene Regulation* (Hong Kong: iConcept Press), ISBN, 978-14775549-4-4

Oberle G. nD. (1938). A genetic study of variations in floral morphology and function in cultivated forms of Vitis. New York State Agric. Exp. Station Tech. Bulletin 250:416-420

Ochiai K., Shimizu A., Okumoto Y., Fujiwara T., and Matoh T. (2011). Suppression of a NAC-like transcription factor gene improves boron-toxicity tolerance in rice. *Plant Physiol*. 156, 1457–1463. Doi: 10.1104/pp.110.171470

OIV (2016). International Organization of Vine and Wine. Available at: http://www.oiv.int/en/the-international-organisation-of-vine-and-wine

Ong C. T. and Corces V. G. (2011). Enhancer function: new insights into the regulation of tissuespecific gene expression. Nat. Rev. Genet. 1

Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., et al. (2003). Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Res*. 10, 239–247. Doi: 10.1093/dnares/10.6.239

Palumbo F., Vannozzi A., Magon G., Lucchin M., & Barcaccia G. (2019). Genomics of flower identity in grapevine. Vitis vinifera, 1-15.

Park P. J. (2009). ChIP–seq: advantages and challenges of a maturing technology. Nature reviews genetics, 10(10), 669-680.

Peremarti A., Twyman R. M., Gómez-Galera S., Naqvi S., Farré G., Sabalza M., Miralpeix B., Dashevskaya S., Yuan D., Ramessar K., Christou P., Zhu C., Bassie L., Capell T. (2010). Promoter diversity in multigene transformation. Plant Mol Biol 73:363–378

Perin C., Fait A., Palumbo F., Lucchin M., Vannozzi A. (2020). The Effect of Soil on the Biochemical Plasticity of Berry Skin in Two Italian Grapevine (*V. vinifera* L.) Cultivars Front. Plant Sci., 11, 822 (2020) http://doi.org/10.3389/fpls.2020.00822

Phillips T. (2008) Regulation of transcription and gene expression in eukaryotes. Nat Educ 1(1):199

Porto M. S., Pinheiro M. P. N., Batista V. G. L., Cavalcanti dos Santos R., de Albuquerque Melo Filho P., de Lima L. M. (2014) Plant promoters: an approach of structure and function. Mol Biotechnol 56:38–49

Pratt C. (1971). Reproductive anatomy in cultivated grapes–A review. Am. J. Enol. Vitic. 22:92-109

Puente P., Wei N., Deng X. W. (1996). Combinatorial interplay of promoter elements constitutes the minimal determinants for light and developmental control of gene expression in Arabidopsis. EMBO J. 151(1):3732–3743.

Qiu Y. & Yu D. (2009). Over-expression of the stress induced OsWRKY45 enhances disease resistance and drought tolerance in Arabidopsis. *Environmental and experimental botany*, *65*(1), 35-47.

Qu J., Dry I., Liu L., Guo Z., Yin L. (2021). Transcriptional profiling reveals multiple defense responses in downy mildew-resistant transgenic grapevine expressing a TIR-NBS-LRR gene located at the MrRUN1/MrRPV1 locus. Horticulture Research 8:161

Ren B., Robert F., Wyrick J. J., Aparicio O., Jennings E. G., Simon I., ... & Young R. A. (2000). Genome-wide location and function of DNA binding proteins. science, 290(5500), 2306-2309.

Reyes J. L., Chua N. H. J. (2007). ABA induction of mir159 controls transcript levels of two myb factors during 37rabidopsis seed germination. Plant J. 49, 592–606.

Riechmann J. L. (2002). Transcriptional regulation: a genomic overview. The *Arabidopsis* Book/American Society of Plant Biologists, 1.

Robatzek S. and Somssich I. E. (2002) Targets of AtWRKY6 regulation during plant senescence and pathogen defense. Genes Dev. 16, 1139– 1149

Robertson G., Hirst M., Bainbridge M., Bilenky M., Zhao Y., Zeng T., ... & Jones S. (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nature methods, 4(8), 651-657.

Rosinski J. A. and Atchley W. R. J. (1998). Molecular evolution of the myb family of transcription factors: Evidence for polyphyletic origin. J. Mol. Evol. 46, 74–83.

Rushton P. J., Bokowiec M. T., Han S., Zhang H., Brannock J. F., Chen X., … & Timko M. P. (2008). Tobacco transcription factors: novel insights into transcriptional regulation in the Solanaceae. *Plant physiology*, *147*(1), 280-295.

Rushton P. J., Somssich I. E., Ringler P. & Shen Q. J. (2010). WRKY transcription factors. Trends in plant science, 15(5), 247-258.

Rushton P. J., Torres J. T., Parniske M., Wernert P., Hahlbrock K., & Somssich I. E. (1996). Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *The EMBO journal*, *15*(20), 5690-5700.

Russell P. J. (1996) Genetics, 4th edn. Harper Collins College, New York

Sakura H., Kanei-Ishii C., Nagase T., Nakagoshi H., Gonda T. J. & Ishii S. (1989). Delineation of three functional domains of the transcriptional activator encoded by the c-myb protooncogene. Proceedings of the National Academy of Sciences, 86(15), 5758-5762.

Schnee S., Viret O., Gindro K. (2008). Role of stilbenes in the resistance of grapevine to powdery mildew. Physiological and Molecular Plant Pathology 72:128−33

Scott M. P. (2000). Development: the natural history of genes. Cell. 1001(1):27–40.

Shelton D., Stranne M., Mikkelsen L., Pakseresht N., Welham T., Hiraka H., Tabata S., Sato S., Paquette S., Wang T. L. J. (2012). Transcription factors of lotus: Regulation of isoflavonoid biosynthesis requires coordinated changes in transcription factor activity. Plant Physiol. 159, 531–547.

Sindhu A., Chintamanani S., Brandt A. S., Zanis M., Scofield S. R., and Johal G. S. (2008). A guardian of grasses: specific origin and conservation of a unique disease-resistance gene in the grass lineage. *Proc. Natl. Acad. Sci. U.S.A*. 105, 1762–1767. Doi: 10.1073/pnas.0711406105

Song L. & Crawford G. E. (2010). DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. Cold Spring Harbor Protocols, 2010(2), pdb-prot5384.

Souer E., van Houwelingen A., Kloos D., Mol J., and Koes R. (1996). The no apical meristem gene of Petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell* 85, 159–170. Doi: 10.1016/S0092-8674(00)81093-4

Spitz F. and Furlong E. E. (2012). Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. 13: 613–626.

Staudt G. (1999). Opening of flowers and time of anthesis in grapevines, Vitis vinifera L. Vitis 38:15-20.

Stracke R., Werber M., Weisshaar B. J. (2001). The r2r3-myb gene family in *Arabidopsis thaliana*. Curr. Opin. Plant Biol. 4, 447–456.

Sultan, S. E. (2000). Phenotypic plasticity for plant development, function and life history. Trends Plant Sci. 5, 537–542. Doi: 10.1016/S1360-1385(00)01797-0

Szathmáry E., Jordán F., Pál C. (2001). Can genes explain biological complexity? Science. 2921(1):1315–1316.

Tello J., Torres-Perez R., Grimplet J., Carbonell-Bejerano P., Martinez-Zapater J. M., Ibanez J. (2015). Polymorphisms and minihaplotypes in the VvNAC26 gene associate with berry size variation in grapevine. BMC Plant Biol. 15, 253.

Tran L. S., Nakashima K., Sakuma Y., Simpson S. D., Fujita Y. and Maruyama K. (2004). Isolation and functional analysis of Arabidopsis stress inducible NAC transcription factors that bind to a drought responsive *cis*-element in the early responsive to dehydration stress 1 promoter. *Plant Cell* 16, 2481–2498. Doi: 10.1105/tpc.104.022699

Twyman R. M. (2003) Growth and development: control of gene expression, regulation of transcription. In: Thomas B, Murphy DJ, Murray B (eds) Encyclopedia of applied plant sciences. Elsevier Science, London, pp 558–567

Ülker B., Shahid Mukhtar M. & Somssich I. E. (2007). The WRKY70 transcription factor of *Arabidopsis* influences both the plant senescence and defense signaling pathways. *Planta*, *226*(1), 125-137.

Van Leeuwen C., Seguin G. (2006). The concept of terroir in viticulture. J. Wine Res. 17, 1–10. Doi: 10.1080/09571260600633135

Vandenbussche M., Zethof J., Souer E., Koes R., Tornielli G. B., Pezzotti M., et al. (2003). Toward the analysis of the petunia MADS box gene family by reverse toward the analysis of the petunia MADS box gene family by and forward transposon insertion mutagenesis approaches: B, C, and D floral organ identity functions require SEPALLATA-like MA. *Plant Cell* 15, 2680–2693. Doi: 10.1105/tpc.017376

Vannozzi A., Wong D. C. J., Höll J., Hmmam I., Matus J. T., et al. (2018). Combinatorial Regulation of Stilbene Synthase Genes by WRKY and MYB Transcription Factors in Grapevine (Vitis vinifera L.). Plant and Cell Physiology 59:1043−59

Vasconcelos M. C., Greven M., Winefield C. S., Trought M. C. T. and Raw V. (2009). The flowering process of *Vitis vinifera*: a review. Am. J. Enol. Vitic. 60, 411–434. Doi: 10.1104/pp.002428.68

Vaughn J. N., Ellingson S. R., Mignone F., von Arnim A. G. (2012) Known and novel post-transcriptional regulatory sequences are conserved across plant families. RNA 18:368–384

Venter M. and Botha F. C. (2010) Synthetic promoter engineering. In: Pua EC, Davey MR (eds) Plant developmental biology—biotechnological perspectives. Springer, Berlin/Heidelberg, pp 393–414

Wang D., Jiang C., Liu W., Wang Y. (2020). The WRKY53 transcription factor enhances stilbene synthesis and disease resistance by interacting with MYB14 and MYB15 in Chinese wild grape. Journal of Experimental Botany 71(10):3211−26

Wang N., Zheng Y., Xin H. et al. (2013). Comprehensive analysis of NAC domain transcription factor gene family in *Vitis vinifera*. Plant Cell Rep 32, 61–75. https://doi.org/10.1007/s00299-012-1340-y

Wang X., Basnayake B. M., Zhang H., Li G., Li W., Virk N., et al. (2009a). The Arabidopsis ATAF1, a NAC transcription factor, is a negative regulator of defense responses against necrotrophic fungal and bacterial pathogens. *Mol. Plant Microbe Interact*. 22, 1227–1238.

Wang X., Goregaoker S. P., and Culver J. N. (2009b). Interaction of the *Tobacco mosaic virus* replicase protein with a NAC domain transcription factor is associated with the suppression of systemic host defenses. *J. Virol*. 83, 9720–9730.

Weston, K.; Bishop J. M. J. Transcriptional activation by the v-myb oncogene and its cellular progenitor, c-myb. Cell 1989, 58, 85–93.

Wong D. C. J., Schlechter R., Vannozzi A., Höll J., Hmmam I., Bogs J., Tornielli G. B., Castellarin S. D., Matus J. T. A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family uncovers new insights into the regulation of stilbene accumulation, DNA Research, Volume 23, Issue 5, October 2016, Pages 451–466, https://doi.org/10.1093/dnares/dsw028

Wu W., Fu P., Lu J. (2022). Grapevine WRKY transcription factors. Fruit Research 2:10 doi: 10.48130/FruRes-2022-0010

Wu X., Shiroto Y., Kishitani S., Ito Y. & Toriyama K. (2009). Enhanced heat and drought tolerance in transgenic rice seedlings overexpressing OsWRKY11 under the control of HSP101 promoter. *Plant cell reports*, *28*(1), 21-30.

Xia N., Zhang G., Liu X. Y., Deng L., Cai G. L., Zhang Y., et al. (2010a). Characterization of a novel wheat NAC transcription factor gene involved in defense response against stripe rust pathogen infection and abiotic stresses. *Mol. Biol. Rep*. 37, 3703–3712.

Xie Q., Frugis G., Colgan D., and Chua N. (2000). Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes Dev*. 14, 3024–3036. Doi: 10.1101/gad.852200

Xie Z., Zhang Z. L., Zou X., Yang G., Komatsu S. & Shen Q. J. (2006). Interactions of two abscisic-acid induced WRKY genes in repressing gibberellin signaling in aleurone cells. *The Plant Journal*, *46*(2), 231-242.

Yamaguchi M., Kubo M., Fukuda H. and Demura T. (2008). Vascular related NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and shoots. *Plant J*. 55, 652–664. Doi: 10.1111/j.1365-313X.2008.03533.x

Yamasaki K., Kigawa T., Inoue M., Tateno M., Yamasaki T., Yabuki T., … & Yokoyama S. (2005). Solution structure of an Arabidopsis WRKY DNA binding domain. The Plant Cell, 17(3), 944-956.

You J., Zong W., Li X., Ning J., Hu H., Li X., et al. (2013). The SNAC1-targeted gene OsSRO1c modulates stomatal closure and oxidative stress tolerance by regulating hydrogen peroxide in rice. *J. Exp. Bot*. 64, 569–583. Doi: 10.1093/jxb/ers349

Yuh C. H., Bolouri H., Davidson E. H. Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. Science. 1998;2791(1):1896–1902.

Zandalinas S. I., Mittler R., Balfagon D., Arbona V., Gomez-Cadenas A., Balfagón D., et al. (2017). Plant adaptations to the combination of drought and high temperatures. Physiol. Plant. 162, 2–12. Doi: 10.1111/ppl.12540

Zhang Y. and Wang L. (2005) The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. BMC Evol. Biol. 5, 1

Zhou L., He Y., Li J., Liu Y., Chen H. Cbfs function in anthocyanin biosynthesis by interacting with myb113 in eggplant (Solanum melongena L). Plant Cell Physiol. 2020, 61, 416–426.

Zhou M., Li D., Li Z., Hu Q., Yang C., Zhu L., et al. (2013). Constitutive expression of a miR319 gene alters plant development and enhances salt and drought tolerance in transgenic creeping bentgrass. *Plant Physiol*. 161, 1375–1391. Doi: 10.1104/pp.112.208702

Zhu Z., Li G., Yan C., Liu L., Zhang Q., Han Z., & Li B. (2019). DRL1, encoding a NAC transcription factor, is involved in leaf senescence in grapevine. International Journal of Molecular Sciences, 20(11), 2678.

Paper published in Horticulture Research (https://doi.org/10.1038/s41438-021-00635-7)

CHAPTER II: The grapevine (*V. vinifera* L.) floral transcriptome in Pinot noir variety: identification of tissue-related gene networks and whorl specific markers in pre- and post-anthesis phases

Alessandro Vannozzi[1#], Fabio Palumbo[1#], Gabriele Magon[1], Margherita Lucchin[1], Gianni Barcaccia[1]

[1] Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Campus of Agripolis, Viale dell'Università 16, 35020 Legnaro (PD), Italy

[*]Correspondence: gianni.barcaccia@unipd.it

[#]Equal contribution

2.1 Abstract

The comprehension of molecular processes underlying the development and progression of flowering in plants is a hot topic, not only because that often the products of interest for human and animal nutrition are linked to the development of fruits or seeds, but also because the processes of gametes formation occurring in sexual organs are at the basis of recombination and genetic variability which constitutes the matter on which evolution acts, whether understood as natural or human driven. In the present study, we used an NGS approach to produce a grapevine flower transcriptome snapshot in different whorls and tissues including calyx, calyptra, filament, anther, stigma, ovary and embryo in both pre- and post-anthesis phases. Our investigation aimed at identifying hub genes that unequivocally distinguish the different tissues providing insights into the molecular mechanisms that are at the basis of floral whorls and tissue development. To this end we have used different analytical approaches, some now consolidated in transcriptomic studies on plants, such as pairwise comparison and weighted gene co-expression network analysis, others used mainly in studies on animals or human's genomics, such as the tau (t) analysis aimed at isolating highly and absolutely tissue-specific genes. The intersection of data obtained by these analyses allowed us to gradually narrow the field, providing evidence about the molecular mechanisms occurring in those whorls directly involved in reproductive processes, such as anther and stigma, and giving insights into the role of other whorls not directly related to reproduction, such as calyptra and calyx. We believe this work could represent an important genomic resource

for functional analyses of grapevine floral organ growth and fruit development shading light on molecular networks underlying grapevine reproductive organ determination.

Keywords: *WGCNA, anther, filament, embryo, calyx, calyptra, ovary, anthesis, expression atlas*

2.2 Introduction

Grapevine (*V. vinifera* L.) is indisputably one of the most popular crops in the world. According to the most updated data of the International Organization of Vine and Wine, vines are spread over an area of more than 7 million hectares and produce annually about 75 million tons of grapes, destined to the production of wine, table grapes, juices and raisins (OIV). Therefore, in a sector that alone is worth billions of dollars, a deep comprehension of each of the thorny phases characterizing berry development is pivotal and several approaches have been applied to elucidate the changes that take place from flowering to ripe berry.

From a molecular point of view, analyses of large gene expression datasets represent a key tool to decipher the biological processes underlying the development of a specific tissue or organ. This would explain the exponential increase in the number of expression atlases developed in recent years. An expression atlas should be meant as a snapshot of the genes expressed in one or more tissues, organs or even cells at a given phenological stage or throughout a developmental kinetics (Papatheodorou et al., 2018). According to the main databases dedicated to this purpose – the Bio-Analytic Resource (BAR) for plant biology (Provart, 2020) and the Expression Atlas platform (Papatheodorou et al., 2018)– expression atlases are available for at least 29 plant species. Among the monocots species used to produce the most recent and complete atlases – in terms of tissues and time point analyzed – stand out *Zea mays* (Stelpflug et al., 2016), *Sorghum bicolor* (Wang et al., 2016) and *Hordeum vulgare* (Mayer et al., 2012). Instead, for dicots species, in addition to *Arabidopsis thaliana* (Cheng et al., 2017) a remarkable effort was done in *Glycine max* (Severin et al., 2010; Libault et al., 2010), *Solanum lycopersicum* (Sato et al., 2012)*, Solanum tuberosum* (Xu et al., 2011) and *V. vinifera* L. As regards this latter, several RNA-seq-based experiments have been conducted in the last 10 years. Considering the “time” variable, transcriptional profiles in temporal kinetics are available for berry as a whole (Corvina cv.) (Venturini et al., 2013), grape skin (Cabernet Sauvignon cv.) (Wang and Li, 2015) and leaf (Summer Black cv.) (Pervaiz et al., 2016). Considering the “cultivar” variable, the grape berry transcriptomes of ten different cultivars were compared to identify cultivar specific-splicing events (Potenza et al., 2015), while Ghan et al. (2017) applied the same approach in 7 wine grape cultivars to identify the common transcriptional subnetworks underlying the berry skin in the late stages of ripening. Also, the comparison among berry transcriptomic profiles of 5 Italian cultivars, each sampled at 4 progressive phenological phases (Palumbo et al., 2014), led to the identification of common switch genes, that seem to regulate the phase transition during berry ripening. In an exhaustive study, Dal Santo et al. (2018) performed an

RNA-seq analysis in two genotypes (Cabernet Sauvignon and Sangiovese) at two developmental stages and cultivated in three different environments over two vintages, in order to elucidate the contribution of genotype, the influence of environment and the effect of their interaction (G×E) on the berry transcriptome. Finally, the most comprehensive atlas so far produced in grapevine is based on 54 samples representing green and woody organs at different developmental phases and it was developed to infer the specific metabolic pathways characterizing each of the samples (Fasoli et al., 2012).

The abundance of transcriptomic data relating to grape berry and its sub-tissues offers an in-depth insight into the molecular processes underlying berry growth and ripening, but it leaves unclarified most of the upstream aspects related to flower development. In fact, except for Fasoli et al. (2012), who considered some of the floral tissues, transcriptional data straddling the anthesis process and related to single grape whorls are lacking. Since the flower gene expression regulation at both temporal and spatial level is the cornerstone for achieving the specification of morphology and physiology of the berry, we attempted to fill this gap by dissecting the transcriptome profiles of six (calyptra, calyx, anther, filament, stigma and ovary) and four floral tissues (calyx, stigma, ovary, stigma and embryo) before and after anthesis, respectively. Making use of analytical tools such as the weighted gene network co-expression analysis (WGCNA), and the analysis tau (t) and crossing obtained results with already available data, we tried to identify those hub genes and molecular networks that specifically characterize different floral organs or tissues. This analysis made it possible to identify enriched ontological categories in the different tissues under examination and to isolate specific transcription factors expressed exclusively or predominantly in a given whorl. Furthermore, limited to genes expressed exclusively in anther or in stigma, we carried out a *de novo cis* regulatory elements (CRE) analysis at the promoter level, in order to identify those motifs linked to the tissue-specific expression of selected genes. This original grapevine atlas could represent an important genomic resource for functional analyses of grapevine floral organ growth and fruit development.

2.3 Results and Discussion

2.3.1 Global RNA-seq analysis of grapevine flower tissues

To obtain gene expression profiles for different whorls and tissues of the grapevine flower, RNA sequencing (RNA-seq) data were generated from 10 different tissues collected from flowers of *V. vinifera* cv Pinot noir plants grown in open fields. Six tissue samples were collected from pre-anthesis ("Before Anthesis" – BA) flowers (EL-18), namely calyptra, calyx, filament, anther, ovary and stigma, and 4 tissue samples were collected after anthesis (AA; EL-26): calyx, ovary, stigma and embryo (Fig. 1.2A). The HiSeq 2500 sequencing run produced a total output of 336 M of 2 x 250 bp reads (on average 12 M reads per sample), while, after filtering steps, approximately 312 M of reads were retained. Given the low number of reads obtained for Stigma BA and Filament BA, the third replicate from both tissues was excluded from further analyses. The cluster dendrogram analysis based on raw counts (Supplementary Figure 1) showed a good correlation among the biological replicates of each sample, except for one replicate of Calyx post-anthesis (Calyx AA) grouped with the Embryo samples, which may be due to the high proximity of the tissues which reside very close to each other. The filtered reads deriving from the 28 samples were combined and assembled into a reference catalogue, composed by 210,674 transcripts and annotated based on the PN40024 12X v1 (Jaillon et al., 2007). After Transcript Per Million (TPM) calculation, 7,802 genes were filtered out while 22,094 genes, corresponding to the 73.8 % of the total number of genes predicted in the PN40024 12X v1 grapevine reference genome, were retained for further analyses, being expressed in at least one tissue with TPM > 1. Pre-anthesis anther (Anther BA) had the lowest number of expressed genes (15,843), whereas Calyx AA had the highest number (19,320). Finally, 13,087 genes scored a TPM value > 1 in all tissues considered (Fig. 1.2B, Supplementary Table 1).

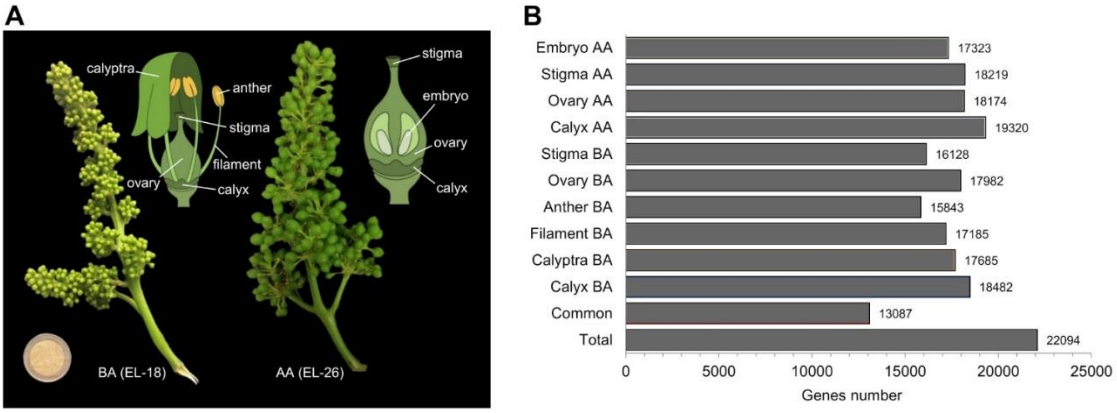


Figure 1.2. Overview of the *V. vinifera* cv Pinot noir floral samples used for RNA-Seq analysis. (A) The pictures show the grapevine inflorescence in pre- and post-anthesis phases, while the schematic illustrations alongside indicate the specific floral tissues sampled for subsequent transcriptional analyses. (B) Number of genes expressed (TPM > 1) in each of the 10 tissue/organs. Total: number of genes expressed in at least one organ (22,094). Common: genes expressed in all 10 organs (13,087).

To isolate tissue-specific genes providing evidence about the molecular mechanisms occurring in different whorls analyzed, we took advantage of different analytical approaches. Some of them are commonly used in plant genomics, such as pairwise comparison and weighted gene co-expression network analysis (WGCNA), some others have been mainly exploited in animals or human genomics, such as the tau (t) analysis. For the sake of clarity in Fig. 2.2, we reported the logical workflow of all analyses performed in this study.

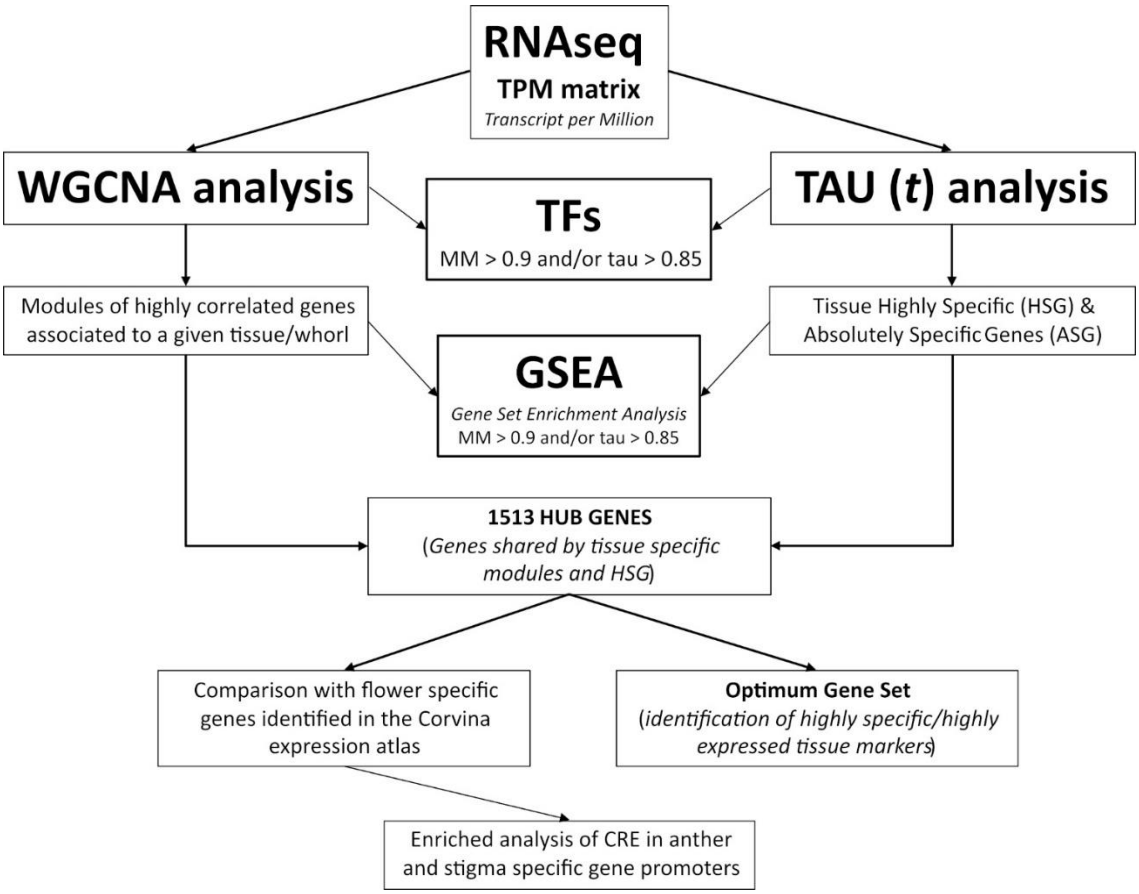


Figure 2.2. Logical workflow of analyses performed in this study. Schematic illustration of the main analyses performed on TPM normalized data.

2.3.2 Weighted Gene Co-expression Network Analysis identified gene modules highly associated with specific grapevine flower whorls/tissues

WGCNA is a systems biology approach aimed at understanding networks of highly correlated genes instead of individual genes, which has been successfully applied in various genomics studies in many plant species including pineapple (Wang et al., 2020), strawberry (Shahan et al., 2018), pear (Li et al., 2019) and grapevine (Guo et al., 2020). In this study, co-expression networks were constructed based on pairwise correlations of gene expression trends across all sampled tissues. The 22,094 genes resulting from TPM normalization and filtering, were analyzed in order to identify gene co-expression modules using the R package WGCNA. Genes showing variance lower than 1 were removed leaving a total number of 19,658 genes. The matrix was raised to a soft-thresholding power β = 12 to ensure a scale-free network (Fig. 3.2A, B). Modules are defined as clusters of highly interconnected genes, such that genes belonging to the same module, highly correlated with each

other. For the present analysis, the minimum module size was set to 30, and modules with highly correlated eigengenes (based on a threshold of 0.25) were merged. The eigengene represents the first principal component of a given module and can be considered as a representative of the expression profiles of genes within that module.

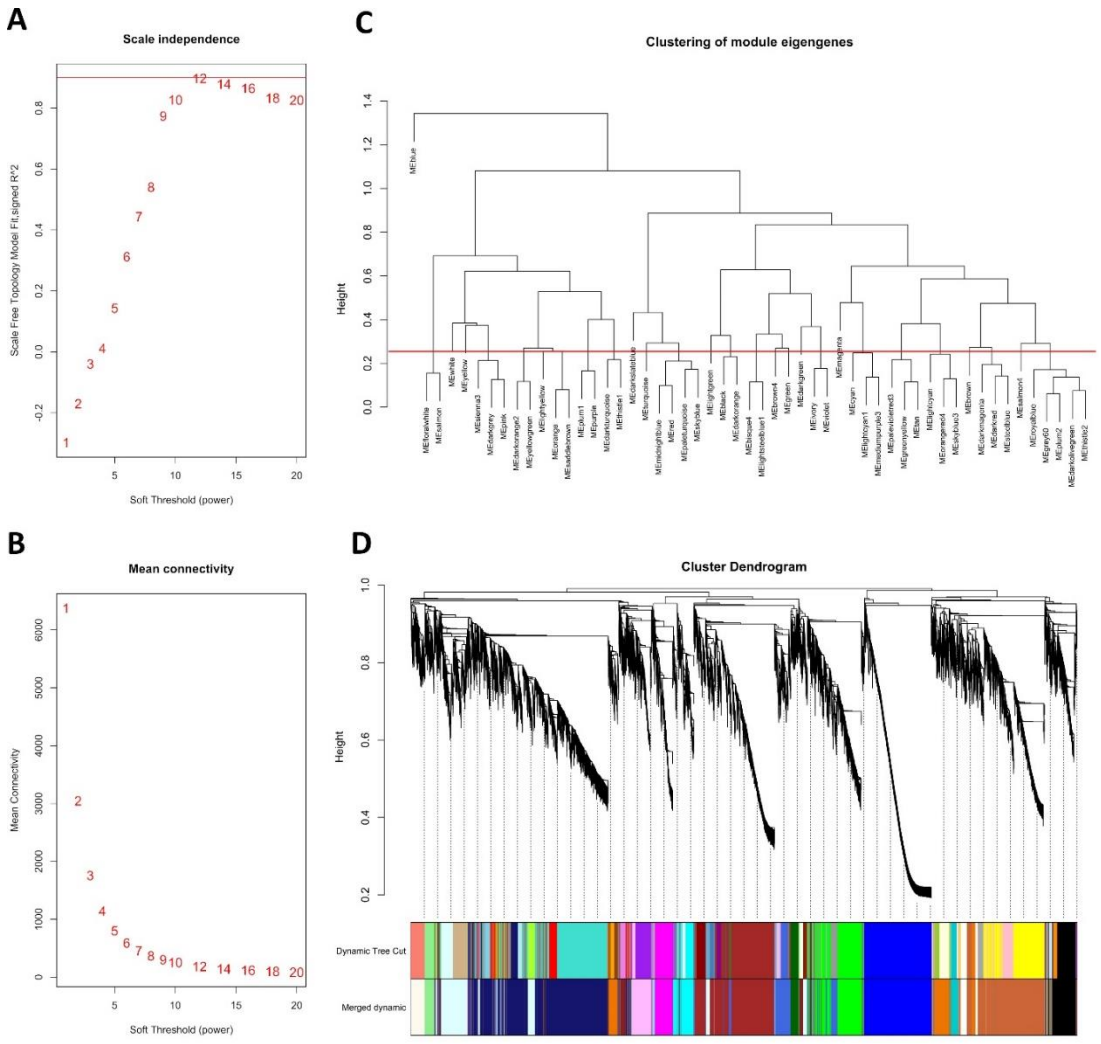


Figure 3.2. Weighted gene co-expression network analysis for grapevine RNA-seq data. Analysis of network topology for various soft-thresholding powers showing (A) the scale-free fit index (y-axis) as a function of the soft-thresholding power (x-axis) and (B) mean connectivity (degree, y-axis) as a function of the soft-thresholding power (x-axis). (C) Cluster dendrogram of module eigengenes. Branches of the dendrogram group together eigengenes that are positively correlated. The red line is the merging threshold, and groups of eigengenes below the threshold represent modules whose expression profiles should be merged due to their similarity. (D) Hierarchical cluster dendrogram showing co-expressed modules identified by weighted gene co-expression network analysis for the grapevine flower RNA-seq data. Each leaf on the tree represents a gene. The major tree branches constitute 20 merged modules (based on a threshold of 0.25), labeled with different colors.

Twenty distinct modules were identified (Fig. 3.2C). The modules are labeled by color and are shown in a hierarchical clustering dendrogram, in which each tree branch constitutes a module and each leaf in the branch is one gene (Fig. 3.2D). Next, we performed a correlation analysis between the 20 distinct modules and the 10 tissues/whorls under study (Fig. 4.2A).

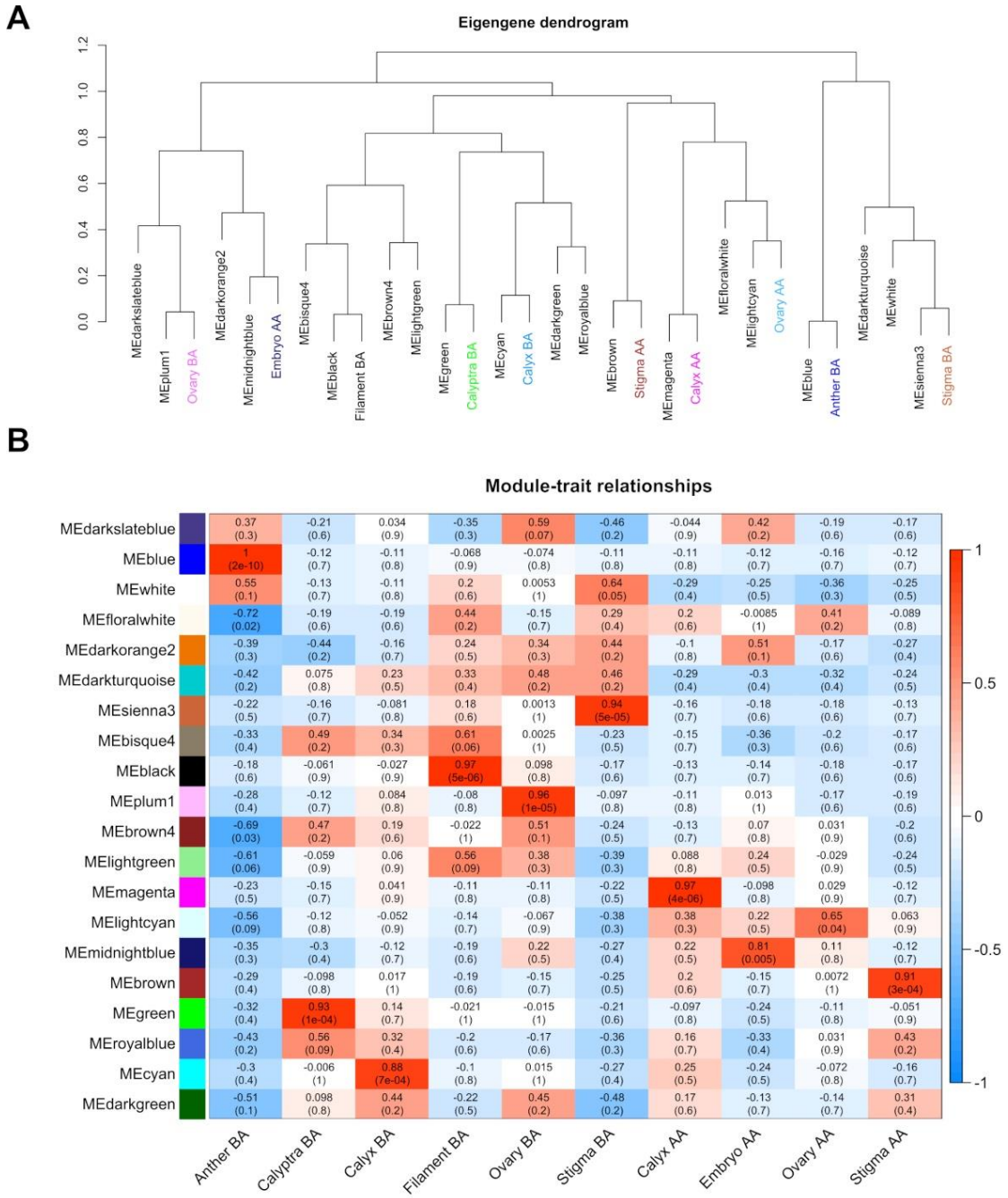


Figure 4.2. Module-tissue association analysis. Visualization of the eigengene network representing the relationships among the modules and the tissues under study. Panel (A) shows a hierarchical clustering dendrogram of the eigengenes in which the dissimilarity of eigengenes E_I, E_J is given by 1−cor (E_I, E_J). The color of the most correlated module was used to color the name of the organ/tissue. Panel (B) heatmap shows the correlation between modules and tissues. Each row corresponds to a module, whereas each column corresponds to a specific tissue. The correlation coefficient between a given module and tissue type is indicated by the color of the cell at the row-column intersection and by text inside cells (p-value is also reported). Red and blue indicate positive and negative correlation, respectively.

For each organ, at least one highly specific module was identified ($r > 0.8$; correlation p-value < 0.01; Fig. 4.2B), although, in some cases, multiple modules showed significant correlations with the same tissue and/or more tissues showed correlations with the same module. For example, the sienna3, cyan, green, and midnightblue modules were specifically correlated with Stigma BA, Calyx BA, Calyptra BA, and Embryo, respectively. The best correlations between module eigengene (ME) and tissue were between the blue module and Anther BA ($r = 1$, $p = 2e\text{-}10$), the black module and Filament BA ($r = 0.97$, $p = 5e\text{-}06$) and finally the magenta module and Calyx AA ($r = 0.97$, $p = 4e\text{-}06$) (Fig. 4.2B, Supplementary Table 2). To further investigate the gene constitution of the 10 modules showing the best correlation with tissues under study, two network unique properties such as gene significance (GS) and module membership (MM) were carried out. The module membership (MM) is a measure of the correlation between the expression profile of a given gene with the considered module eigengene. The gene significance (GS) is an additional network parameter, that can be also defined by the minus log of a p-value and give an estimation of the biological significance of a gene. The higher the absolute value of GS*i*, the more biologically significant is the *i*-th gene.

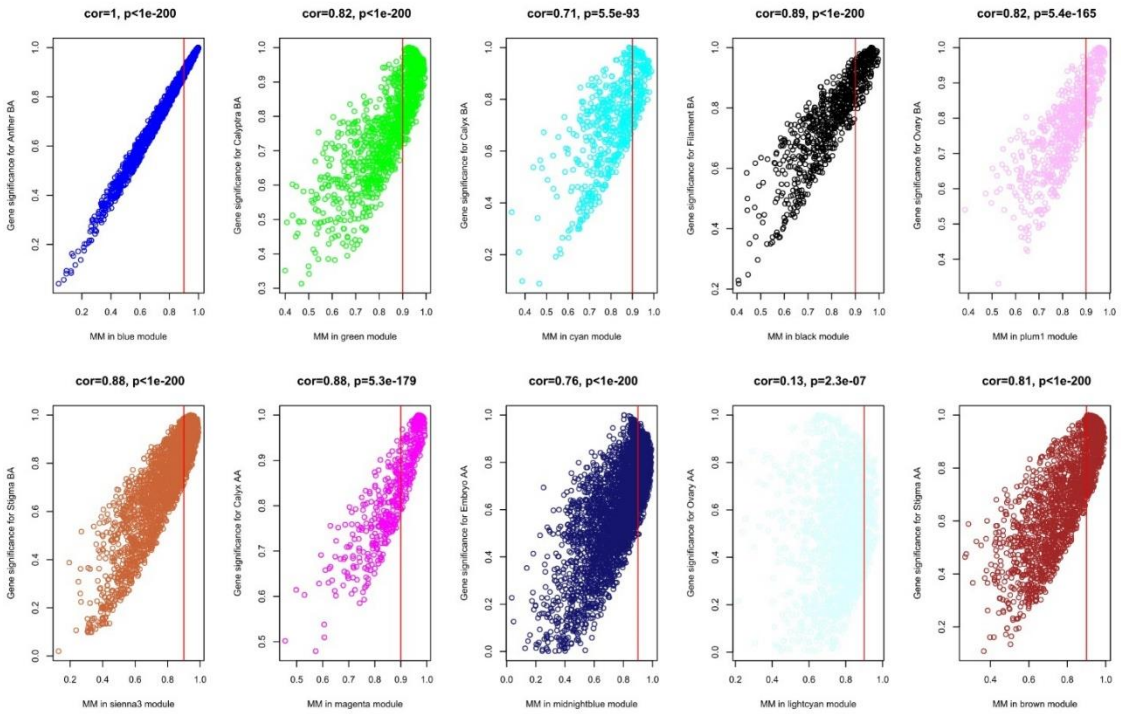


Figure 5.2. Scatterplots of gene significance (GS) vs. module membership (MM). The plots show the correlation between GS and MM in the ten modules that best correlate with the different tissues analyzed illustrating that gene highly significantly associated with a trait are often also the most important (central) elements of modules associated with the trait. Genes showing a MM > 0.9 can be considered as hub-genes.

Abstractly speaking, if a gene has higher GS and MM, it is more meaningful with the phenotypical trait (Lou et al., 2017; Langfelder et al., 2013). Thus, a specific module whose MM or GS were significantly connected and associated with the anther tissue may play a more important biological role on anther determination or functionality (Lou et al., 2017). Although all 10 modules considered showed extremely significant correlations between GS and MM, the blue (Anther BA), black (Filament BA) and sienna3 (Stigma BA) ones showed the best correlations between MM and GS (Fig. 5.2). Overall, module blue was observed as the best meaningful module by its strongly positive correlations (r = 1, p < E-200 in GS vs. MM) indicating its strict involvement in anther specific molecular mechanisms. In order to understand if modules associated with different tissues were enriched in genes belonging to determined ontological categories, we conducted a gene set enrichment analysis (GSEA) on those genes showing a MM > 0.9 (Fig. 6.2), which were considered as WGCNA hub genes.

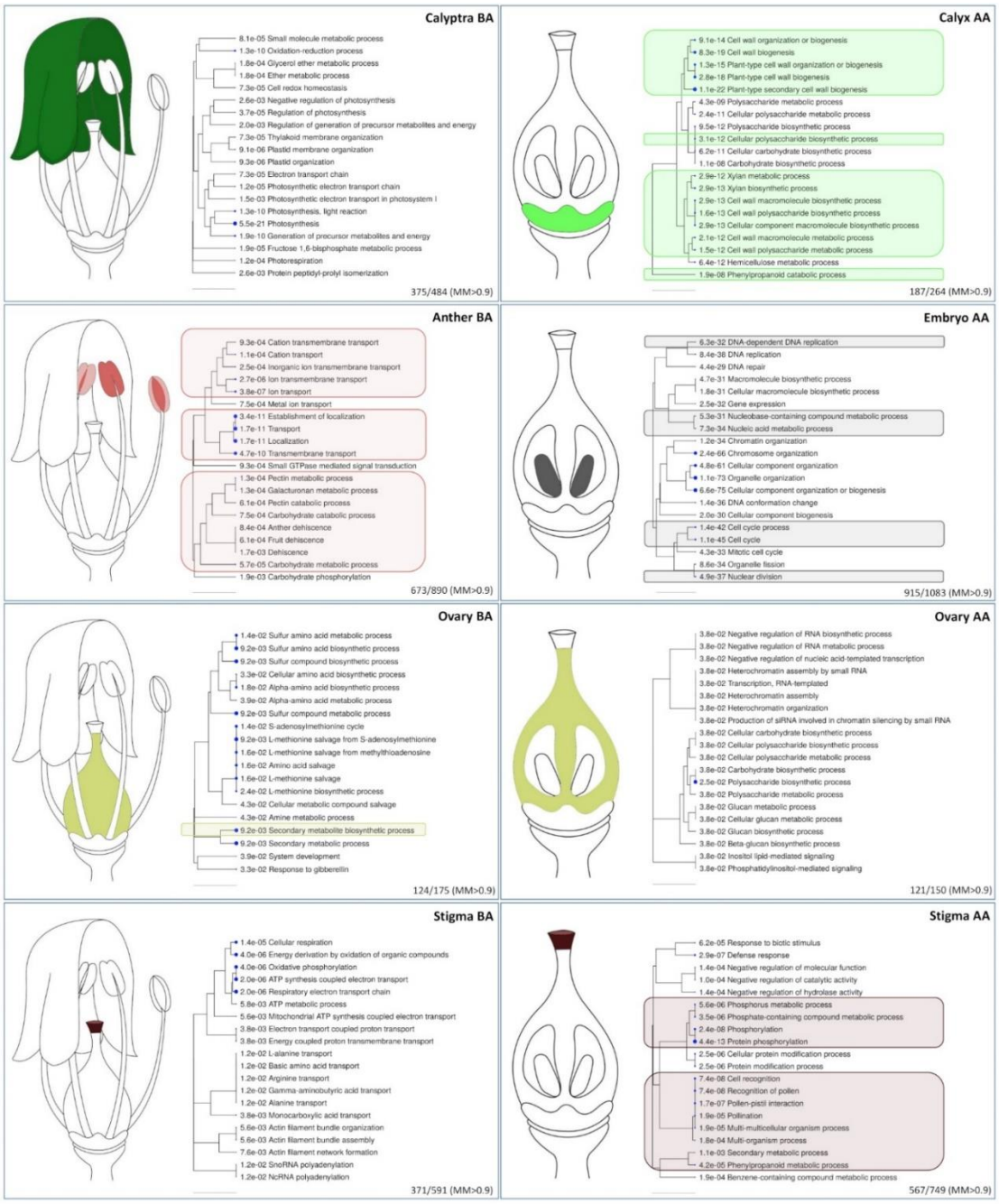


Figure 6.2. Gene set enrichment analysis. Genes showing a module membership (M) higher than 0.9 were subjected to GSEA and the most relevant biological networks for Anther BA, Calyptra BA, Calyx AA, Stigma BA, Ovary BA, Stigma AA, Ovary AA and Embryo AA are reported. The hierarchical clustering trees summarize the correlation among significant pathways identified. Pathways with many shared genes are clustered together. Bigger dots indicate more significant p-values. GO categories highlighted by colored boxes were detected also in the GSEA on highly specific genes (HSG) obtained by the tau (t) analysis.

The analysis allowed to identify the most relevant biological networks for all the tissues except for Filament BA and Calyx AA. This could be due to both the small size of the modules and the absence of specific processes occurring in these whorls. In Calyptra BA we identified a consistent number of biological networks associated to “photosynthesis” (GO:0015979) and “photosynthesis light reaction” (GO:0019684). Within these GO categories, two genes encoding light-harvesting chlorophyll A-binding proteins (*LHCA1* and *LHCA4*) stand out. Their expression resulted sensibly higher in calyptra (1843 and 1153 TPM, respectively; Supplementary Table 3) compared to any other tissue, in accordance with the active role of this green tissue in the photosynthesis process (Lebon et al., 2005). The most enriched GO categories in Anther BA were “localization” (GO:0051179), “transport” (GO:0006810), and establishment of localization (GO:0051234). Something similar was observed in the Corvina expression atlas, where pollen and stamen were characterized by the strong expression of genes related to transport and cell wall structure (Fasoli et al., 2012). In addition, “anther dehiscence” (GO:0009901), “fruit dehiscence” (GO:0010047), and “dehiscence” (GO:0009900) were the most enriched categories in Anther BA. Five polygalacturonase (PG) encoding genes (Supplementary Table 3) were included in all the categories abovementioned and exhibited the highest expression levels in Anther BA. *VIT_13s0064g00760*, a member of the polygalacturonase GH28 sub-family, showed an impressive transcripts accumulation (3310 TPM) in the male reproductive organ whereas its levels were always very low (from 0 to 17 TPM) in all the other tissues. The expression of PG genes in tapetum, pollen grains, stigmas and pollinated pistils has been described in various species and implies their role in tapetum degradation, pollen maturation, pollen tube growth, and pollination (Yang et al., 2018). In *Arabidopsis*, suppressing the expression of *QRT3* interferes with microspore separation after the tetrad stage (Rhee et al., 2003), whereas in Chinese cabbage *BcMF6* silencing determines smaller floral organs and a lower pollen germination rate caused by the disruption of microspore maturation (Zhang et al., 2008). In the same species, *BcMF2* and *BcMF9* RNA antisense lines showed disturbed development of the pollen wall intine layer and of the pollen tube wall (Huang et al., 2009; Huang et al., 2009) and the downregulation of *BcMF2* caused pollen deformity and balloon-like swelling in the pollen tube tip, together with premature tapetum formation (Huang et al., 2009). When a soybean PG is heterologously overexpressed in Arabidopsis, inflorescence mortality is over 50%, and siliques and seeds significantly decrease in number (Wang et al., 2016). With regards to post-anthesis tissues, “recognition of pollen” (GO:0048544), “pollen-pistil interaction” (GO:0009875) and “cell recognition” (GO:0008037) were unarguably the most interesting terms enriched in Stigma AA. Most of the genes highly expressed in this tissue (and scarcely or no detected

within the other tissues), as expected, were linked to the self-incompatibility locus (S-locus) and mainly represented by kinases. In Stigma AA, of interest was also the detection of 99 genes included in the “protein phosphorylation” (GO:0006468) and “phosphorylation” (GO:0016310) categories, both involved in pollen/stigma interaction processes (Hiscock et al., 1995; Rundle et al., 1993). Calyx AA showed a significant number of genes covering categories such as “plant-type secondary cell wall biogenesis” (GO:0009834) and “cell wall biogenesis” (GO:0042546). Among them, most noteworthy is the almost exclusive expression of three distinct genes (*VIT_06s0004g03050*, *VIT_08s0040g01970* and *VIT_08s0040g02030*) all belonging to the fasciclin arabinogalactan-proteins (FLA11) family. In this regard, the Arabidopsis orthologous (*AT5G03170*) seems to be pivotal for tensile strength and tensile modulus of elasticity, two features that match with the mechanical sustain role of calyx after fertilization and fruit set. Finally, the midnightblue module, the one that best correlates with Embryo AA, included genes mainly involved in cell ontogenesis such as “cellular component organization or biogenesis” (GO:0071840), “organelle organization” (GO:0006996), and “chromosome organization” (GO:0051276), coherently with the intense embryonal development activity (Supplementary Table 3).

2.3.3 WGCNA and tau analyses identified whorl-specific transcriptional regulators

In order to identify transcriptional regulators specifically expressed in different tissues and whorls analyzed in this study, hub genes belonging to the different tissue-specific modules showing MM > 0.9, were screened based on functional annotation reported in the Plant Transcription Factor database (Plant TFDB; Jin et al., 2017). Considering specific genes identified by WGCNA we selected 251 TF genes representing 36 TF families. In absolute terms, the Stigma AA, Embryo AA and Anther BA were the tissues with the largest number of specific TF-coding genes (83, 40 and 31, respectively). In contrast, Ovary AA and Calyptra BA were the tissues with the lowest number of specific TFs (7). Overall, the MYB-R2R3 (41 genes), WRKY (24), bHLH (20), NAC (16), MICK-MADS (12), ERF (12) families were the most abundant, although with differences depending on the tissue considered (Fig. 7.2A; Supplementary Table 4). For example, whether in Stigma AA there was a consistent number of MYB-R2R3 (18) and WRKY (16), in Ovary BA they were lacking, leaving room to genes belonging to minor TF families, such as B3 or G2-like. While in Supplementary Table 4 we listed all tissue-specific TFs identified, in Table 1 we reported the most relevant ones based on module membership (MM > 0.9) and expression. Overall, a rather large number of transcription factors identified appeared to be related to roles in flower development, determination or identity.

Amongst them it is worth mentioning *VviMYB108A* (*VIT_05s0077g00500*) (Wong et al., 2016), highly expressed in Anther BA, whose *Arabidopsis* orthologous is a JA-inducible TF gene with an important role in stamen development and male fertility, being involved in three main aspects: filament elongation, anther dehiscence, and pollen viability (Mandaokar and Browse, 2009). One of the most expressed TFs in stigma BA, the bHLH gene *BEE1*, is orthologous of the BR-responsive gene *AtBEE1* that regulates stigmatic cell development in Arabidopsis (Crawford et al., 2011). The occurrence of *VIT_17s0000g03580*, another BR-responsive bHLH, in stigma best ranked TFs, raises questions about the involvement of these hormones in the processes of flower development and fertilization.

Table 1. Top 5 tissue-specific TFs identified by WGCNA analysis (MM > 0.9) ordered by descending TPM values.

ID	TF family	Grape name	Functional annotation	MM	Analysis	Calyx BA	Cap BA	Filament BA	Anther Ba	Ovary BA	Stigma BA	Calyx AA	Ovary AA	Stigma AA	Embryo AA
VIT_04s0023g01910	MYB_related	-	Myb family	0.97	WGCNA	552.2	130.6	12.0	3.4	100.2	80.3	217.0	32.1	16.7	0.0
VIT_01s0011g00100	MIKC_MADS	*VviAP1*	MADS-box APETALA 1	0.94	WGCNA	390.1	159.4	27.9	5.8	9.6	2.2	152.4	1.8	0.5	0.2
VIT_14s0060g01180	TALE	*VvHB49*	KNAT2 (knotted1-like homeobox gene 6)	0.92	WGCNA	154.8	58.6	52.8	7.1	54.0	1.3	41.5	2.5	0.5	8.4
VIT_12s0059g01190	TALE	*VvHB43*	Homeobox shoot MERISTEMLESS (STM)	0.93	WGCNA	93.6	3.3	0.7	1.5	10.6	14.4	60.8	11.2	10.7	2.4
VIT_05s0124g00240	bHLH	*bHLH025*	Basic helix-loop-helix (bHLH) family	0.92	WGCNA	79.0	29.9	8.4	1.0	38.8	22.5	42.3	20.2	32.1	11.6
VIT_15s0048g02870	HD-ZIP	*VvHB54*	Homeobox-leucine zipper protein HB-7	0.96	WGCNA	35.5	101.0	31.8	9.3	11.9	12.6	8.0	6.5	10.9	0.1
VIT_14s0068g00330	TCP	-	PTF1 (plastid transcription factor 1) TCP13	0.91	WGCNA	46.2	82.6	1.2	0.3	3.1	0.7	21.4	22.8	24.0	2.6
VIT_04s0023g01020	TALE	*VvHB17*	BEL1 homeobox 2 (BLH2) (SAWTOOTH 1)	0.91	WGCNA	11.3	51.1	4.6	10.4	1.4	0.2	9.9	1.7	0.3	0.4
VIT_07s0031g01710	WRKY	*VvWRKY22*	WRKY DNA-binding protein 51	0.96	WGCNA	16.9	39.4	0.0	0.1	6.7	0.0	8.3	1.0	6.4	0.2
VIT_06s0004g03780	WOX	*VvHB24*	Wuschel-related homeobox 3 pressed flower	0.92	WGCNA/TAU	9.0	14.3	1.9	0.5	1.2	0.0	0.0	0.0	0.0	0.2
VIT_09s0002g06750	ERF	*VvERF042*	ERF/AP2 transcription factor sub B-6 SHINE	0.94	WGCNA	83.5	137.3	751.9	6.5	256.1	212.7	7.8	107.9	133.2	1.5
VIT_11s0016g03560	bHLH	-	Ethylene-responsive protein	0.93	WGCNA	57.5	18.3	177.7	51.6	84.3	1.7	42.0	27.0	0.2	13.1
VIT_04s0023g01380	GRAS	*GRAS25*	Scarecrow-like	0.92	WGCNA	41.6	30.1	111.3	46.2	39.8	17.1	25.9	10.4	10.9	6.0
VIT_15s0048g01240	MIKC_MADS	-	MADS-box protein AGL20	0.94	WGCNA	21.8	11.4	82.4	26.1	33.6	12.4	8.1	0.0	1.3	2.9
VIT_04s0008g01820	MYB	*MYBPA9*	TT2 (transparent testa 2)	0.97	WGCNA/TAU	1.6	6.7	29.4	0.1	2.8	2.8	0.4	0.4	0.1	0.5
VIT_05s0077g00500	MYB	*MYB108A*	Myb domain protein 108	1.00	WGCNA	8.7	13.4	31.6	307.6	4.3	0.3	5.0	0.3	14.8	0.3
VIT_00s0313g00070	MIKC_MADS	*VviSVP1*	MADS-box protein SVP	1.00	WGCNA	11.9	6.8	13.7	191.0	14.8	11.8	9.2	12.7	19.8	4.9
VIT_07s0031g01870	C3H	-	Zinc finger (CCCH-type) family protein	1.00	WGCNA/TAU	0.3	0.5	1.0	142.2	0.6	0.0	0.7	0.0	0.3	0.6
VIT_12s0028g03050	NAC	*VvNAC34*	NAC domain-containing protein 73	0.99	WGCNA	2.7	1.6	3.5	116.5	3.1	0.6	12.4	2.3	1.3	0.5
VIT_12s0059g02500	CO-like	-	Constans-like 11	1.00	WGCNA/TAU	0.3	0.6	0.5	104.6	0.8	0.2	0.3	0.2	0.1	0.7
VIT_14s0083g01050	MIKC_MADS	*VviSEP1*	SEPALLATA1	0.93	WGCNA	423.7	171.0	147.2	153.0	624.4	222.9	148.5	113.5	154.5	230.5
VIT_01s0011g00140	YABBY	-	CRABS CLAW	0.97	WGCNA	29.2	8.0	24.6	2.1	262.9	8.7	0.1	0.4	0.0	0.1
VIT_05s0077g01860	ERF	*VvERF058*	ERF (ethylene response factor) sub B-2	0.98	WGCNA	15.3	3.1	16.8	5.0	107.8	14.8	16.1	7.4	7.0	23.3
VIT_08s0007g06270	SBP	-	Squamosa promoter-binding protein 9 (SPL9)	0.93	WGCNA	24.9	10.7	32.5	1.4	66.5	15.4	20.9	18.5	20.5	27.9
VIT_08s0007g02940	NAC	*VvNAC62*	NAC domain containing protein 90	0.96	WGCNA	7.0	2.3	0.3	0.2	52.4	0.9	6.4	4.4	6.3	3.9
VIT_01s0011g03720	bHLH	-	BEE1 (BR Enhanced expression 1)	0.94	WGCNA	2.9	1.3	1.3	1.4	48.2	1334.1	0.9	2.4	10.2	0.1
VIT_17s0000g05900	bHLH	-	Basic helix-loop-helix (bHLH) family	0.94	WGCNA	10.8	0.5	4.5	0.3	3.1	1305.8	12.9	2.9	19.6	0.0
VIT_12s0142g00360	MIKC_MADS	*VviAG1*	SHATTERPROOF 2	0.95	WGCNA	139.2	27.6	496.3	140.2	558.0	1248.6	105.6	271.6	203.4	247.9
VIT_17s0000g03580	bHLH	-	BEE3 (BR Enhanced expression 3)	0.94	WGCNA	15.1	11.5	0.7	0.9	24.4	693.8	5.7	13.6	19.3	1.3
VIT_15s0021g02510	GATA	*VvGATA4*	GATA transcription factor 2	0.98	WGCNA	43.6	51.1	97.1	2.4	48.6	502.0	9.2	27.8	12.5	2.6
VIT_17s0000g01230	MIKC_MADS	*VviTM8a*	MADS-box protein AGL20	0.96	WGCNA	73.0	0.1	0.0	1.9	3.9	1.4	191.7	3.5	1.1	0.2
VIT_18s0001g10160	WOX	*VvHB65*	Wuschel homeobox 4	0.95	WGCNA	31.8	3.9	13.0	0.7	2.3	0.5	63.4	19.2	14.0	6.0
VIT_00s0291g00020	ERF	-	Unknown protein	0.98	WGCNA/TAU	5.0	6.0	0.0	0.0	2.7	0.0	45.7	4.4	0.5	0.4
VIT_18s0001g12610	MYB_related	-	Radialis-like protein 6	0.93	WGCNA/TAU	22.6	0.4	0.0	0.5	1.0	0.0	43.6	1.9	0.4	0.0
VIT_14s0006g02450	C3H	-	Zinc finger (CCCH-type) family protein	0.96	WGCNA/TAU	3.8	0.0	8.1	0.6	1.3	0.0	38.4	1.0	0.0	1.2
VIT_16s0050g02530	MYB_related	-	Myb Triptychon	0.90	WGCNA	26.8	19.7	8.2	5.7	29.2	2.4	38.7	90.3	22.0	39.0
VIT_14s0030g01860	Trihelix	-	Transcription factor	0.93	WGCNA	20.0	24.8	29.3	1.4	34.8	0.5	46.8	77.1	42.2	66.1
VIT_08s0007g00410	MYB	*MYB91A*	Myb domain protein 91	0.93	WGCNA	30.6	26.5	33.0	1.7	37.7	6.9	42.4	76.2	34.1	62.7
VIT_13s0067g02860	SRS	-	LRP1 (lateral root primordium 1)	0.95	WGCNA	21.4	13.0	9.4	2.3	11.4	5.3	31.1	31.6	17.6	20.3
VIT_06s0004g02800	HD-ZIP	-	leucine zipper protein Revoluta (REV)	0.94	WGCNA	9.7	5.0	8.6	1.1	12.5	2.7	25.0	25.5	9.1	13.1
VIT_12s0028g03270	ERF	-	Ethylene-responsive 9	0.95	WGCNA	480.6	277.9	255.2	15.0	256.9	65.0	376.9	246.1	961.9	209.6
VIT_18s0001g05250	ERF	-	DREB sub A-6 of ERF/AP2 (RAP2.4)	0.94	WGCNA	386.5	151.8	150.3	29.6	140.8	235.7	392.3	353.2	839.3	331.4
VIT_01s0026g02710	NAC	*VvNAC26*	NAC domain-containing protein 29	0.96	WGCNA	61.2	52.5	6.2	2.2	7.4	4.1	56.9	23.6	360.5	7.9
VIT_16s0098g01170	HD-ZIP	*VvHB56*	Homeobox-leucine zipper protein HB-12	0.92	WGCNA	1.1	1.1	5.4	8.7	0.9	0.0	19.9	87.9	358.4	0.8
VIT_04s0008g05760	WRKY	*VvWRKY08*	WRKY DNA-binding protein 18	0.90	WGCNA	218.2	109.4	17.6	2.7	62.6	9.6	133.9	33.7	350.3	44.3
VIT_18s0001g12530	MYB_related	-	Retrotransposon protein	0.93	WGCNA	57.1	1.0	3.9	7.2	65.8	12.8	101.7	26.5	2.2	282.8
VIT_06s0061g01240	C2H2	*VviZFP42*	Histone deacetylase (HD2A)	0.98	WGCNA	93.5	46.2	61.3	31.2	148.5	78.4	149.8	152.1	110.3	275.6
VIT_10s0042g00820	MIKC_MADS	*VviABS1*	Transparent testa16	0.92	WGCNA	0.0	0.2	4.3	1.5	73.1	3.4	42.1	0.9	0.2	146.8
VIT_17s0000g00330	bHLH	*bHLH075*	Inducer of CBF expression 1 ICE1	0.90	WGCNA	1.7	0.5	4.1	0.3	11.9	1.0	38.6	4.1	0.6	105.6
VIT_08s0007g01920	MYB	-	MYB divaricata	0.92	WGCNA	40.7	20.6	27.4	5.4	44.6	12.7	59.8	61.7	21.2	82.4

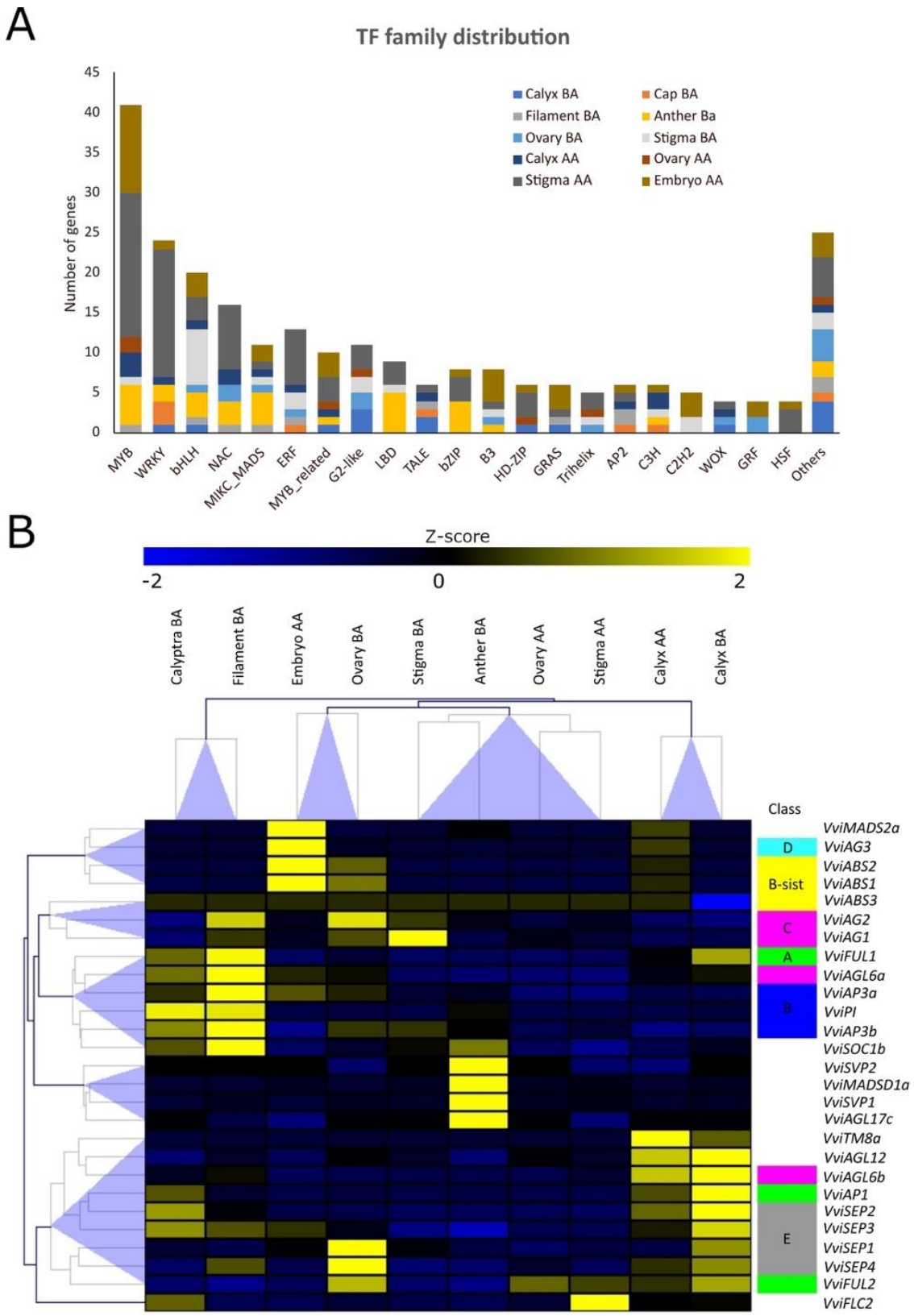

A
TF family distribution
Number of genes
Calyx BA
Cap BA
Filament BA
Anther Ba
Ovary BA
Stigma BA
Calyx AA
Ovary AA
Stigma AA
Embryo AA
MYB
WRKY
bHLH
NAC
MIKC_MADS
ERF
MYB_related
G2-like
LBD
TALE
bZIP
B3
HD-ZIP
GRAS
Trihelix
AP2
C3H
C2H2
WOX
GRF
HSF
Others
B
Z-score
Calyptra BA
Filament BA
Embryo AA
Ovary BA
Stigma BA
Anther BA
Ovary AA
Stigma AA
Calyx AA
Calyx BA
Class
VviMADS2a
VviAG3
VviABS2
VviABS1
VviABS3
VviAG2
VviAG1
VviFUL1
VviAGL6a
VviAP3a
VviPI
VviAP3b
VviSOC1b
VviSVP2
VviMADSD1a
VviSVP1
VviAGL17c
VviTM8a
VviAGL12
VviAGL6b
VviAP1
VviSEP2
VviSEP3
VviSEP1
VviSEP4
VviFUL2
VviFLC2

Figure 7.2. Tissue-specific transcription factors based on WGCNA. (A). Distribution of tissue-specific TF families across the 10 tissues related WGCNA modules. Only TF having a module membership higher than 0.9 were considered. (B) Heatmap showing the behavior of the main homeotic genes described in [40] together with those additional MADS box identified by the WGCNA analysis. Data were normalized using the gene/row normalization provided by T-mev software. This approach transforms values using the mean and the standard deviation of the row of the matrix to which the value belongs, using the following formula: Z-score = [(value) – mean(row)]/[standard deviation(row)]. Colored boxes close to gene names indicate the homeotic class of appartenance (for genes that have one attributed to). Hierarchical clustering of both genes and samples grouped genes/samples showing similar behavior.

Vogler et al. (2014) hypothesized that the growth-promoting properties of the reproductive tract of *Arabidopsis* depend, at least partly, on BR compounds, which are provided by the cells of the reproductive tract to promote pollen germination on the stigmatic papillae, and to boost pollen tube growth for rapid double fertilization (Vogler et al., 2014). Another TF, *VviAG1* (*VIT_12s0142g00360*), was highly expressed in stigma BA and its orthologous in Arabidopsis, the MADS box gene *SHATTERPROOF 2*, was demonstrated to be involved in promoting stigma, style and medial tissue development (Colombo et al., 2010). Certainly, a gene family of particular interest when it comes to floral identity is represented by floral homeotic genes, which are the basis of the ABCDE model and are well-studied genes involved in flower development (Krizek et al., 2005). These genes, belonging to the MICK_MADS family, already characterized at the genomic level by Grimplet et al. (2016), have recently been characterized from their transcriptional sub-functionalization by Palumbo et al. (2019), who selected 18 MADS boxes belonging to different classes (A, B, C, D, E) and analyzed their expression in different whorls during Pinot noir flower development. Within the tissue-specific TF obtained in this study, we identified 12 TF belonging to the MICK_MADS box family. Amongst them, only 5 genes fall within those homeotic genes considered by Palumbo et al. (2019), namely *VviAG1* (*VIT_12s0142g00360*), a class C gene listed in the Stigma BA related module (sienna3), *VviSEP1* (*VIT_14s0083g01050*), a class E gene belonging to the Ovary BA module (plum1), *VviAP1* (*VIT_01s0011g00100*), a class A gene belonging to the Calyx BA module (cyan), *VviABS1* and *VviABS2* (*VIT_10s0042g00820* and *VIT_01s0011g01560*), both belonging to the B-sister class and detected in the Embryo module (midnightblue). The absence of the other homeotic genes studied by Palumbo et al. (2019) from those identified in this study is likely due to the approach used to analyze the data. By filtering by module membership, most of those genes expressed simultaneously in more than one tissue, an intrinsic characteristic of some homeotic genes on which the ABCDE model is based, were in fact set aside. Nonetheless, we

identified some other MADS box genes which appear to be expressed in different tissues and which deserve a thought. Amongst these are *VviAGL17c* (*VIT_00s0211g00180*), *VviSVP1* (*VIT_00s0313g00070*), *VviMADSD1a* (*VIT_07s0031g01140*), *VviTM8a* (*VIT_17s0000g01230*), *VviSVP2* (*VIT_18s0001g07460*), *VviFLC2* (*VIT_14s0068g01800*) and *VIT_15s0048g01240*, an AGL20-like gene. Although some of these genes showed limited expression, some others, such as *VviSVP1*, *VviTM8a* and *VviFLC2* were significantly induced in specific whorls. *VvSP1* belongs to the Anther BA specific module (blue) and showed a transcript accumulation approximately 16 times higher with respect to the mean transcript accumulation of all other tissues. *VviTM8a* was the first ranked TF gene for expression in the calyx AA module (magenta), being approximately 20 times more expressed than in all other tissues. Finally, *VviFLC2* was expressed preferentially in stigma AA (Fold Change = 9 compared to other tissues). To provide a global view of the behavior of all homeotic genes of interest, the heatmap in Fig. 7.2B reports all the genes analyzed by Palumbo et al. (2019), together with those that have emerged from the WGCNA in this study.

VviAP1, a class A homeotic gene, was highly expressed in Calyx (both BA and AA) and in Calyptra BA, in agreement with what was observed by Palumbo et al. (2019) and with previous observations in other plant species such as *Arabidopsis* (Bowman et al., 1993), *Camellia japonica* (Sun et al., 2014) and *Medicago trucantula* (Roque et al., 2018). *VviFUL1* (*VIT_17s0000g04990*) and *VviFUL2* (*VIT_14s0083g01030*), the two grapevine orthologues of Arabidopsis *FRUITFULL* (*FUL*) (Grimplet et al., 2016), showed distinctive expression patterns. *VviFUL2* was expressed in ovary and calyx before anthesis, whereas *VviFUL1*, was expression was generally much lower, was expressed in Calyx BA, Calyptra BA and, most intriguing, in Filament BA. Class B genes, namely *VviPI*, *VviAP3a* and *VviAP3b* confirmed their role in petal and stamen identity with the highest transcript accumulation detected in Calyptra BA and stamen tissues (anther and/or filament). Amongst the B-sister genes, *VviABS1* and *VviABS2* perfectly matched what was expected based on Palumbo et al. (2019), being exclusively expressed in Ovary BA and in embryo. For what concerns the class C genes, *VviAG2* and *VviAGL6a* were preferentially expressed in stamen, at the level of filament, whereas *VviAG1* and *VviAGL6b* were expressed in Stigma BA and in Calyx BA and AA, respectively. *VvAG3*, a class D gene, was switched off in all tissue except for embryo, as previously described by Palumbo et al. (2019) and Boss et al. (2003). Finally, amongst the class E genes, *VvSEP1-4* transcripts were accumulated preferentially in Calyx BA and AA, with the exclusion of *VviSEP1* which was only detected in pre anthesis phase. Moreover, a relevant expression in Ovary BA was detected for *VviSEP1* and *VviSEP4*.

2.3.4 Isolation of whorls/tissue-specific gene markers using the tau (t) analysis

The WGCNA analysis, based on the identification of clusters of highly correlated genes sharing similar expression patterns across all samples, allowed the determination of groups of genes closely associated with a specific phenotypic character, represented in this study by a determined floral tissue or whorl.

This analysis has proved particularly effective in numerous studies (Wang et al., 2020; Li et al., 2019), nevertheless, the fact that a particular gene is highly expressed in one tissue compared to others is a relative parameter and, in some cases, it may be more useful to identify genes that are exclusively expressed in one organ and not in others: in other words, specific tissue/organ gene markers. For this purpose, we applied an algorithm generally used in transcriptomic studies on animals or humans. Such an algorithm, defined as tau (t) algorithm (Kryuchkova-Mostacci and Robinson-Rechav, 2017), can determine the tissue-specificity level of each predicted gene of a given genome.

After the quantile normalization of 22,094 genes (selected because showing TPM values equal or higher than 1 in at least one of the 10 samples) and the creation of BIN profiles, the implementation of the t algorithm led to the assignment of a value ranging from 0 (constitutively expressed in all or most of the tissues) to 1 (absolutely specific for a given tissue) to each gene. The uneven occurrence of the t values throughout the gene set is illustrated in Fig. 8.2A and is coherent with what is expected based on Kryuchkova-Mostacci and Robinson-Rechav (2017). Overall, 3,575 genes proved to be highly specific (HSG, tau > 0.85) and, among them, 1,514 resulted absolutely specific (ASG, t = 1). The tau value only defined the "specificity" of a gene, whereas to determine which tissue the gene is specific for, the tau expression fractions (t_{ef}) were calculated.

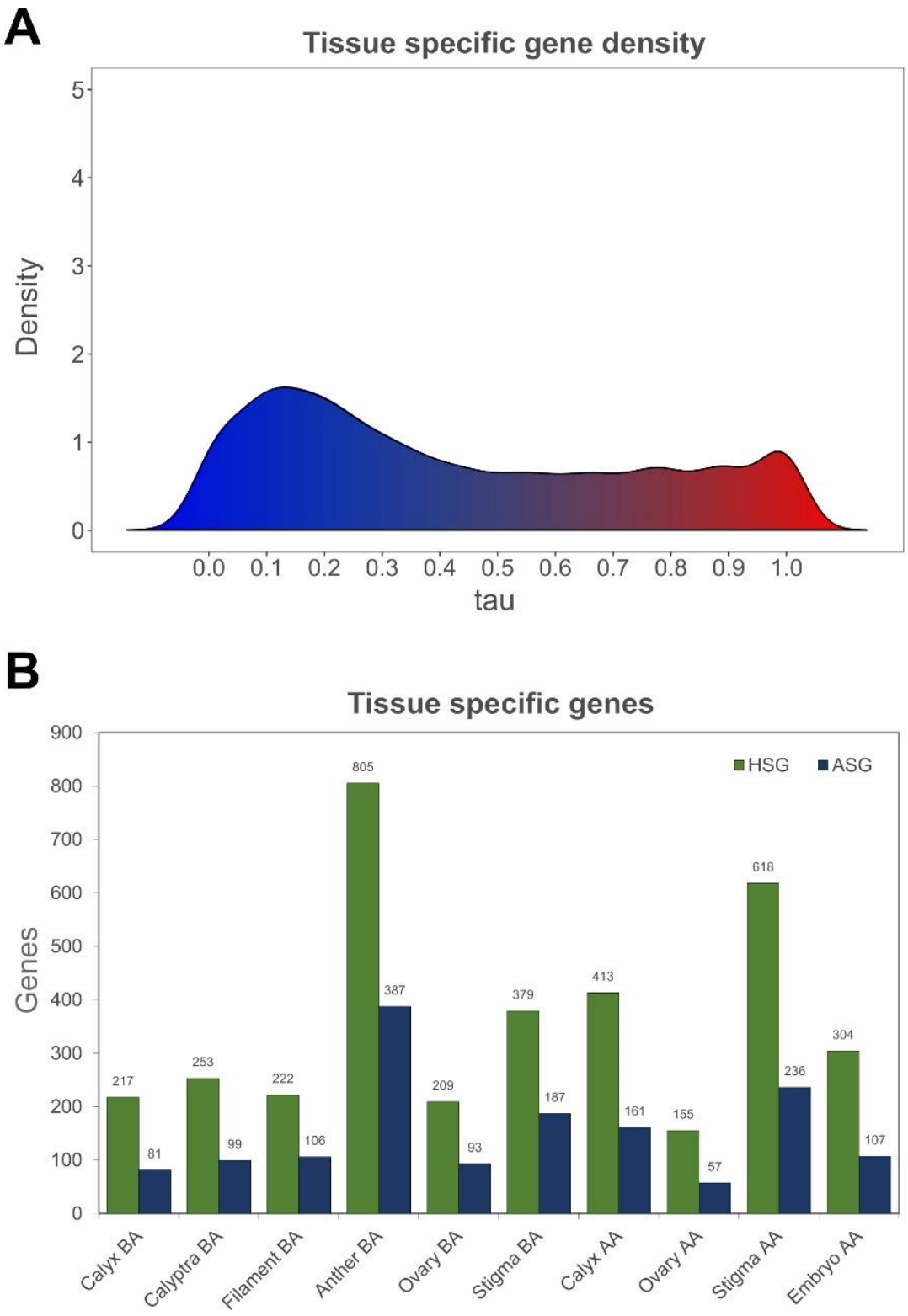


Figure 8.2. Tissue-specific gene distribution. (A) Distribution of tissue-specificity tau parameter over the 22,094 genes considered (TPM > 1). The shape of this plot and density distribution is coherent to what is expected based on Kryuchkova-Mostacci and Robinson-Rechav (2017). (B) Bar graph

showing the distribution of absolutely specific genes (ASG; tau = 1) and highly specific genes (HSG; tau > 0.85) over the ten tissues/organs considered.

Anther BA was the tissue displaying the highest number of HSG (805) and ASG (307) while, on the contrary, the lowest HSG and ASG values (155 and 57, respectively) were identified in Ovary AA (Fig. 8.2B). This observation is intriguing since anther BA represents the tissue that showed the lowest number of expressed genes (TPM> 1; Fig. 1.2), whereas Ovary AA was one of those tissues with the highest number of expressed genes. Nevertheless, this observation is partially confirmed in the Corvina expression atlas, where, out of 516 genes identified as specific flower, 229, equal to 44%, were specific for stamens and pollen (Fasoli et al., 2012). For each tissue, the list of HSG and ASG is available as Supplementary Table 5. Similar to what was done for the genes obtained through WGCNA, also in this case the highly specific genes (HSG) resulting from the tau analysis were subjected to a GSEA to verify the presence of enriched ontological categories and the possible overlap with those obtained in the WGCNA analysis. As a matter of fact, many enriched terms identified in HSG were common with those highlighted in the WCGNA analysis (Fig. 6.2). This observation conferred greater robustness to the results obtained and laid the foundations for a subsequent step aimed at further narrowing down the list of key genes of interest.

2.3.5 Comparison between tau and WGCN analyses and determination of best optimum tissue-specific genes

The comparison between the highly specific genes (HSG) identified by the tau analysis and the tissue-specific modules identified by WGCNA (MM > 0.9) led to the identification of 1513 genes shared by the two approaches. Looking at the specific tissues under study the percentage of common genes found its maximum in Anther BA (47.4%; Supplementary Table 2**;** Supplementary Figure 2.2). This result is not surprising, considering that the modules isolated by WGCNA contain genes that can show high levels of expression even in those tissues that are not associated with the module itself, whereas highly specific genes (HSG) identified by (t) algorithm represent genes that are almost exclusively expressed in that specific tissue and not in others. Nevertheless, although the results provided by the two analyses have different biological meanings, we considered those genes shared between the two approaches to be of particular interest, defining them as key hub genes. Ranking of genes by both expression and specificity is useful for anyone working on a single tissue wanting to identify a set of genes that are highly specific to the tissue, that are expressed in high enough quantities (facilitating bench work in the laboratory), and with minimal expression in

other tissues (limiting off-target effects). With this aim, we retrieved the quantile normalized expression and tissue-specificity of every key hub gene detected by the WGCNA and tau analyses comparison and we used this information to create a score column. Each gene's score was between 0 and 2 and was the sum of its tau expression fraction value (t_{ef}) and its 0-1 ranged normalized expression value. For each tissue, we then considered the top 10 ranking genes based on score values, in other words, those genes with both the highest expression and specificity. Whilst the complete list of these genes is available as Supplementary Table 6 and it is graphically represented in Supplementary Figure 3, in Table 2 we reported the first ranking best optimum specific gene for each different tissue.

Table 2. Best optimum gene for each tissue under study based on tau analysis.

	ID	Function	τ_{ef}	Score	TPM before anthesis						TPM after anthesis			
					Calyx	Calyptra	Filament	Anther	Ovary	Stigma	Calyx	Ovary	Stigma	Embryo
Optimum tissues specific genes	*VIT_13s0106g00350*	Lipase GDSL	**0.99**	**1.41**	66	2	0	0	0	1	0	0	0	0
	VIT_00s0194g00180	Unknown protein	**0.90**	**1.42**	10	175	2	0	3	2	0	0	2	0
	VIT_15s0021g01510	Phospholipase A1	**1.00**	**1.34**	1	0	32	0	0	0	0	0	0	0
	VIT_07s0141g00030	fatty acid elongase	**0.87**	**1.62**	3	7	11	1727	2	0	3	0	1	0
	VIT_15s0046g03420	Unknown protein	**0.94**	**1.46**	4	2	0	0	170	1	2	1	0	0
	VIT_18s0001g14760	Lipase 3 (EXL3)	**0.89**	**1.58**	1	3	3	0	10	1034	1	2	0	0
	VIT_02s0154g00300	snRNP Sm D3	**0.86**	**1.43**	8	4	0	0	1	0	281	15	4	0
	VIT_18s0164g00100	Laccase	**0.87**	**1.37**	2	0	0	0	8	1	15	148	1	2
	VIT_16s0100g01100	Stilbene synthase	**0.97**	**1.55**	0	1	4	1	0	0	1	0	314	0
	VIT_13s0019g01520	lipid transfer protein	**0.89**	**1.50**	0	0	0	0	0	0	84	1	1	459

Among the most interesting optimum genes in pre anthesis, the best optimum gene in filament was a *Phospholipase A1* (*VIT_15s0021g01510*; t_{ef} = 1 and no expression in all the other tissues). Although no information on this gene is available in grapevine, its Arabidopsis orthologous, namely *AT2G44810*, encodes DEFECTIVE ANTHER DEHISCENCE 1 (DAD1), a protein located in the chloroplast (Seo et al., 2009) whose phospholipase A1 activity catalyzes the late phase of the jasmonate biosynthetic pathway (Hyun et al., 2008). *DAD1* expression appears to be restricted to stamen filaments immediately before flower opening and *dad1* mutants show defects in flower opening as well as anther dehiscence and pollen maturation (Ishiguro et al., 2002). The massive occurrence of transcript encoding for fatty acid elongase (*VIT_07s0141g00030*) in anther BA (TPM = 1726, compared to an average TPM = 3 in all the others tissues) is in agreement with the grapevine expression atlas published by Fasoli et al. (2012), where this gene was highly expressed in stamen, considering both anther and filament tissues, and in pollen grain. Transcripts of *A. thaliana* orthologous *AT2G26640* encoding a 3-ketoacyl-CoA-synthase 11 (KCS11) are accumulated, above all the other tissues, in stamens and pollen (Joubès et al., 2008). It's also worth noting that *AT1G68530*, another *Arabidopsis* orthologous encoding KCS6 another protein belonging to the ketoacyl-CoA-synthase family, is the major condensing enzyme involved in stem wax and pollen coat lipid biosynthesis (Kunst and Samuel, 2003) and is highly expressed in the tapetum of anthers near maturity (Hooker et al., 2002). *Kcs6* mutants are male sterile (Millar et al., 1999). *VIT_18s0001g14760* transcript was found to be the best optimum gene marker in stigma BA. This observation is consistent with what was observed by Fasoli et al. (2012) in the Corvina expression atlas. The related Arabidopsis orthologous *AT1G75900* encodes a GDSL esterase/lipase EXL3, a protein belonging to the extracellular lipases (EXLs). This class of proteins is abundant in pollen coat,

and its combination with lipids can interact with stigma cells, bringing the recognition signal and triggering a mechanical conduit that leads to pollen hydration (Preuss, 2002). It's interesting to note that EXL4, another GDSL esterase/lipase protein, is localized in small granules in the tapetal cells of pollen coat (Mayfield et al., 2001) it is required for its formation and is involved in male fertility. Mutants show reduced pollen fertility, underdeveloped pollen grain coat as well as impaired water absorption and germination capacities. Amongst the best optimum genes detected in post-anthesis phase, it is noteworthy the presence of a laccase encoded by *VIT_18s0164g00100* in the ovary, as confirmed in the Corvina atlas by Fasoli et al. (2012), where a high expression of this gene was detected in the pericarp and in the pulp of all berry developmental stages, above all in fruit set and post fruit set. The upregulation of *VIT_18s0164g00100* was associated with processes directly involved in berry ripening (Guo et al., 2020). Finally, *VIT_13s0019g01520* transcript was highly and specifically accumulated in embryo, again in agreement with the observation of the massive occurrence of this gene mRNA in fruit set and post fruit set seed made by Fasoli et al. (2012). The specific accumulation of a transcript coding for a stilbene synthase, *VvSTS36* (*VIT_16s0100g01100*) (Vannozzi et al., 2012), in the stigma AA is curious both for the fact that the induction of this gene and its paralogues had never previously been observed in this tissue, and because generally *STSs* are expressed in response to biotic or abiotic stresses. It is conceivable that in the post-anthesis phase, once fertilization has taken place, the stigma undergoes a senescence process. Several studies reported the accumulation of *STS* transcript and, consequently, of basic and complex stilbenes, in the senescence phase, probably as a response to the oxidative processes and the production of ROS which characterize senescing tissues but also in response to the accumulation of one of the main hormones linked to this process: ethylene, closely linked to the transcriptional activation of these genes. Ultimately, although the discussion of the results obtained only touched on the best optimum genes identified through the tau approach, it seems evident that most of the genes identified found confirmation in the literature. On this basis, through this tool, we believe we have made available a pertinent list of specific tissue/organ genes that may be of interest to the scientific community.

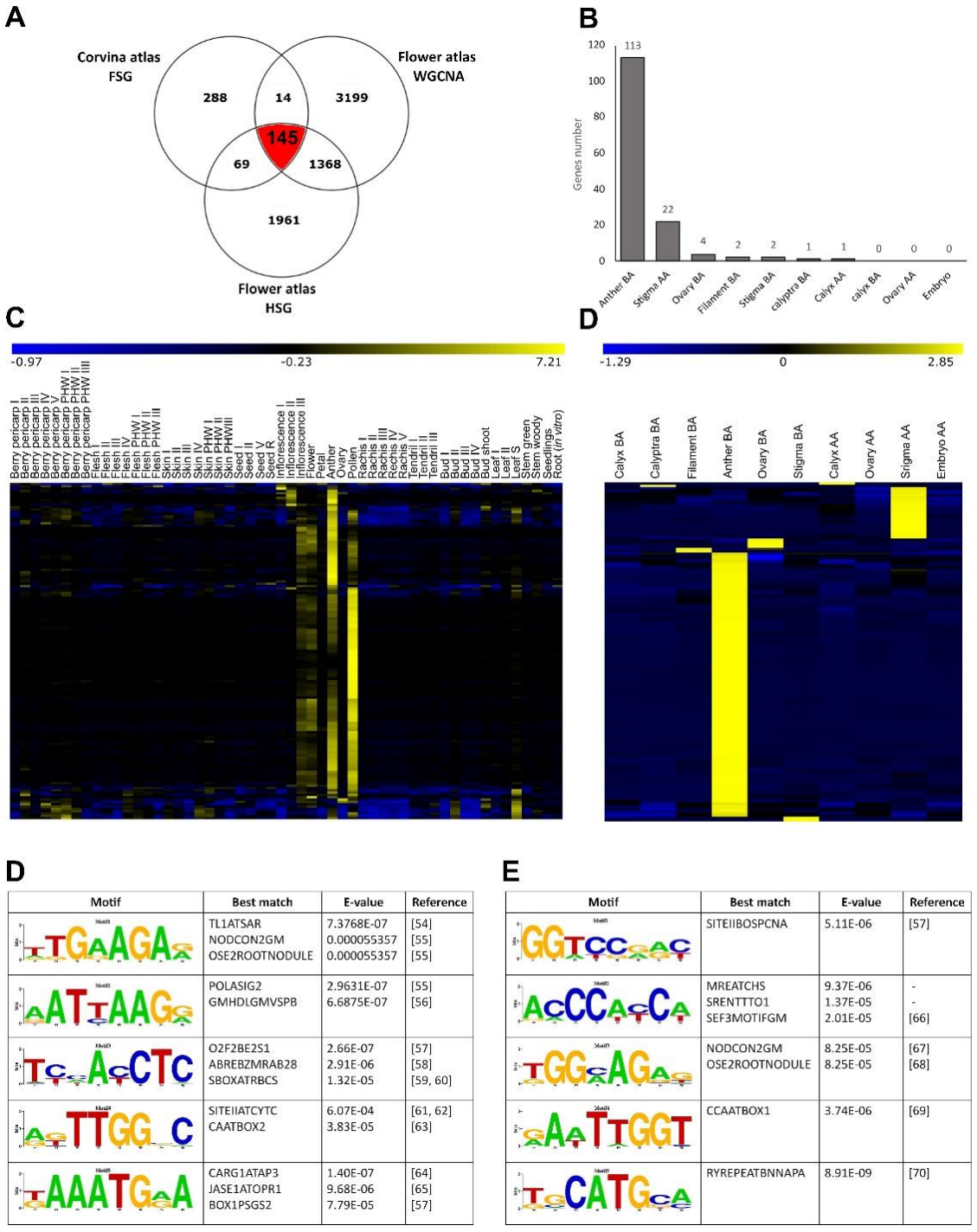
2.3.6 Integration of the flower and the Corvina expression atlases and Identification of enriched *cis*-regulating elements in flower-specific genes

Although WGCNA and tau analyses provided lists of genes of interest related to specific floral tissues or expressed exclusively in one whorl rather than another, based on samples considered in this study it was not possible to rule out the fact that these genes are also expressed in other tissues of

the plant. In order to further circumscribe the number of genes of interest at the floral level, we did an additional step, excluding those transcripts whose expression was reported also in other tissues based on the *V. vinifera* cv Corvina expression atlas (Fasoli et al., 2012). Based on this analysis, carried out using microarray technology, the number of flower specific genes (FSG) was estimated in 516 (Fasoli et al., 2012). We then crossed these data with those resulting from the tau and WGCN analyses, considering all HSG genes (tau > 0.85) identified in this study and all those genes belonging to the 10 tissue-specific modules identified in the WGCNA (MM > 0.9). Overall, 145 transcripts were found to be common between the three datasets (approximately 28.1% of the specific flower genes identified in the Corvina atlas; Fig. 9.2A, Supplementary Table 7), representing genes that are expressed only in flower based on the Corvina atlas and specifically expressed in one whorl rather than another based on our data. Two-hundred-eighty-eight genes were expressed only in flower based on the Corvina atlas but did not show specificity for any whorl based on the tau/WGCNA analyses, representing genes that are homogeneously expressed in all flower tissues considered. Of the 145 genes shared by the 3 datasets, the majority was highly specific for Anther BA (113 genes; 78%) and post-anthesis stigma (22 genes; 15%). The remaining ones were HSG for Ovary BA (4 genes), Filament BA (2 genes), Stigma BA (2 genes), Calyptra BA (1 gene), and Calyx AA (1 gene) (Fig. 9.2B). The heatmaps in Fig. 9.2 report the behavior of these genes across the 54 tissues considered in the Corvina atlas (Fig. 9.2C) and in the different floral tissues/whorls considered in this study (Fig. 9.2D) whereas Supplementary Table 7 provides a list of these key hub genes for each tissue considered.

In order to study the structure of co-expressed genes promoters and to identify genetic determinants of the tissue-specific expression of tissue-specific genes, we conducted a *de novo* motifs discovery analysis considering the 2kb sequences upstream the 113 and 22 found to be expressed only in anther and stigma, respectively, based on the WCGNA, tau analysis and Corvina expression atlas. The analysis was carried out exclusively in these two tissues because they were the only ones having enough genes shared between the WGCNA analysis, the Corvina atlas and the present flower atlas, to allow a reliable *de novo* motifs discovery. For this purpose, we retrieved the promoter sequences of selected genes previously isolated from the 12x V1 prediction of the PN40024 genome. Using the DECOD software, we screened these sequences for novel enriched motifs with k-mer equal to 8 with respect to 10,000 2kb promoter sequences randomly retrieved from the grapevine genome. We then used STAMP (Mahony et al., 2007) to match the motifs discovered against PLACE, a database of motifs found in plant *cis*-acting regulatory DNA elements collected from previously published studies (Higo et al., 1999). For both tissues analyzed, 10 motifs

enriched in the promoters of tissue-specific genes were identified. The comparison with *cis* element already deposited in the database via STAMP highlighted the 5 best matches. Among these, we only considered those CREs with a biological meaning.



Motif	Best match	E-value	Reference
	TL1ATSAR	7.3768E-07	[54]
	NODCON2GM	0.000055357	[55]
	OSE2ROOTNODULE	0.000055357	[55]
	POLASIG2	2.9631E-07	[55]
	GMHDLGMVSPB	6.6875E-07	[56]
	O2F2BE2S1	2.66E-07	[57]
	ABREBZMRAB28	2.91E-06	[58]
	SBOXATRBCS	1.32E-05	[59, 60]
	SITEIIATCYTC	6.07E-04	[61, 62]
	CAATBOX2	3.83E-05	[63]
	CARG1ATAP3	1.40E-07	[64]
	JASE1ATOPR1	9.68E-06	[65]
	BOX1PSGS2	7.79E-05	[57]

Motif	Best match	E-value	Reference
	SITEIIBOSPCNA	5.11E-06	[57]
	MREATCHS	9.37E-06	-
	SRENTTTO1	1.37E-05	-
	SEF3MOTIFGM	2.01E-05	[66]
	NODCON2GM	8.25E-05	[67]
	OSE2ROOTNODULE	8.25E-05	[68]
	CCAATBOX1	3.74E-06	[69]
	RYREPEATBNNAPA	8.91E-09	[70]

Figure 9.2. Comparison between flower-specific genes in the Corvina atlas and highly specific genes detected in this study. (A) Venn diagram showing specific or common genes between the Corvina expression atlas (Fasoli et al., 2012) and the highly specific genes detected by tau (HSG) and the

WGCNA analyses in this study. One-hundred-forty-five genes were expressed exclusively in flower based on the Corvina atlas and at the same time turned out to be tissue-specific based on the tau and WGCNA analyses. (B) Distribution of the 145 common genes in the different floral tissues based on tissue-specificity; (C) heatmap showing the expression of 145 common genes in the 54 tissues/organs analyzed in the Corvina expression atlas; (D) heatmap showing the expression of the same genes in the flower expression atlas object of this study. For both (C) and (D) panels data were normalized using the gene/row normalization provided by T-mev software. This approach transforms values using the mean and the standard deviation of the row of the matrix to which the value belongs, using the following formula: value = [(value) – mean(row)]/[standard deviation(row)]. (E) Top-5 *cis* regulatory elements (CREs) detected in 113 anther BA specific genes shared by the three datasets considered. (F) Top-5 *cis* regulatory elements (CREs) detected in 22 stigma AA specific genes shared by the three datasets considered.

In Anther BA (Fig. 9.2E) the first ranked motif based on the score, motif 1 (score = 4.41E-04), was similar to motifs TL1ATSAR (E-value 7.38 E-07) and NODCON2GM (E-value 5.54E-05), both enriched in promoter regions of genes involved in reproductive processes. TL1ATSAR was identified in *Arabidopsis* and is associated with male meiocyte-expressed genes (Li et al., 2014) whereas NODCON2GM is a cis-regulatory element of the gene TM6, directly involved in stamens petaloidy and flower shape formation In *Peonia* spp. (Shu et al., 2012). Motif 2 (3.26E-04) was associated with POLASIG2 and GMHDLGMVSPB. POLASIG2 is highly related to a large number of flowering genes of Arabidopsis and GMHDLGMVSPB was found to be highly and specifically related to pollen tapetum-expressed genes in rice (*O. sativa*) (Hobo et al., 2008). Something similar was observed for O2F2BE2S1, the best match of Motif 3. The ABREBZMRAB28 motif was found to be enriched in the promoter region of bZIP genes involved in floral development in six strawberry species and it is presumed that they also play important role in the male gametophyte (Liu et al., 2017). Moreover, in Arabidopsis, several AtbZIP genes were selected for their putative involvement in pollen development (Liu et al., 2017). Moving from anther to stigma AA, the other tissue that allowed, given the number of specific tissue genes, to carry out the analysis of the promoters, among the identified motifs whose biological significance has been confirmed by the similarity with already characterized CREs stand out SEF3MOTIFGM, NODCON2GM and RYREPEATBNNAPA. SEF3MOTIFGM was found in the promoter region of SEP3. SEP3 sequence of Platanus acerifolia was heterologously expressed in tobacco and through a GUS assay observed to be expressed in reproductive tissues, among them also stigma (Lu et al., 2009). NODCON2GM motif is enriched in the promoter region of the poplar gene PtaRHE1, coding for a RING-H2 protein, ectopically

expressed in tobacco and observed highly expressed in the stigma through a GUS assay (Mukoko Bopopi et al., 2010). Finally, RYREPEATBNNAPA in *Arabidopsis*, was found to be enriched in ABA-related differentially expressed genes which were observed to be highly affected by the overexpression of MINI ZINC FINGER 1 (MIF1), a putative zinc finger protein (Hu and Ma, 2006). For mutants, a lot of defects were observed in floral whorls, including stigma, showing also reduction in fertility (Hu and Ma, 2006).

Ultimately, many of the regions identified through this approach found confirmation in the literature. An interesting investigation would be to evaluate the possible relationships existing between the transcription factors identified through WGCNA analysis (for example *MYB108A*, *VviSVP1* and *VvNAC34*) and the *cis* elements identified here. To date, there are several NGS approaches that could shed light on the physical interaction of these *trans* and *cis* factors, first the DNA Affinity Purification Protocol (DAPseq) (Bartlett et al., 2017). This could validate the possible interaction between these specific TFs and specific tissue/organ genes and, in association with further RNAseq analyses, could provide numerous information about the tissue-specific transcriptional networks of the flower, as well as about the cistrome landscape of candidate TFs.

2.3.7 Isolation of novel housekeeping genes based on the floral and Corvina expression atlases

Looking at constitutive (housekeeping) genes, we identified 662 genes that exhibited a τ value = 0 and TPM values > 100 in all tissues. The complete list (in ascending order based on standard deviation of TPM values), is available as Supplementary Table 8, while Supplementary Figure 4 depicts the main biological networks resulting from the GO enrichment analysis of the 662 housekeeping genes here identified. The GO categories more represented were “positive regulation of translational termination” (GO:0045905), “proteasomal ubiquitin-independent protein catabolic” (GO:0010499) and “assembly of large subunit precursor of pre-ribosome” (GO:1902626) that scored respectively fold enrichment values of 34.15, 27.32 and 22.77. It is not surprising that the most represented categories (and the related genes) resulting from the GO enrichment analysis are involved in molecular processes common to all tissues and strictly required for cellular survival (protein synthesis and degradation). A further step was taken to verify whether the 662 constitutive floral genes found in all the tissues here analyzed, were also constitutively expressed in the 54 tissues considered in the Corvina atlas (Fasoli et al., 2012): in this way, we managed to ascertain the qualitative expression of the entire gene list also in the Corvina atlas and to reclassify therefore

these loci as “whole plant housekeeping genes”. Due to their crucial role in cellular survival and to their constitutive expression, housekeeping genes or control genes play a decisive role for mRNA levels normalization in qPCR studies. At this aim, several studies attempted the identification of optimal housekeeping genes to be used for specific grapevine tissues (González-Agüero et al., 2013), developmental stages (Reid et al., 2006) and variable physiological conditions (Luo et al., 2018; Selim et al., 2012). Some of these genes, obtained by cross-checking our RNA-seq data with the Corvina atlas, matched with the findings already available in the scientific literature regarding the ubiquitous expression of genic loci such as *ACTIN 7* (*VIT_04s0044g00580*) (Selim et al., 2012), *RIBOSOMAL PROTEIN 60S* (*VIT_18s0001g06410*) (Gamm et al., 2011), *ELONGATION FACTOR 1-alpha 1* (*VIT_06s0004g03220*) (Reid et al., 2006), *GAPDH* (*VIT_17s0000g10430*) (Reid et al., 2006) and *AQUAPORIN PIP2B* (*VIT_13s0019g04280*) (Reid et al., 2006). On the other side, an exhaustive and novel list of candidate control genes never considered before is presented here. Among the best 10 optimum housekeeping genes (in terms of transcripts abundance and comparable expression levels in all tissues of this study and in the Corvina atlas too; Supplementary Figure 4 and Supplementary Table 8) stand out four proteasome-related genes, genes involved in exocytosis (*RAB GTPase ARA3*), RNA transcription (*CTV.22*) and protein synthesis (*EIF-3E* and *EIF-4A3*) stands out a member of the Pollen Ole e 1 allergen and 70xtension (*AtPOE1*) family, initially identified as a group of allergens and recently recognized as developmental regulators in many plant tissues (Mandaokar and Browse, 2009). A further qPCR validation step is needed to evaluate whether these genes could represent a valid alternative to those commonly used in expression analyses.

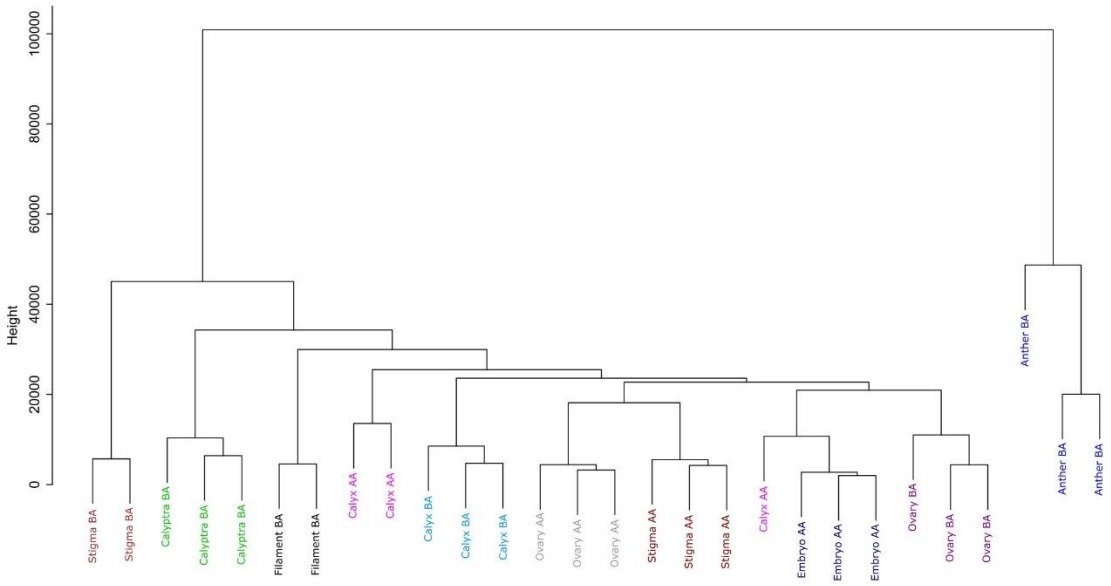


Figure S1: Cluster dendrogram based on raw counts (Supplementary Table 1) showing correlation among biological replicates of each tissue considered in this study.

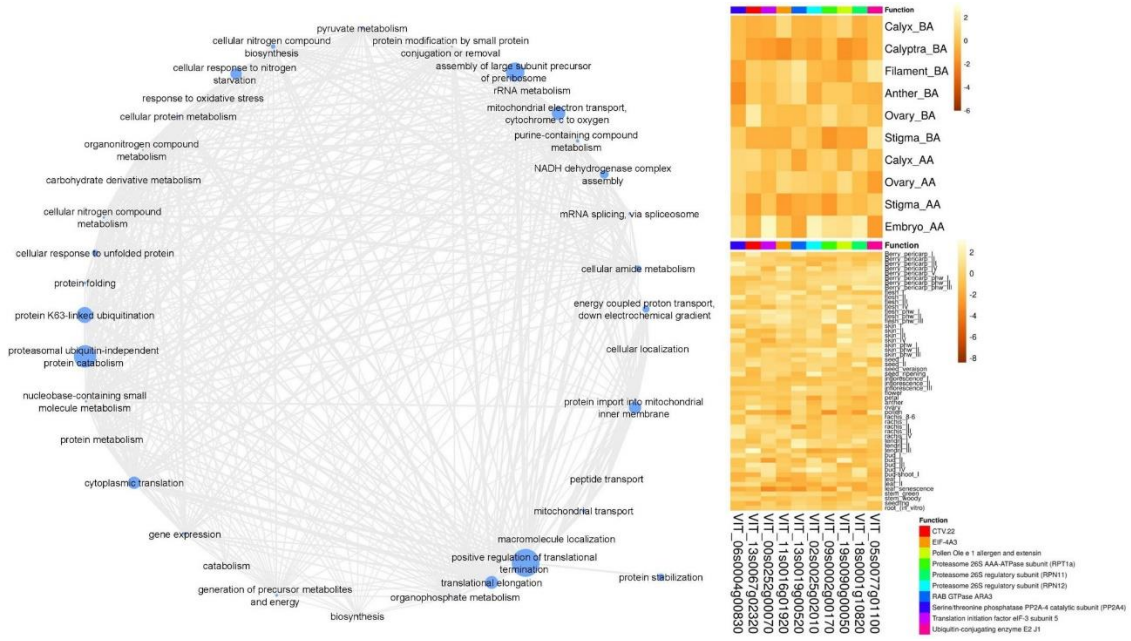


Figure S4: Main biological networks resulting from the GO enrichment analysis of the 662 housekeeping genes identified.

2.3.8 Conclusions

Although grapevine is a plant species mainly propagated by agamic way, nowadays, the understanding of the molecular mechanisms that lead to the ontogenetic determination of the different organs within the flower is a topic of great interest. The reasons behind this statement are many: i) the importance of conventional genetic improvement of varieties as well as rootstocks is increasingly affirming, in contrast to the historical and cultural legacies that see viticulture as an extremely conservative discipline; ii) flowering represents the first step of the reproductive phase that will lead to the development of the main vine product, the berry; iii) the structure of the flower and the architecture of the inflorescence influence the organization of the bunch, with many consequences at the production level; iv) the flower represents the main source of tissues (anthers and filaments) used for *in vitro* regeneration and, indirectly, for genetic improvement through new breeding techniques. In this study, we generated 30 grapevine RNA-seq datasets for different whorls and tissue of *V. vinifera* cv Pinot noir flowers in pre-anthesis (E-L 18, 8 days before anthesis) and post-anthesis (EL 26, 6 days after anthesis) stages and integrated them with a previously published grapevine cv Corvina expression atlas, which included several flower samples at different development stages (Fasoli et al., 2012). Both WGCNA and tau analysis were used to analyze the RNA-seq data and identify tissue-specific gene modules or marker genes and many of these were identified by both methods. Both analyses have advantages and disadvantages, depending on the objectives of the work and the biological questions to be answered. In this study, which has as its main objective the development of a transcriptomic reference for functional studies on flower-specific genes in grapevine, we tried to combine both approaches, identifying key hub genes which specifically characterize different flower organs before and after anthesis. Although this work is configured as a descriptive study at a wide genome level, the provision of numerous data relating to specific genes for every single tissue or whorls considered represents an important resource for the scientific community of the vine.

2.4 Materials and methods

2.4.1 Plant Material and Sample Collection

Flower materials (*V. vinifera* L. cv Pinot noir, clone 115, grafted onto Kober 5BB rootstock) were retrieved on May 2018 from a germplasm collection established in 2009 in the experimental farm "Lucio Toniolo" in Legnaro (University of Padova, Padova, Italy; 45°21'5,68"N 11°57'2,71"E). The soil texture was as follows: 46% sand, 24% clay, and 30% loam; pH = 7.9; electric conductivity, 112 µS; and organic carbon, 1.1%. Specifically, on May 14th (E-L 18, 8 days before anthesis) and May 28th (E-L 26, 6 days after anthesis), three inflorescences, each of which collected from an individual plant (1 inflorescence x 3 plants) were sampled and snap-frozen in liquid nitrogen. Each flower was then rapidly dissected in cold conditions into the relative whorls with the aid of a stereoscope and a scalpel. From pre anthesis flowers, calyptra (or cap), calyx, anther, filament, ovary and stigma were collected whereas after anthesis - being the cap and the male tissues released - the inflorescence was dissected into calyx, ovary, stigma and embryo. Considering both stages, ten tissues were isolated, each in three biological replicates (n = 30).

2.4.2 RNA purification, library preparation and sequencing

For each sample, approximately 50 mg of tissue were ground in liquid nitrogen, and total RNA was purified using the "Spectrum Plant Total RNA Kit" (Sigma-Aldrich, St. Louis, MO, USA) following the instruction provided by the manufacturer. The integrity of total RNA was checked on 1 % (w/v) agarose gel (Life Technologies, Carlsbad, CA, USA) stained with 1 x SYBR Safe DNA Gel Stain (Life Technologies) while the quality (in terms of 260/280 and 260/280 ratios) and the quantity were spectrophotometrically evaluated using NanoDrop-1000 (Thermo Scientific, Wilmington, MA, USA). RNA was stored at −80 °C until use.

cDNA libraries construction and sequencing were performed as described by Chitarrini et al. (2020). Briefly, 1 µg of total RNA was used to construct stranded mRNA-seq libraries (KAPA Stranded mRNA-Seq Kit, Kapa Biosystems, Woburn, MA, USA), that were later barcoded using the KAPA Dual-Indexed Adapter Kit (Kapa Biosystems). Libraries were then quantified (KAPA Library Quantification Kit, Kapa Biosystems) using a LightCycler 480 (Roche, Mannheim, Germany), checked in terms of correct size (250 – 280 bp) with a Tapestation 2200 (Agilent Technologies, Santa Clara, CA, USA) and High Sensitivity D1000 ScreenTape assay (Agilent Technologies) and, finally, multiplexed. Sequencing was performed on an Illumina HiSeq 2500 platform (Rapid Run Mode, Illumina, Inc.,

San Diego, CA, USA) to generate 2 × 250 bp reads. All raw reads were deposited in the NCBI SRA database with accession numbers SRR14777742 - SRR14777769.

2.4.3 RNA-seq analysis

FastQC software v.0.11.9 (Ewels et al., 2016) was used to summarize analysis results and to verify the overall quality of the sequencing output while fastp v.0.36 (Chen et al., 2018) was used to trim the Illumina adapters, merge the reads and filter the sequences based on phred quality score (removed if Q<30). Trinity software v2.8.5 (Haas et al., 2016) was used in *de novo* mode to assemble raw reads deriving from the 28 samples (two samples, one from stigma and one from filament were excluded) into a single reference catalogue as demanded by Salmon (Patro et al., 2017) for quantifying transcript abundance from RNA-seq reads. Trinity was run with default parameter by setting the minimum contig length to 200 and k-mer value at 25. The resulting catalogue was then annotated based on of the PN40024 12X v1 grapevine reference genome assembly (29,971 genes; Canaguier et al., 2017) and using the BLASTn algorithm (Altschul et al., 1990). Since the average number of raw reads produced per sample (~12 million) did not reach the minimum threshold required to estimate possible alternative transcripts (i.e., 30-60 million per samples; Illumina), all the putative isoforms (e.g., i1, i2, i3) produced by Trinity and therefore deriving from the same gene locus (e.g., VIT_18s0166g00210), were annotated under the same transcript name (e.g., VIT_18s0166g00210.01). To quasi-map and quantify RNA-seq reads with Salmon software v.0.14.1 (Patro et al., 2017), we built an index based on the newly assembled catalogue of flower transcripts. The 'decoys' option was used to build a *decoy-aware* index by employing the entire genome as the decoy sequence. The RNA-seq reads of each sample were then quantified and their abundance in terms of Transcripts Per Million (TPM) was calculated. As recommended when using the *decoy-aware* index, we used the 'validateMappings' option to mitigate potential spurious mapping of reads arising from unannotated genomic loci sequence-similar to annotated transcriptomic loci.

2.4.4 Weighted Gene Correlation Network Analysis (WGCNA)

In order to identify clusters (modules) of highly correlated genes attributable to a specific tissue, co-expression networks were constructed using the WGCNA 1.70-3 (https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/) package (Langfelder et all., 2008) in R-studio Version 1.3.1093, R version 4.0.3 (R Core Team). The analysis was performed on 19658 genes showing a mean TPM equal to or greater than 1 in at least one

tissue and variance higher than 1, while the remaining 10230 genes were filtered out. Parameters used in the analysis were set as follows: weighted network, signed; hierarchical clustering tree, Dynamic Hybrid Tree Cut algorithm; power = 12; minModuleSize = 30. As a first step in the analysis, a matrix of pairwise correlations between all genes across the 10 tissues was built. Then, the matrix was raised to a given soft-thresholding power based on the criterion of approximate scale-free topology and pickSoftThreshold function ($R^2 > 0.9$) to obtain an adjacency matrix. In order to identify modules of co-expressed genes, the topological overlap-based dissimilarity was constructed (Zhang et al., 2005; Ravasz et al., 2002) and used as input to perform the average linkage hierarchical clustering. Modules with highly correlated eigengenes were then merged (mergeCutHeight = 0.25). The association between merged modules and tissues/organs was tested calculating each module eigengene, defined as the first principal component of a PCA on the gene expression of all genes within the module. For each gene, total and intramodular connectivity (function softConnectivity), kME (for modular membership, also known as eigengene-based connectivity) and kME-*P* value were calculated, resulting in 20 tissue-specific modules. Genes belonging to each module were subjected to a GO enrichment analysis using the online tool ShinyGO (Ge et al., 2020).

2.4.5 Identification of tissue-specific genes and genes constitutively expressed in all floral tissues

The tissue-specificity level of each gene was calculated according to the tau (τ) algorithm (Yanai et al., 2005), which was demonstrated to be the best performing method to measure expression specificity in a benchmark study by Kryuchkova-Mostacci & Robinson-Rechavi (2017). Tau, whose values vary from 0 (broadly expressed) to 1 (tissue-specific), was calculated using the *tispec* R-package (https://rdrr.io/github/roonysgalbi/tispec). Data were firstly normalized removing all genes whose expression was < 1 TPM in any tissue and then, in order to make cross-tissue comparisons possible, a quantile normalization on the entire dataset was accomplished. Thereafter, for each tissue, a BIN value ranging from 0 (lowest expression) to 10 (highest expression) was attributed to each gene. The specificity of each gene (considering all tissues) was calculated implementing the τ algorithm:

$$\tau = \frac{\sum_{i=1}^{N} (1 - x_i)}{N - 1}$$

where N is the number of tissues and x_i is the expression value normalized by the highest expression.

Absolutely specific genes (ASGs) were defined as genes expressed in a single tissue only and indicated by a τ value of 1; highly specific genes (HSGs) were genes with relatively highly enriched expression in a few tissues and defined by a τ value of at least 0.85. Finally, genes were considered constitutively expressed in all floral tissues if τ value was < 0.2. The plotDensity function was then used to plot the tau value of every gene and visualize which tau values occur most often. Finally, for each tissue, the specificity of each gene was calculated as τ expression fraction (τ_{ef}):

$$\tau_{\text{ef}} = \tau\frac{qn}{max}$$

where *qn* is the quantile normalized expression and max is the highest quantile normalized expression. The function getTissue was used to retrieve the quantile normalized expression and tissue-specificity of every gene in each tissue and to create a score value between 0 and 2. This value represents the sum of its τ_{ef} value and its 0-1 ranged normalized expression value. Ranking of genes by both expression and specificity was used to identify a set of 10 optimal genes that were highly specific for a given tissue, highly expressed, and with minimal expression in other tissues. The online tool VENNY 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/) was used to highlight any possible overlapping among HSGs (resulting from the τ_{ef} analysis) of a given tissue and genes belonging to the cluster (resulting from the WGCNA analysis) significantly most associated with the tissue. Finally, a list of constitutive genes was drawn up retaining those loci that - in all tissues – scored a τ_{ef} value = 0 and a TPM value > 100.

2.4.6 Identification of *cis* and *trans* regulating factors in genes of interest

To identify genes coding for transcription factors, gene IDs identified by WGCNA and tau analysis were screened against the PlantTFDB (Jin et al., 2017). To identify cis regulatory element (CRE), the promoter sequences (2 Kb) of the specific genes selected for anther BA and stigma AA were retrieved from the 12x V1 annotation of PN40024 (Jaillon et al., 2007). The *de novo* identification of motifs enriched in the promoters of these genes was carried out using the DECOD software (Huggins et al., 2011), using as a background a collection of 10,000 promoter sequences obtained randomly from the grape genome. The analysis was performed using k-mers of 8 nucleotides (default parameter), a maximum of 10 motifs identified and 50 iterations. Once the enriched motifs were identified, they were submitted to the online tool STAMP (Mahony and Benos, 2007), in order to identify any CRE already characterized in previous studies.

DATA AVAILABILITY

Raw Illumina sequence data were deposited in the National Center for Biotechnology Information (NCBI) and can be accessed in the sequence read archive (SRA) database (https://www.ncbi.nlm.nih.gov/sra). The accession number is PRJNA736298 and includes 28 accession items (SRR14777742 - SRR14777769). All data generated or analyzed during this study are included in this published article and its supplementary information files.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Sara Sgubin who was of great help in the collection of the flower samples and in the dissection of the whorls and flower tissues on which the analyses were conducted and Dr. Luca Meloni who has kindly contributed to the drafting of Fig. 1 and Fig. 6. The authors are also grateful to Prof. Sara Zenoni, which provided transcriptional data from the *V. vinifera* cv Corvina atlas.

CONFLICT OF INTERESTS

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

CONTRIBUTIONS

GB and ML designed the research. AV, FP and GM conducted and controlled the experiments and analyzed the data. AV and FP carried out the bioinformatics analyses. GM performed GO and CRE analyses. FP and AV wrote the manuscript. All authors contributed to editing the manuscript.

2.5 Supplementary Information

Additional supporting information too big to be displayed here, please consult at the following link: https://doi.org/10.1038/s41438-021-00635-7

Supplementary Table 1: Mean (n = 3) transcript per million (TPM) counts in the ten tissues under study. Only those loci exhibiting TPM>1 in at least one sample were retained.

Supplementary Table 2: List of genes belonging to the 10 modules showing the best correlation with tissues under study. For each gene ID, gene significance (GS), module membership (MM) and functional annotation are reported along with the putative orthologs in Arabidopsis.

Supplementary Table 3: Detail of genes identified within enriched GO categories in WGCNA tissue-related modules.

Supplementary Table 4: List of TF coding genes (with MM > 0.9) identified in tissue-specific modules based on PlantTFDB. Together with Gene ID, the family of appurtenance, Functional annotation, grapevine gene name (when available based on previous studies), Module membership (MM) and TPM are reported. The "Analysis" column indicates the approach that led to the identification of gene. Gene written in red were found to be flower specific also in the Corvina expression atlas[19].

Supplementary Table 5: Tau values of genes in all tissues considered in this study.

Supplementary Table 6: List of ten best ranking genes based on a score value (0-2) corresponding to the sum of the quantile normalized expression of a given gene and its tau expression factor.

Supplementary Table 7: Genes in common between WGCNA, tau analysis and flower-specific genes in the Corvina expression atlas.

Supplementary Table 8: List of 662 genes (in ascending order) that exhibited a τ value = 0 and TPM values > 100 in all tissues.

Figure S2: Venn diagrams showing common and specific genes between the highly specific genes (HSG) identified by the tau analysis and the tissue-specific modules identified by WGCNA.

Figure S3: Top 10 best optimum ranking genes based on score values (0-2) for each tissue considered.

BIBLIOGRAPHIC REFERENCES

Altschul S. F., Gish W., Miller W., Myers E. W. & Lipman D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

Bartlett A. *et al.* Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* **12**, 1659–1672 (2017).

Boss P. K., Buckeridge E. J., Poole A. & Thomas M. R. New insights into grapevine flowering. *Funct. Plant Biol.* **30**, 593–606 (2003).

Bowman J., Alvarez, J., Weigel D., Meyerowitz E. & Smyth D. R. Control of flower development in Arabidopsis thaliana by APETALA1 and interacting genes. *Development* **119**, 721–743 (1993).

Canaguier A. *et al.* A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). *Genomics Data* **14**, 56–62 (2017).

Chen S., Zhou Y., Chen Y. & Gu J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

Cheng C. Y. *et al.* Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant J.* **89**, 789–804 (2017).

Chitarrini G. *et al.* Two-omics data revealed commonalities and differences between Rpv12- and Rpv3-mediated resistance in grapevine. *Sci. Rep.* **10**, 1–15 (2020).

Colombo M. *et al.* A new role for the SHATTERPROOF genes during Arabidopsis gynoecium development. *Dev. Biol.* **337**, 294–302 (2010).

Crawford B. C. W. & Yanofsky, M. F. Half filled promotes reproductive tract development and fertilization efficiency in Arabidopsis thaliana. *Development* **138**, 2999–3009 (2011).

Dal Santo S. *et al.* Grapevine field experiments reveal the contribution of genotype, the influence of environment and the effect of their interaction (G×E) on the berry transcriptome. *Plant J.* **93**, 1143–1159 (2018).

Ewels P., Magnusson M., Lundin S. & Käller M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

Fasoli, M. *et al.* The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. *Plant Cell* **24**, (2012).

Gamm M. *et al.* Identification of reference genes suitable for qRT-PCR in grapevine and application for the study of the expression of genes involved in pterostilbene synthesis. *Mol. Genet. Genomics* **285**, 273–285 (2011).

Ge S. X., Jung D., Jung D. & Yao R. ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).

Ghan R. *et al.* The common transcriptional subnetworks of the grape berry skin in the late stages of ripening. *BMC Plant Biol.* **17**, 94 (2017).

González-Agüero M. *et al.* Identification of two putative reference genes from grapevine suitable

for gene expression analysis in berry and related tissues derived from RNA-Seq data. *BMC Genomics* **14**, (2013).

Grimplet J., Martínez-zapater J. M. & Carmona M. J. Structural and functional annotation of the MADS-box transcription factor family in grapevine. *BMC Genomics* 1–23 (2016) doi:10.1186/s12864-016-2398-7.

Guo D. L., Wang Z. G., Pei M. S., Guo L. L. & Yu Y. H. Transcriptome analysis reveals mechanism of early ripening in Kyoho grape with hydrogen peroxide treatment. *BMC Genomics* **21**, 1–18 (2020).

Haas B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).

Higo K., Ugawa Y., Iwamoto M. & Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* **27**, 297–300 (1999).

Hiscock S. J., Doughty J. & Dickinson H. G. Synthesis and phosphorylation of pollen proteins during the pollen-stigma interaction in self-compatible Brassica napus L. and self-incompatible Brassica oleracea L. *Sex. Plant Reprod.* **8**, 345–353 (1995).

Hobo T. *et al.* Various spatiotemporal expression profiles of anther-expressed genes in rice. *Plant Cell Physiol.* **49**, 1417–1428 (2008).

Hooker T. S., Millar A. A. & Kunst L. Significance of the expression of the CER6 condensing enzyme for cuticular wax production in Arabidopsis. *Plant Physiol.* **129**, 1568–1580 (2002).

Hu W. & Ma H. Characterization of a novel putative zinc finger gene MIF1: Involvement in multiple hormonal regulation of Arabidopsis development. *Plant J.* **45**, 399–422 (2006).

Huang L. *et al.* The polygalacturonase gene BcMF2 from Brassica campestris is associated with intine development. *J. Exp. Bot.* **60**, 301–313 (2009).

Huang L. *et al.* BcMF9, a novel polygalacturonase gene, is required for both Brassica campestris intine and exine formation. *Ann. Bot.* **104**, 1339–1351 (2009).

Huggins P. *et al.* DECOD: Fast and accurate discriminative DNA motif finding. *Bioinformatics* **27**, 2361–2367 (2011).

Hyun Y. *et al.* Cooperation and Functional Diversification of Two Closely Related Galactolipase Genes for Jasmonate Biosynthesis. *Dev. Cell* **14**, 183–192 (2008).

Illumina Inc. Considerations for RNA-Seq read length and coverage. *Different RNA-Seq experiment types require different sequencing read lengths and depth* https://emea.support.illumina.com/bulletins/2017/04/considerations-for-rna-seq-read-length-and-coverage-.html (2021).

Ishiguro S. *et al.* SHEPHERD is the Arabidopsis GRP94 responsible for the formation of functional CLAVATA proteins. *EMBO J.* **21**, 898–908 (2002).

Jaillon O. *et al.* The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463–7 (2007).

Jin J. *et al.* PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions

in plants. *Nucleic Acids Res.* **45**, D1040–D1045 (2017).

Joubès J. *et al.* The VLCFA elongase gene family in Arabidopsis thaliana: Phylogenetic analysis, 3D modelling and expression profiling. *Plant Mol. Biol.* **67**, 547–566 (2008).

Krizek B. A. & Fletcher J. C. Molecular mechanisms of flower development: An armchair guide. *Nat. Rev. Genet.* **6**, 688–698 (2005).

Kryuchkova-Mostacci N. & Robinson-Rechavi M. A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.* **18**, 205–214 (2017).

Kunst L. & Samuels A. L. Biosynthesis and secretion of plant cuticular wax. *Prog. Lipid Res.* **42**, 51–80 (2003).

Langfelder P. & Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, (2008).

Langfelder P., Mischel P. S. & Horvath S. When Is Hub Gene Selection Better than Standard Meta-Analysis? *PLoS One* **8**, (2013).

Lebon G., Brun O., Magné C. & Clément C. Photosynthesis of the grapevine (Vitis vinifera) inflorescence. *Tree Physiol.* **25**, 633–639 (2005).

Li J., Yuan J. & Li M. Characterization of Putative cis-Regulatory Elements in Genes Preferentially Expressed in Arabidopsis Male Meiocytes. *Biomed Res. Int.* **2014**, (2014).

Li X. *et al.* Comparative transcriptomic analysis provides insight into the domestication and improvement of pear (P. pyrifolia) fruit. *Plant Physiol.* **180**, 435–452 (2019).

Libault M. *et al.* An integrated transcriptome atlas of the crop model Glycine max, and its use in comparative analyses in plants. *Plant J.* **63**, 86–99 (2010).

Liu H. *et al.* Genome-wide analysis and evolution of the bZIP transcription factor gene family in six Fragaria species Plant Syst Evol (2017) 303 : 1225-1237 nil OnęcMart Genome-wide analysis and evolution of the bZIP transcript factor gene family in six Fragana speci. **303**, 1225–1237 (2021).

Lou Y. *et al.* Characterization of transcriptional modules related to fibrosing-NAFLD progression. *Sci. Rep.* **7**, 1–12 (2017).

Lu S. *et al.* Isolation and functional characterization of the promoter of SEPALLATA3 gene in London plane and its application in genetic engineering of sterility. *Plant Cell. Tissue Organ Cult.* **136**, 109–121 (2019).

Luo M. *et al.* Selection of reference genes for miRNA qRT-PCR under abiotic stress in grapevine. *Sci. Rep.* **8**, 1–11 (2018).

Mahony S. & Benos P. V. STAMP: A web tool for exploring DNA-binding motif similarities. *Nucleic Acids Res.* **35**, 253–258 (2007).

Mandaokar A. & Browse J. MYB108 acts together with MYB24 to regulate jasmonate-mediated stamen maturation in Arabidopsis. *Plant Physiol.* **149**, 851–862 (2009).

Mayer K. F. X. *et al.* A physical, genetic and functional sequence assembly of the barley genome.

Nature **491**, 711–716 (2012).

Mayfield J. A., Fiebig A., Johnstone S. E. & Preuss D. Gene families from the Arabidopsis thaliana pollen coat proteome. *Science (80-.).* **292**, 2482–2485 (2001).

Millar A. A. *et al.* CUT1, an Arabidopsis gene required for cuticular wax biosynthesis and pollen fertility, encodes a very-long-chain fatty acid condensing enzyme. *Plant Cell* **11**, 825–838 (1999).

Mukoko Bopopi J. *et al.* Ectopic expression of PtaRHE1, encoding a poplar RING-H2 protein with E3 ligase activity, alters plant development and induces defence-related responses. *J. Exp. Bot.* **61**, 297–310 (2010).

OIV. International Organization of Vine and Wine. (2016).

Palumbo F., Vannozzi, A., Magon, G., Lucchin, M. & Barcaccia, G. Genomics of flower identity in grapevine (Vitis vinifera L.). *Front. Plant Sci.* **10**, 1–15 (2019).

Palumbo M. C. *et al.* Integrated network analysis identifies fight-club nodes as a class of hubs encompassing key putative switch genes that induce major transcriptome reprogramming during grapevine development. *Plant Cell* **26**, 4617–4635 (2014).

Papatheodorou I. *et al.* Expression Atlas: Gene and protein expression across multiple studies and organisms. *Nucleic Acids Res.* **46**, D246–D251 (2018).

Patro R., Duggal G., Love M. I., Irizarry R. A. & Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

Pervaiz T. *et al.* Transcriptomic analysis of grapevine (cv. Summer Black) Leaf, using the illumina platform. *PLoS One* **11**, e0147369 (2016).

Potenza E. *et al.* Exploration of alternative splicing events in ten different grapevine cultivars. *BMC Genomics* **16**, 706 (2015).

Preuss D. Sexual signaling on a cellular level: Lessons from plant reproduction. *Mol. Biol. Cell* **13**, 1803–1805 (2002).

Provart N. The Bio-Analytic Resource. *Gene expression and protein tools* (2020).

R Core Team. R: a language and environment for Statistical computing. (2020).

Ravasz E., Somera A. L., Mongru D. A., Oltvai Z. N. & Barabási A. L. Hierarchical organization of modularity in metabolic networks. *Science (80-.).* **297**, 1551–1555 (2002).

Reid K. E., Olsson N., Schlosser J., Peng F. & Lund S. T. An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. *BMC Plant Biol.* **6**, 1–11 (2006).

Rhee S. Y., Osborne E., Poindexter P. D. & Somerville C. R. Microspore Separation in the quartet 3 Mutants of Arabidopsis Is Impaired by a Defect in a Developmentally Regulated Polygalacturonase Required for Pollen Mother Cell Wall Degradation. *Plant Physiol.* **133**, 1170–1180 (2003).

Roque E., Gómez-Mena C., Ferrándiz C., Beltrán J. & Cañas L. Functional genomics and genetic control of flower and fruit development in Medicago truncatula: an overview. in *Functional*

Genomics in Medicago truncatula - Methods and Protocols (eds. Cañas, L. A. & Beltrán, P. J.) 273–290 (Humana Press, 2018).

Rundle S. J., Nasrallah M. E. & Nasrallah J. B. Effects of inhibitors of protein serine/threonine phosphatases on pollination in Brassica. *Plant Physiol.* **103**, 1165–1171 (1993).

Sato S. *et al.* The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**, 635–641 (2012).

Selim M. *et al.* Identification of suitable reference genes for real-time RT-PCR normalization in the grapevine-downy mildew pathosystem. *Plant Cell Rep.* **31**, 205–216 (2012).

Seo Y. S., Kim E. Y., Kim J. H. & Kim W. T. Enzymatic characterization of class I DAD1-like acylhydrolase members targeted to chloroplast in Arabidopsis. *FEBS Lett.* **583**, 2301–2307 (2009).

Severin A. J. *et al.* RNA-Seq Atlas of Glycine max: A guide to the soybean transcriptome. *BMC Plant Biol.* **10**, 160 (2010).

Shahan R. *et al.* Consensus coexpression network analysis identifies key regulators of flower and fruit development in wild strawberry. *Plant Physiol.* **178**, 202–216 (2018).

Shu Q. *et al.* Analysis of the formation of flower shapes in wild species and cultivars of tree peony using the MADS-box subfamily gene. *Gene* **493**, 113–123 (2012).

Stelpflug S. C. *et al.* An Expanded Maize Gene Expression Atlas based on RNA Sequencing and its Use to Explore Root Development. *Plant Genome* **9**, 1–16 (2016).

Sun Y., Fan Z., Li X., Li J. & Yin H. The APETALA1 and FRUITFUL homologs in Camellia japonica and their roles in double flower domestication. *Mol. Breed.* **33**, 821–834 (2014).

Vannozzi A., Dry I. B., 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.* **12**, (2012).

Venturini L. *et al.* De novo transcriptome characterization of Vitis vinifera cv. Corvina unveils varietal diversity. *BMC Genomics* **14**, 41 (2013).

Vogler F., Schmalzl C., Englhart M., Bircheneder M. & Sprunck S. Brassinosteroids promote Arabidopsis pollen germination and growth. *Plant Reprod.* **27**, 153–167 (2014).

Wang B. *et al.* A comparative transcriptional landscape of maize and sorghum obtained by single-molecule sequencing. *Genome Res.* **28**, 921–932 (2018).

Wang F. *et al.* A global analysis of the polygalacturonase gene family in soybean (glycine max). *PLoS One* **11**, 1–23 (2016).

Wang L. *et al.* Floral transcriptomes reveal gene networks in pineapple floral growth and fruit development. *Commun. Biol.* **3**, (2020).

Wang Y. & Li R. Grape skin transcriptome to reveal genes related with resveratrol accumulation. (2015).

Wong D. C. J. *et al.* A systems-oriented analysis of the grapevine R2R3-MYB transcription factor

family uncovers new insights into the regulation of stilbene accumulation. *DNA Res.* **00**, dsw028 (2016).

Xu X. *et al.* Genome sequence and analysis of the tuber crop potato. *Nature* **475**, 189–195 (2011).

Yanai I. *et al.* Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650–659 (2005).

Yang Y., Yu Y., Liang Y., Anderson C. T. & Cao J. A profusion of molecular scissors for pectins: classification, expression, and functions of plant polygalacturonases. *Front. Plant Sci.* **9**, 1–16 (2018).

Zhang B. & Horvath S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, (2005).

Zhang Q., Huang L., Liu T., Yu X. & Cao J. Functional analysis of a pollen-expressed polygalacturonase gene BcMF6 in Chinese cabbage (Brassica campestris L. ssp. chinensis Makino). *Plant Cell Rep.* **27**, 1207–1215 (2008).

CHAPTER III: The DAP-Seq as item to explore grapevine cistrome

3.1 Introduction

In the post-genomics Era, the enormous amount of data generated by omics technologies being stored in public databases considerably exceeds the analytical capacities of humans, making it imperative to use increasingly powerful computational resources to process, analyze and store this information. Nowadays, several studies aim at specifically addressing this issue in grapevine (*V. vinifera* L.), a plant which is quickly establishing itself as an appealing 'model' species for studying non-climacteric fleshy fruits. Indeed, the increasing amount of data being continuously generated within the grapevine community, after this species genome was sequenced and released more than ten years ago, has certainly helped in this appointment. Genomics resources for *Vitis* species have increased promptly within the last fifteen years, beginning with the sequencing of expressed sequence tags (ESTs. Da Silva et al., 2005; Moser et al., 2005). These resources allowed us to quantitatively assess the grape transcriptome by aiding the development of cDNA and oligonucleotide microarrays (Terrier et al., 2005; Waters et al., 2005). However, after the concomitant release of two *V. vinifera* genome sequences, one of a near homozygous genotype closely related to Pinot Noir, PN40024 (Jaillon et al., 2007) and one of the cv. 'Pinot Noir' (Velasco et al., 2007), that a burst of new transcriptomic technologies emerged for this species. The reference grape genome (PN40024), currently on its third assembly (PN40024.v4) and its fourth annotation (PN40024.v4.2) comprises 35,230 genes (Velt et al., 2022). In addition to Pinot Noir, there has recently been a proliferation of projects aimed at sequencing the genomes of other grapevine accessions (*i.e.,* cultivars), some of which are based on third generation sequencing technologies. One of the main advantages of such technologies based on long-reads, such as PacBio, is to retain haplotype information after using intense-computing methods (*e.g.,* FALCON unzip method, Chin et al., 2016). At present, the PN40024 reference genome is being updated using this technology, in order to retain information from both haplotypes in the remaining heterozygous regions. In addition to genomics and transcriptomics data, the recent development of regulatory genomics (also known as regulomics) has gained much of the focus in crop species since transcriptional regulation of genes plays a very important role in phenotypic plasticity and quality of fruits (Dal Santo et al., 2013). Transcriptional regulation can be accomplished through several processes, including epigenetic modifications, microRNA and transcription factor mediated control of gene expression. In this regard, DNA-Affinity Purification Sequencing (DAP-Seq), a high-throughput TFs binding site discovery method, interrogates genomic DNA with *in vitro*-expressed TFs (Bartlett et al., 2017). Despite ChIP-seq which is limited in scale as it is complicated to execute,

depends on antibody quality and is rare for low abundant TFs, DAP-Seq is the ultimate approach for TF binding site discovery and provides a scalable alternative for non-conventional model species where genetic transformation is difficult, to rapidly and inexpensively interrogate large number of TFs. The idea behind this assay is to combine next-generation sequencing of a genomic DNA library with *in vitro* expression of affinity-purified TFs to generate cistrome maps for a wide range of species. It basically works by generating the 200 bp DNA fragments ligated to the adaptors for Illumina platform sequencing, which constitute the library to which the *in vitro* expressed TF protein is hybridized. So, after several washing steps and the ligation of some indexes for multiplexing, the fragments can be processed at Illumina platform (Fig. 1.3).

In this regard, we took advantage of this assay to investigate the cistrome landscape of a representative subset of 17 whorls specific TFs individuated in Chapter II: MYB91A, MYB108A, MYB143, MYB145, MYB148, MYB150, MYB154, MYB192, MYBA7, MYBA8, MYBPA7, MYBPA9, MYBPAL3, NAC62, WRKY15, WRKY22 and WRKY42. The challenge was not only to fit the DAP-Seq protocol itself by adapting it to a different species from the one on which the protocol was designed, but also to test several gDNA extraction methods. Since the subsequent passages are widely affected by gDNA quality, it is pivotal to define a good nucleic acid extraction methodology to obtain the maximum efficiency by the downstream applications. Several protocols were tested to find out the one most suitable for our purposes, which can conciliate quantitative yield and qualitative purity. For this purpose, a double quantification was achieved, by checking with Nanodrop (Thermo Fisher) and QuBit (Thermo Fisher). Since the QuBit is based on fluorescence issued by a dye specifically bound to a certain analyte, the quantification is more accurate than the one made with Nanodrop, which measures can be easily invalidated from contaminants absorbing at the same wavelength of nucleic acid. Due to the different technologies on which the two instruments rely, the more the measurements are similar, the more efficient is the extraction method and the final eluted pure. After these considerations, libraries were constructed using the obtained gDNA from young leaves of *V. vinifera* cv Cabernet Franc. The ORF of each TF was transferred into the Gateway-compatible pIX-HALO expression vector, which contains an N-terminal HALO affinity tag sequence. HALO tag is a self-labeling protein tag made of 297 residue protein (33 kDa) derived from a bacterial enzyme, designed to covalently bind to a synthetic ligand (Los et al., 2008). The pIX-HALO-TF constructs were then expressed using a rabbit reticulocyte system enabling the *in vitro* transcription and translation coupled reactions. The achieved gDNA libraries were hybridized with HALO-tagged *in vitro* expressed TFs, and TF-DNA complexes were purified using magnetic separation of the affinity tag. The bound gDNA was eluted and sequenced. Sequence reads were mapped to the PN40024

reference genome, identifying genome-wide binding sites for each TF. The use of a negative control sample (pIX-HALO-GST vector) dramatically reduced background noise.

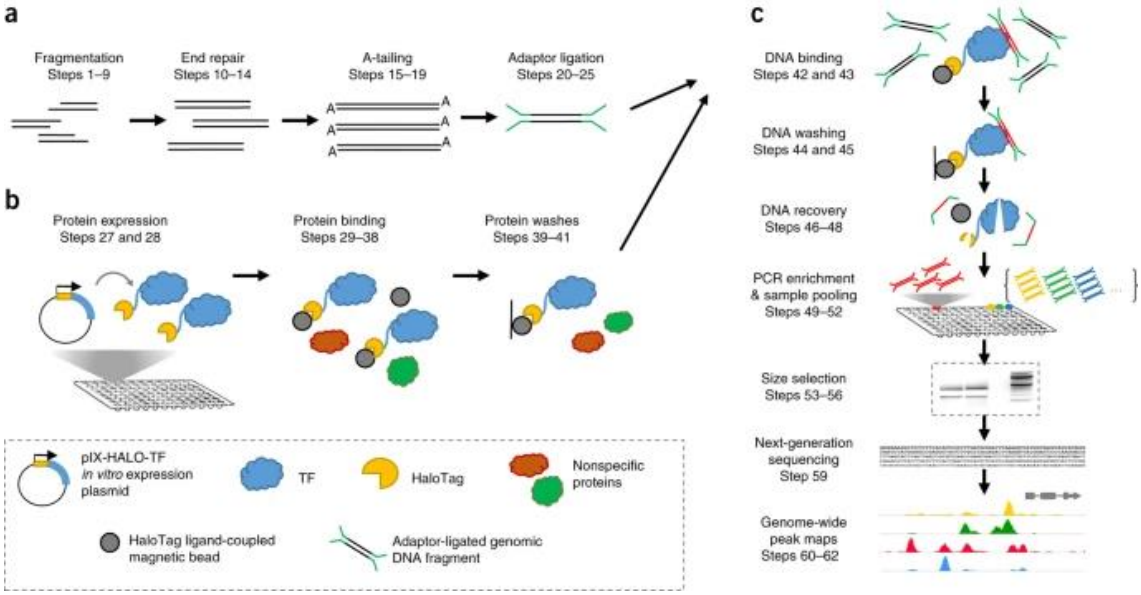


Figure 1.3. General overview of DAP-Seq procedures workflow (A) An adaptor-ligated DNA library is prepared by shearing native genomic DNA into ~200-bp fragments and ligating Illumina-based sequencing adaptors to the repaired ends. (B) Transcription factor (TF) ORF clones fused to the Halo affinity tag are expressed *in vitro* and bound to ligand-coupled beads, whereas nonspecific proteins are washed away. (C) HaloTag-TF fusion proteins are incubated with an adaptor-ligated genomic DNA library, and unbound DNA fragments are washed away. Samples are heated to release TF-bound DNA, and the recovered DNA is PCR-amplified to attach indexed sequencing primers. Indexed DNA samples are subsequently combined and size-selected to remove residual adaptor dimers. Purified DNA libraries are sequenced using next-generation sequencing, and the resulting genome-wide binding events are analyzed (picture taken from Bartlett et al., 2017)

3.2 Material and methods

3.2.1 Alleles frequencies screening

In order to select the most representative allele on which to base the ORF of each TF construct, it was decided to analyze the allelic frequency of the selected TFs by analyzing the genome of 131 accessions of *V. vinifera*, basing on Magris et al. (2021). For each of the selected TFs, the sequence was extracted from PN40024 reference genome and for each of them it was calculated the allelic frequency on the total of the considered population. If that allelic frequency value was greater or equal to the threshold of 0.5, it was considered as sufficiently representative for the whole species and therefore that sequence was kept and used to generate the entry vector. In case of allelic frequency less than 0.5, the ORF was compiled by changing the nucleotide, evaluating the polymorphic position of the Cabernet Franc specific case.

3.2.2 Plasmids preparation

All the TFs ORFs were synthesized directly in pDONR221 vector by ordering to GeneArt service (Thermo Fisher) and recombined in pIX-HALO vector by using Gateway LR Clonase II Enzyme Mix (Thermo Fisher) setting up the following protocol as provided by manufacturer:

I. 50 µl reaction setup:
   - 1-7 µl of entry clone (total amount of 150 ng of pDONR221_ORF)
   - 1 µl of empty pIX-HALO (150 ng)
   - Up to 8 µl of TE buffer
   - 2 µl of LR Clonase II enzyme mix

II. Incubate at 25°C for 1 hour

III. Add 1 µl of Proteinase K and incubate at 37°C for 10 minutes to inactivate the reaction

3.2.3 Preparation of chemically competent *Escherichia coli* strain TOP10 cells (modified from Dagert and Ehrlich, 1979)

REAGENTS

Luria-Bertani broth (LB, pH7)	
Reagent	**Concentration (g/L)**
Triptone	10
NaCl	10
Yeast Extract	5
milliQ H_2O	up to volume

- 100 mM $MgCl_2$
- 100 mM $CaCl_2$ + 15% glycerol

PROTOCOL

I. Pick up a single colony of *Escherichia coli* and grow it in 5 ml LB at 37°C by constant shaking at 200 rpm overnight
II. Dilute 1:100 in LB and let grow at 37°C by constant shaking at 200 rpm
III. After about 2 hours, the OD_{600} should be around 0.5-0.6 and after having checked the cells have reached the proper density, transfer the culture to centrifuge bottles and spin at 5000 rpm for 10 minutes at 4°C
IV. Remove supernatant and keep the cells on ice
V. Resuspend in 100 ml of ice cold 100 mM $MgCl_2$ and homogenize by pipetting
VI. Incubate on ice for 20-30 minutes and spin at 4000 rpm for 10 minutes for 4°C
VII. Discard supernatant and resuspend in 10 ml of 100 mM $CaCl_2$ + 15% glycerol
VIII. Aliquot 40 µl into 1.5 ml Eppendorf tubes and store at -80°C

3.2.4 *Escherichia coli* transformation and plasmid collection

I. Add 5 µl of LR reaction obtained as described in 3.2.2 pIX-HALO cloning paragraph to 40 µl of chemically competent *Escherichia coli* cells and mix gently not by pipetting up and down
II. Incubate on ice for 30 minutes
III. Heat-shock in 42°C bath for 30 seconds and back in in ice for 2 minutes
IV. Add 200 µl of LB and incubate at 37°C for 1 hour with constant shaking at 200 rpm
V. Plate on a 1.5% agarose LB Petri + Ampicillin (100 mg/ml)
VI. Incubate overnight at 37°C
VII. Assess the presence of the plasmid with a colony-PCR. In our case, M13 generic primers were used to confirm the presence of the predicted size amplicon (M13 forward: 5'-GTAAAACGACGGCCAG-3'; M13 reverse: 5'- CAGGAAACAGCTATGAC-3')
VIII. Pick up a single transformed colony, put in 5 ml of LB + Ampicillin (100 mg/ml) and incubate overnight at 37°C with constant shaking at 200 rpm
IX. Plasmid purification miniprep using PureLink Quick Plasmid Miniprep Kit (Thermo Fisher) and sequencing to assess the correctness of the vector

3.2.5 gDNA extraction method

After having tested many gDNA extraction protocols, it was identified as the one described by Thomas et al., (1993) as the most suitable for our purposes. Anyway, it was necessary to make some adaptations in order to optimize it on grapevine leaf tissue. The final protocol is reported below.

REAGENTS

Crude Nuclei Buffer (CNB) pH7,6	
Reagent	**Concentration**
NaCl	0,25M
Tris	0,2M
Na_2EDTA	0,05M
Polyvinylpyrrolidone	2,5%
β-mercaptoethanol*	0,1%

DNA Extraction Buffer (DEB) pH8,0	
Reagent	**Concentration**
NaCl	0,5M
Tris	0,2M
EDTA	0,05%
Polyvinylpyrrolidone	2,5%
N-Lauroyl-Sarcosine	3%
Ethanol*	20%
β -mercaptoethanol*	1%

*Add just before use

PROTOCOL

I. Collect 2 g of leaf tissue, grind it as extremely fine powder in a mortar with liquid nitrogen and put it in a 50 ml falcon tube
II. Add 25 ml of CNB and homogenize by vortexing
III. Spin at 4000 rpm for 10 minutes at 4°C
IV. Discard supernatant, add 10 ml of DEB and homogenize by vortexing having care to entirely disrupting the pellet
V. Incubate at 55-60 °C for 30 minutes with a constant 100-150 rpm shaking
VI. Add 1 volume of Chloroform : Isoamyl alcohol (24:1) solution and vortex well to mix
VII. Spin at 5000 rpm for 15 minutes at room temperature (RT) and transfer the upper phase in a new 50 ml falcon tube having care not to disturb the pellet
VIII. Repeat steps 6 and 7

IX. Add 1 volume of Isopropanol, mix well by inverting the tube and leave at -20 °C for 10-15 minutes
X. Spin at 5000 rpm for 15 minutes at 4°C
XI. Pour off the supernatant and transfer the pellet in a 2 ml Eppendorf tube
XII. Add 500 µl of 70% ethanol
XIII. Spin at maximum speed for 10 minutes
XIV. Remove supernatant with a pipette by having care not to leave even a drop inside the tube
XV. Let airdry the pellet under a hood for 2-3 minutes being careful not to over dry it
XVI. Resuspend in 600 µl of EB and let overnight at 4°C for the pellet to dissolve with 6 µl RNase A (10 mg/ml)
XVII. Add 0.5 volume of 7.5M NH_4OAc
XVIII. Spin at max speed for 10 minutes at RT
XIX. Transfer the supernatant to a new 1.5 ml Eppendorf tube
XX. Add 1 volume of Isopropanol and mix well by inverting the tube
XXI. Spin at maximum speed for 10 minutes at 4°C
XXII. Remove supernatant, add 250 µl of 70% ethanol and spin at maximum speed for 10 minutes at RT
XXIII. Remove supernatant with a pipette by having care not to leave even a drop inside the tube
XXIV. Let airdry the pellet under a hood for 2-3 minutes being careful not to over dry it
XXV. Resuspend pellet in 150 µl of EB (Elution Buffer, Qiagen)

The final concentrations were checked both at Nanodrop and QuBit and were observed to perfectly collimate in both the quantification methods. The final gDNA yield estimated with these conditions is about 30-65 µg per g of ground leaf tissue.

3.2.6 gDNA library preparation and DAP

1) gDNA Fragmentation

1. By always basing on QuBit measurement, dilute 7.5 µg of gDNA in a final volume of 130 µl of EB and transfer in a Covaris tube (microtube AFA Fiber Pre-Slit Snap-Cap 6x16 mm, Covaris)
2. Our recommendation is to use low bind Eppendorf tubes for all the process and to transfer the gDNA in Covaris tube just immediately before the fragmentation process and to put it back in a new 1.5 ml Eppendorf tube immediately after finishing it. Always

work on ice. Set up the following conditions on a Covaris M220 Focused-ultrasonicator in order to obtain 200 bp fragments, as reported in manufacturer's guide:

- Target (peak): 200 bp
- Peak Incident Power (W): 50
- Duty Factor: 20%
- Cycles per Burst: 200
- Treatment time: 150 seconds
- T_{max} = 8°C, T_{min} = 4°C, $T_{setpoint}$ = 6°C

3. Check the fragmentation has occurred correctly (Fig.2.3). In this regard, we used a Tape Station 4150 (Agilent)

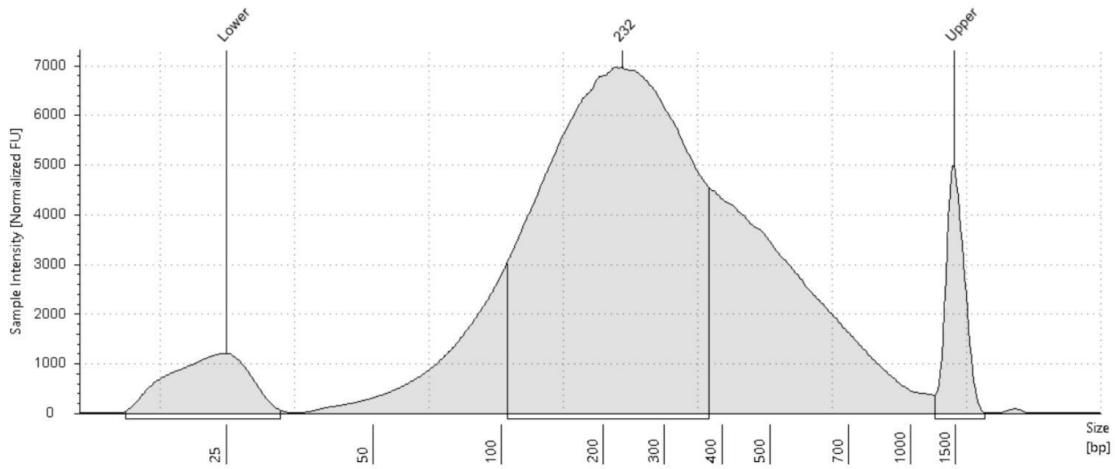


Figure 2.3. Graph showing the fragments size. The two extern peaks named as "Lower" and "Upper" show the run limits and are due to technical buffers. The middle peak shows the correct size of the fragments obtained with ultrasonication procedure.

4. In order to achieve fragments purification, the best way to proceed is to use the magnetic beads AMPure XP (Beckman-Coulter) which – due to its combined action of cleaning and size selection efficiency on 200 bp fragments – is highly recommended in NGS library preparation:
   I. Add 1 volume of AMPure XP beads, vortex and let incubate at RT for 10 minutes
   II. Position on magnetic rack (DynaMag-2, Thermo Fisher) for 5 minutes
   III. Remove supernatant with a pipet. In this step is better to let a little bit of liquid rather than accidentally carryover a little bit of beads for aspiring all the supernatant
   IV. Add 500 µl of freshly prepared 80% ethanol. Remove the tube from the magnetic rack, rotate it 180 degrees and position it back. Repeat for 6 times. This movement will make the beads to migrate from one wall to the other of

the Eppendorf tube and to purify through the scrolling movement inside the ethanol solution. Wait 1 minute, remove ethanol solution, and repeat this step by having care to remove entire volume till the last drop after the second wash

V. Let the tube with the lid opened and dry at 37°C, until in the beads stain are visible dryness cracks (about 10 minutes)
VI. Resuspend the beads in 35 µl of EB mixing well by pipetting
VII. Incubate at RT for 10 minutes to let the beads release fragments in the solution
VIII. Place the tube on the magnetic rack for 3 minutes and subsequently transfer the supernatant in a new 1.5 ml Eppendorf tube being extremely careful not to carryover beads
IX. Check the concentration on QuBit with Qubit dsDNA BR Assay Kit (Thermo Fisher): it is common to have loss till 50% of the starting 7.5 µg of gDNA

2) End Repair

Since, due to sonication, gDNA fragments ends do not have full nucleotide pairing between the two strands (so-called “sticky ends”), it is necessary to repair them by removing overhangs in 3’ with a 3’->5’ activity exonuclease and fulfill the ones in 5’ with a polymerase, obtaining in this way the “blunt ends”, or rather fragments ending with a coupled nucleotide pair (End-It DNA End-Repair Kit Manual). To achieve this, we used the End-It kit (Lucigen) as follows:

I. 50 µl reaction setup:
    - gDNA fragments template: 34 µl
    - 10x Buffer: 5 µl
    - dNTPs mix: 5 µl
    - ATP: 5 µl
    - End-It enzyme: 1 µl
II. Incubate at 22 °C for 45 minutes
III. Perform purification with Qiaquick PCR purification kit (Qiagen) as follow:
    - Add 5 volumes of Buffer PB and 10 µl of 3M NaOAc pH5.5 to each 50 µl end repaired sample
    - Mix by pipetting and charge to purple column
    - Spin at maximum speed for 1 minute and throw the flow through
    - Add 750 µl of Buffer PE and incubate at RT for 2 min
    - Spin at maximum speed for 1 minute, throw the flow through and spin at maximum speed for 2 minutes to remove residuals
    - Transfer the column to a new 1.5 ml Eppendorf tube

- Add 33 µl of EB, incubate 5 minutes at RT and Spin at maximum speed for 1 minutes

3) A-tailing

A-tailing is an enzymatic method for adding a non-templated adenine to the 3' end of a blunt, double-stranded DNA molecule and it is needed for the subsequent passage of adapter ligation (A-tailing Application Overview, NEB). The passages using End-It kit (Lucigen) are the following:

I. 50 µl reaction setup:
   - gDNA blunted fragments template: 32 µl
   - NEBuffer2: 5 µl
   - 1mM dATP: 10 µl
   - Klenow (3'-5' exo-, NEB): 3 µl

II. Incubate at 37°C for 30 minutes

III. Perform purification with Qiaquick PCR purification kit as described in step III of 2) End Repair section and elute in a final EB volume of 32 µl

4) Adapter Ligation

Adapter ligation is a process through which synthetic oligonucleotides are ligate to the fragments allowing to bind the sequencing platform flow cell. In this case were used the truncated Illumina TruSeq adapter, also called "Y-adapter", whose sequences are:

- Adapter strand A: 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'
- Adapter strand B: 5'-P-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-3' (where 'P' indicates a 5' phosphate group)

For the ligation reaction was used the T4 DNA Ligase (NEB) with the following steps:

I. 50 µl reaction setup:
   - gDNA A-tailed fragments template: 30 µl
   - 10x Ligase Buffer: 5 µl
   - 30 µM Y-adapter: 10 µl
   - T4 DNA Ligase (NEB): 5 µl (better to add individually to each reaction to avoid Y-adapter dimers)

II. Incubate at 16°C overnight

III. Incubate at 70°C for 10 minutes to achieve enzymatic inactivation

IV. Proceed with AMPure XP beads purification as described in step 4 of 1) gDNA Fragmentation section and elute in a final EB volume of 32 µl. This step is required to get rid of unligated adapters and adapters dimers

V. Check the concentration on QuBit with Qubit dsDNA BR Assay Kit (Thermo Fisher): it is common to have loss till 50% of the amount measured in step IX of the passage 4 in 1) gDNA Fragmentation section

5) *In vitro* TF protein expression

Due to the ease of execution and the possibility to perform transcription and translation within the same reaction, we chose TNT Coupled Reticulocyte Lysate Systems (Promega) as an in *vitro* protein expression system. Since each TNT Coupled Reticulocyte Lysate tube contains the amount for 8 reactions, our recommendation is to express and manage 8 TFs proteins at a time, in order to avoid freeze thawed. Passages are the following:

I. 8x reaction setup:
   - TNT Coupled Reticulocyte Lysate: 200 µl
   - TNT Reaction Buffer: 16 µl
   - Minus Met 1mM: 8 µl
   - Minus Leu 1mM: 8 µl
   - RNasin Ribonuclease Inhibitor: 8 µl
   - milliQ H_2O: 72 µl

II. Homogenize the mastermix by pipetting being careful not to produce bubbles since the high viscosity of the mix does not allow to remove them

III. Aliquot 40 µl of the mastermix into each tube of an 8 wells strip tube

IV. Dilute 1 µg of pIX-HALO_ORF in a final volume of 10 µl of EB and add it to the mix (final volume of 50 µl per single reaction)

V. Incubate at 30°C for 2 hours. After the incubation, if it is not possible to go on with the following steps, let the expressed TFs at RT for a little while because – if placed on ice or at 4°C – a precipitate will form. We recommend saving 5 µl of each TNT reaction to eventually check the correct expression of the protein on a Western Blot. Store this aliquot at -20°C.

6) DNA Affinity Purification

REAGENTS

PBS (pH7,4)

Reagent	Quantity (g/L)
NaCl	8
KCl	0,2
Na_2HPO_4	1,44
KH_2PO_4	0,24

PBST buffer: 1 volume of PBS + 0.025%NP40 (NP40 Surfact-Amps Detergent, Thermo Fisher)

PROTOCOL

I. Put 80 µl of Magne Halo-Tag Beads (Promega) in a 1.5 ml Eppendorf tube and add 1 ml of PBST. Magne Halo-Tag Beads provide a convenient method to covalently capture HaloTag fusion proteins with magnetic particles for protein pull-downs and purification. These magnetic beads offer a high binding capacity (≥20mg/ml) for purified Halo-Tag fusion proteins with low nonspecific protein binding. The magnetic handling properties allow streamlined protein purification and eliminate the need for multiple centrifugation steps (Magne Halo-Tag Beads Technical Manual, Promega)

II. Mix by pipetting and place on magnetic rack, remove supernatant and add 1 ml of PBST buffer

III. Remove the tube from the magnetic rack, rotate it 180 degrees and position it back. Repeat for 6 times. Wait 1 minute, remove supernatant, and repeat from step I.

IV. Resuspend beads in a final volume of 360 µl of PBST buffer (45 µl per reaction)

V. Add 45 µl of washed Magne-HaloTag beads to each 45 µl TNT reaction and mix well by pipetting. We advise to try to introduce a single big bubble in every well: this allows a correct and homogeneous movement of the beads along all the protein solution in the next step and avoids the stratification on the cap or in the bottom of the tube. This last situation must be absolutely avoided because otherwise the proteins will be bound only by the beads in the interface area between stratification and solution, losing in this way most of the transcription factor just expressed

VI. Rotate at RT for 1 hour to let the Halo-Tag beads to bind the protein. The rotation movement must be circular and slow, in order to obtain a constant and complete migration of the beads along the solution

VII. Place the strip on a PCR tube magnet, remove supernatant, take off the strip from the magnet, add 90 µl of PBST buffer in each tube, let the beads settle down, put again on the magnet and remove the supernatant. Repeat this entire step for 4 times and resuspend in a final volume of 40 µl of PBST buffer

VIII. Dilute 500 ng of Y-adapter ligated gDNA library in a final volume of 40 µl of EB per reaction, add it to each 40 µl PBST resuspended beads gained in the previous step. Let rotate at RT for 1 hour as described in step VI.

IX. Place the strip on a PCR tube magnet, remove supernatant, take off the strip from the magnet, add 90 µl of PBST buffer in each tube, let the beads settle down, put again on the magnet and remove the supernatant. Repeat this entire step for 6 times. After the 5th wash, transfer each 90 µl reaction to a new 8 wells strip and resuspend in a final volume of 25 µl of EB and transfer to a new 8 wells strip

X. Heat at 98°C for 10 minutes to release DNA and denature proteins

XI. Incubate on ice for 5 minutes

XII. Place the strip on a PCR tubes magnet and transfer supernatant in a new 8 wells strip

7) PCR Enrichment

This step is required to add the sequences necessary for sequencing on Illumina platform and to attach indexes for samples multiplexing by using single or dual barcode indexes. It is fundamental not to pool samples with the same index, otherwise it will not be possible to demultiplex them in sequencing data. We used Phusion High-Fidelity DNA Polymerase (NEB). Proceed as follows:

I. 50 µl reaction setup:
   - DAP sample: 25 µl
   - 5x HF Buffer: 10 µl
   - 10mM dNTPs: 2.5 µl
   - Primer i5 (25 µM): 1 µl
   - Primer i7 (25 µM): 1 µl
   - Phusion Polymerase 1 µl
   - milliQ H_2O: 9.5 µl

II. Thermocycler settings:
   - 95°C - 2 minutes
   - 98°C - 30 seconds
   - 20 cycles:
     - 98°C – 15 seconds
     - 60°C – 30 seconds
     - 72°C – 1 minute
   - 72°C – 10 minutes
   - 4°C - ∞

III. Transfer each PCR reaction in a 1.5 ml Eppendorf tube
IV. Proceed with AMPure XP beads purification as described in step 4 of 1) gDNA Fragmentation section, and elute in a final EB volume of 22 µl
V. Check the concentration on QuBit with Qubit dsDNA HS Assay Kit (Thermo Fisher). Final concentrations of DAP-seq samples can vary, but typical amounts range from 3 to 30 ng/µl

3.2.7 Sequencing and Bioinformatic Pipeline

Sequencing was carried out using Illumina NextSeq 500 set up to obtain 30 million 1 × 75 bp paired end reads. pIX-HALO-GST expression vector was used as a negative control to account for non-specific DNA binding as well as copy number variation at specific genomic loci. DAP-Seq reads were mapped to the 'PN40024' 12X.v2 reference genome using bowtie2 (Langmead & Salzberg, 2012), version 2.0-beta7, with default parameters and post-processing to remove reads that have MAPQ scores lower than 30. Peak detection was performed using GEM peak caller (Guo et al., 2012) version 3.4 with the 12X.v2 genome assembly using the following parameters: '*-q 1 -t 1 -k_min 6 -kmax 20 -k_seqs 600 -k_neg_dinu_shuffle*', limited to nuclear chromosomes. Peak summits called by GEM were associated with the closest gene model in the custom annotation file using the BioConductor package ChIPpeakAnno (Zhu et al., 2010) with default parameters (*i.e.* NearestLocation).

3.2.8 *De novo* motif discovery

The coordinates of the DAP-Seq alignment were used to extract from the reference genome (doi.org/10.1016/j.gdata.2017.09.002) the 200 bp peak-centered at GEM-identified binding events sequence for every single peak of every single TF, using the software bedtools getfasta (10.1093/bioinformatics/btq033). For each transcription factor studied, a fasta file was built collecting all the genomic sequences identified by the DAP-seq peaks. These files were subsequently used as input for the motif discovery using the tool" MEME-chip" of the MEME suite v5.5.0 (doi.org/10.1093/bioinformatics/btr189), with default setting. Finally, once the enriched motifs were retrieved, they were submitted to the online tool New PLACE (Higo et al., 1999) to identify any *cis*-regulatory element already characterized in previous studies.

3.3 Results

As general consideration, DAP-Seq is a relatively cheap, easy and fast method to discover the total population of binding sites for a given TF within a genome, defined as cistrome. It is remarkable the fundamental role assumed by the technique applied. Since DAP-Seq is quite new, very few is available in scientific literature. In this case, it was the first time in which DAP-Seq was applied in grapevine and the fifth time applied to a plant species, after *Arabidopsis thaliana*, maize, rice and wheat (Zhang et al., 2022; Zander et al., 2020; Cerise et al., 2020; Galli et al., 2018). For this reason, it was necessary to spend a lot of time making proves aimed at optimizing and standardizing the entire protocol to make it suitable for this species. Since this assay is strictly dependent on library quality (Bartlett et al., 2017), the first step to achieve was the definition of a well working protocol to extract high quality genomic DNA from grapevine tissues. In this regard, we proposed a modification of the protocol proposed by Thomas et al. (1993) that worked particularly well extracting from grapevine leaf. Our proves focused on three different grapevine tissue: young leaf, senescent leaf and softening berry. If from one side the extraction protocol that we have proposed works well for both young and senescent leaf, on the other hand no suitable extraction procedure was individuated for softening berry in order to obtain a final gDNA product featured by high concentration together with appreciable purity levels. About that, we suggest quantifying the final eluted first with a spectrophotometer approach, like Nanodrop for example, because – in addition to the concentration of the nucleic acid - it gives back a measure of the quality, but it always revealed a good practice to flank the Nanodrop measurement also with another one made with QuBit, in order to have a stronger value on which basing the constitution of the samples for library preparation because Nanodrop – in presence of contaminants – gives back an inaccurate quantification (Bartlett et al., 2017). Among the subsequent steps, it is important to highlight that during the purification of the sonicated gDNA with AMPure XP (step 4, paragraph “3.2.6 gDNA library preparation and DAP”) it could happen that the beads do not migrate or do it with difficulty. This situation is due to the low quality of the genomic eluted, specifically for the presence of polyphenols and other secondary metabolites that form a sort of net matrix disturbing the right passages of libraries preparation. In this case, it is fundamental to fix the problem at the basis, or rather review the purification procedure during the genomic extraction process in order to efficiently remove contaminants. Another point on which is necessary to linger is the *in vitro* expression of the protein. As reported in step V, section “5) *In vitro* TF protein expression”, paragraph “3.2.6 gDNA library preparation and DAP”, it is recommended to save 5 µl of each TNT reaction to eventually check the protein on a Western blot, in fact it could happen that even applying correctly the DAP-Seq protocol, at the end there are no gained results. This could also occur because of the

unhappened or incorrect translation of the protein that brings to a lack of binding events between the TF and the genome. Performing a Western blot can help assessing the correctness of the polypeptide synthetization. In particular, it is important to consider that the localization at C- or N-terminal of the HALO tag could differentially affect the correct expression of the TF protein and also the efficient binding due to the steric hindrance exerted by that residue. In this regard, it is advisable to change the location of the HALO tag in case of unsatisfactory results by using C- or N-terminal pIX-HALO expression vector depending on the need. About that, another delicate step is the hybridization between TF-protein and the fragmented gDNA. The calibration of the correct amount of DNA libraries to interrogate with the expressed protein is a critical point, indeed too small quantities of genomic DNA is suboptimal for the detection of all the possible TF binding sites and, on the contrary, an excessive quantity of genomic DNA could lead to many false positives (Bartlett et al., 2017). To this purpose, different amounts of *Arabidopsis thaliana* gDNA have been tested (from 500 pg to 100 ng) and the best results are those given by the hybridization of 50-100 ng of library per each single transcription factor (Bartlett et al., 2017). Considering that the genome size of *A. thaliana* is approximately 125 MB, scaling up the amount on the size of the genome of *Vitis vinifera* (about 500 MB) and considering the heterozygosis level of some grapevine cultivars, the optimum amount was individuated around 200-500 ng library per TF. Moreover, as general rule, if we consider an average loss of 50-70% of gDNA during the entire libraries preparation process, the starting gDNA amount should be considered around 500-1000 ng per transcription factor.

Despite the huge size of the datasets generated from the DAP-Seq experiments defining the cistrome of each TF, it was decided not to filter data based on peak distance from TSS. Although the peak proximity to the gene TSS represents strong evidence for the existence of a cis regulatory elements (CRE) and of a TF-CRE interaction, there is an increasing number of studies describing long-range DNA interactions in gene regulation. Thus, even if these mechanisms have not been fully elucidated, we choose not to filter out distal peaks to avoid losing information. For this reason, for each TF, we conducted the analysis considering the total number of peaks A Hight variability was observed in the total number of binding events amongst different TFs (Fig. 3.3), with WRKY15 showing the highest number of peaks equal to 25096 and MYBA8 the lowest, with only 15 peaks. The distribution of binding events along the assigned gene followed a normal distribution centered in the TSS, in correspondence of which the peak frequencies were higher (see related Figures in the TF-dedicated chapters). The majority of TFs showed a percentage of the total binding events between -5000 - +2000 bp raging between 50-59% (Fig. 3.3), confirming a more intense regulative activity in the perigenic regions. The only exceptions are MYBA7, MYBA8, MYBPA7 and WRKY42, whose poorness in peaks number does not allow to draw any

reliable conclusion regarding peaks distribution. As described in paragraph "3.2.7 Sequencing and Bioinformatic Pipeline" and basing on the total amount of peaks defined with each single DAP-Seq assay, they were identified those genes which are the potential targets of the related TF. Each of these lists was crossed both with the related tissue gene list of the previously defined HSG and WGCNA analysis (Chapter II), obtaining in this way the High Confidence Targets (HCTs), a set of target genes confirmed by two (DAP-Seq + HSG or DAP-Seq + WGCNA) or three (DAP-Seq + HSG + WGCNA) levels of confidence (Tab. 1.3). This multilevel approach which is analyzed and discussed separately for every single TF in the following chapters. For those TFs which did not show triple validated gene targets, genes shared by the DAP-Seq targets list with HGS and/or WGCNA individually were considered as HCTs. GO enrichment analysis was performed on each TF HCTs by using ShinyGO software (Ge et al., 2020). Finally, considering the peaks retrieved by DAP-Seq, a *De novo* motifs discovery analysis was carried out by using MEME-suite to identify the best representative binding motif. The obtained sequences were subsequently matched against the New PLACE database to find motifs previously characterized and the ones with biological meaning were discussed in the related TF paragraph.

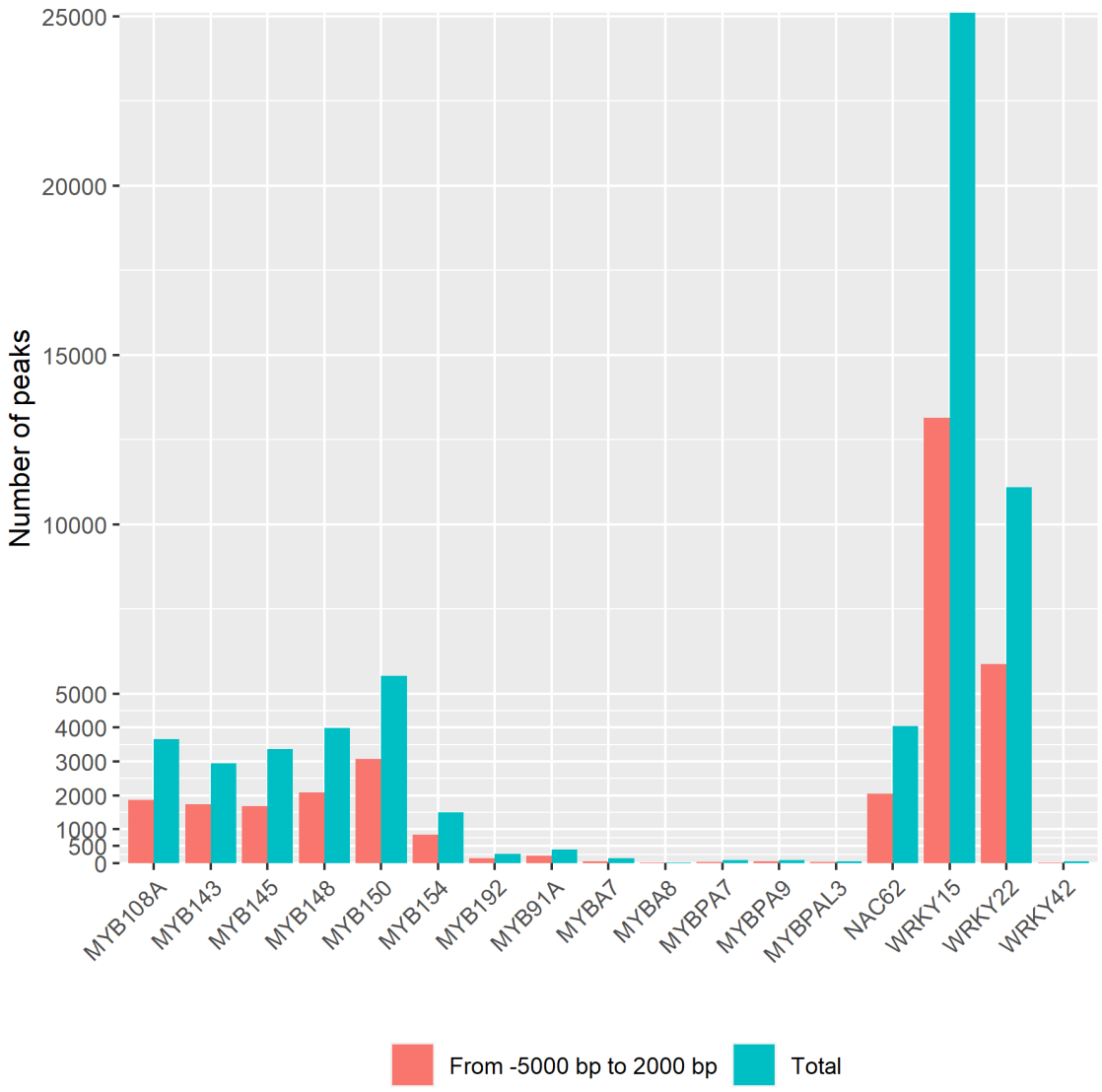


Figure 3.3. General overview of peaks detected for each of the 17 TFs, indicating both the total amount and the amount observed in the region between -5000 bp from the TSS and 2000 from the gene end.

3.3.1 MYB91A

DAP-Seq analysis conducted on MYB91A TF produced a total number of 404 binding events. Of them, 36% were located in the upstream region, 3% overlapped the TSS, 42% were retrieved inside the gene, 1% overlaps with transcription end site and 18% downstream the gene (Fig. 4.3B). Crossing the DAP-Seq results with the Ovary_AA-specific HSG and with the Ovary_AA-related WGCNA module (lightcyan module), no HCT was identified. Thus, we focused on those genes shared between DAP-Seq and HSG, or those shared between DAP-Seq and WGCNA module, separately. In the first case, three genes were identified (Fig. 5.3A). Amongst them VIT_02s0025g03170 it's worth to be noted. This gene belongs to the APETALA2 related Ethylene

responsive transcription factors family (AP2/ERF) and was observed to be downregulated in an early ripening mutant of *V. vinifera* compared to the wild type, suggesting a regulatory role in berry development and ripening (Ma and Yang, 2019), also revealing to be consistent with the specificity of MYB91A for the after anthesis pheno-phase of the ovary. In *Arabidopsis thaliana*, the ortholog is Ethylene-responsive transcription factor ABI4 (ERF052, At2g40220) and, amongst the other described functions, it is involved in starch catabolism (Ramon et al., 2007), seed development (Finkelstein et al., 1998) and ABA and Ethylene mediated signaling pathways (Arenas-Huertero et al., 2000), suggesting a possible similar role in the very early grapevine berry developmental stages. Interestingly, the binding motif obtained for MYB91A (Fig. 6.3) showed similarity (E value = 8,7693E-04) to the motif RAV1AAT, described as the binding consensus sequence of *Arabidopsis* transcription factor RAV1, which contains AP2-like and B3-like domains (Kagaya et al., 1999). This is in line with our findings and is a further suggestion of the regulatory activity shown by MYB91A for AP2/ERF in grapevine.

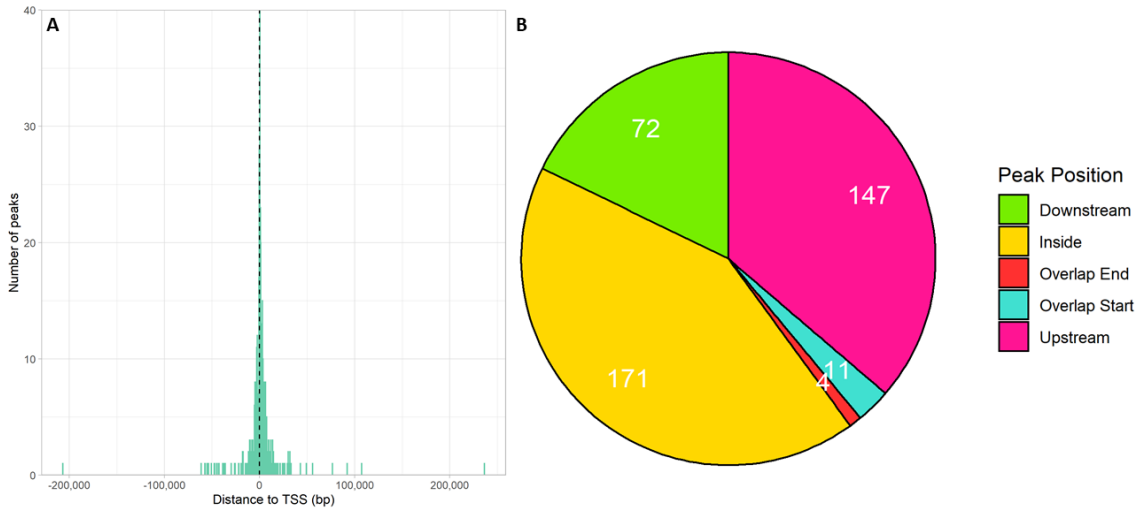


Figure 4.3. MYB91A DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

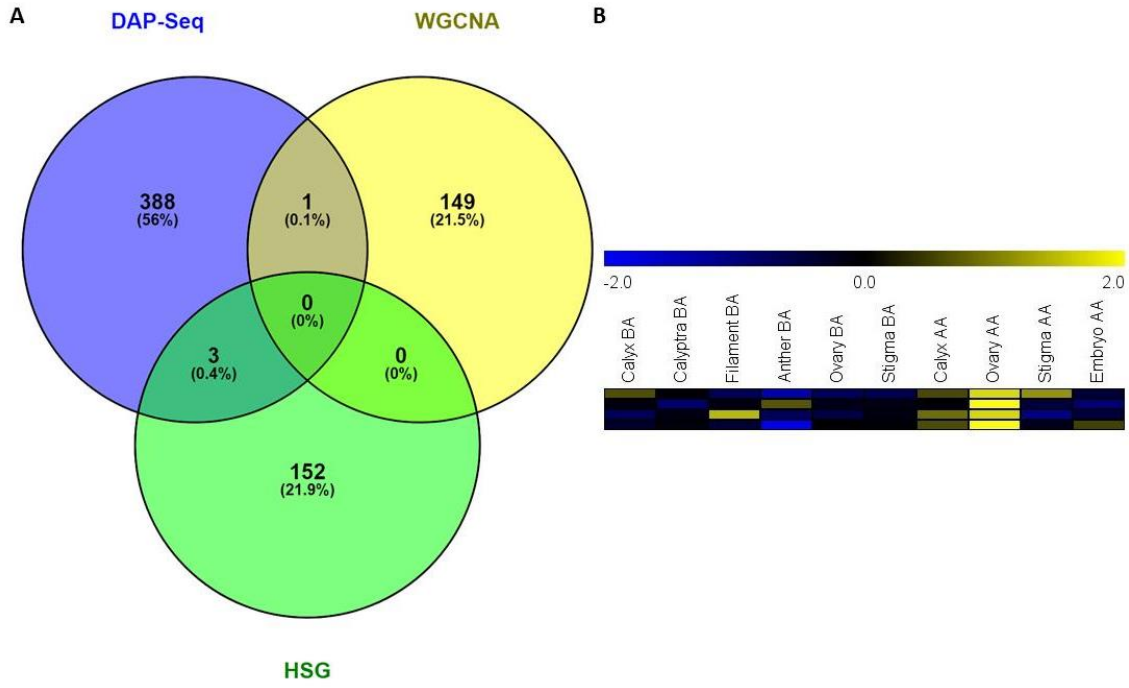


Figure 5.3. Individuation of MYB91A High Confidence Targets. (A) Shared genes between DAP-Seq and Ovary_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

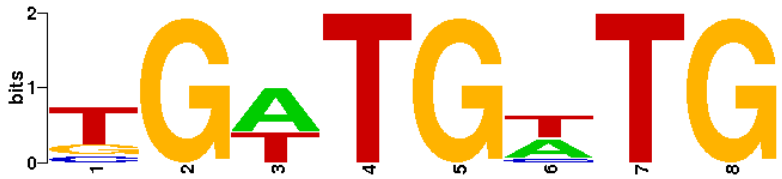


Figure 6.3. *De novo* motif discovery analysis on the whole cistrome of MYB91A discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.2 MYB108A

MYB108A showed a total amount of 3657 peaks, corresponding to an equal number of DNA-binding events. Thirty-one percent of peaks were located upstream the transcription starting site (TSS) of the nearest gene, whereas 20% was detected downstream the stop codon. Remaining peaks were located within the genes: 4% overlapping with TSS, 44% inside the gene and 1% overlapping with stop codon (Fig. 7.3B). To identify Highly Confident Target (HCT) genes we crossed data obtained by DAPseq with those obtained in the flower atlas (Chapter II). More

in detail, we considered as HCTs those genes shared between the MYB108A DAP-Seq analysis, the WGCNA analysis (blue module, Anther_BA specific) and the tau analysis (Anther_BA, HSG). In total we identify 69 genes (Fig. 8.3A). Figure 6.3B illustrates the behavior of MYB108A together with HCT genes in the P. noir flower expression atlas. Interestingly, GO enrichment analysis on HCTs showed an overrepresentation of the Plant Ontology categories related to "anatomy of the pollen", "anatomy of the stamen", "anatomy of the petal" and "whole plant flowering stages" (Fig. 9.3A). Investigating more in detail HCTs list, it was observed that the gene VIT_17s0000g09770 was detected to be only expressed in inflorescence if compared with tendril and leaf (Bio Project PRJNA149111) and was highly expressed in flower tissues of both Corinto Bianco and Pedro Ximenes varieties (Royo et al., 2016). Moreover, based on the Corvina expression atlas it seems to be expressed in stamens (Fasoli et al., 2012). Surprisingly, Corinto Bianco is a somatic variant of Pedro Ximenes which shows parthenocarpy due to absence of microsporogenesis and, moreover, its pollen grains are completely sterile (Royo et al., 2016). The VIT_17s0000g09770 ortholog in Arabidopsis is *CYSTEYNE ENDOPEPTIDASE 1* (*CEP1*; At5g50260), a gene involved in the programmed cell death (PCD) of tapetum cells during pollen formation, whose product is involved in the cleavage of extensins supporting the final collapse of cells (Zhang et al., 2014; Helm et al., 2008). VIT_04s0023g01290 another HCT gene, resulted to be the most highly expressed HCT in Anther_BA in the P. noir flower expression atlas, was observed to be significantly downregulated in Corinto Bianco compared to Pedro Ximenes 10 DBA (CIT.) and to be strictly expressed in stamens (Fasoli et al., 2012), suggesting a putative role in anther development with direct influence on male fertility. This gene was also observed to be expressed in post-véraison berries and in pre-véraison berries subjected to heat stress (Bio Project PRJNA148951). VIT_01s0127g00800 is another HCT specifically expressed in Anther in the P. noir flower expression atlas. In a study aimed at identifying the genetic determinants conditioning the development of loose and compact bunch architecture in grapevine, it was found to be more expressed in inflorescence of loose clustered grapevine (Richter et al., 2020). In the Corvina atlas (Fasoli et al. 2012), this gene was highly expressed in flower and stamens, in agreement with our results, but also in tendril, validating more which was observed by Richter et al., (2020). This result is not surprising, considering that within the *Vitaceae* family tendril constitutes a modified inflorescence (Boss and Thomas, 2002). VIT_12s0028g00210 is a Protein-S-isoprenylcysteine O-methyltransferase B, ortholog of the *Arabidopsis* ICMTB (At5g08335). ICMTB is involved in correct development of flower organs, indeed *icmtb* mutants show flowers developed from the same region of the inflorescence stem with minimal or no elongation of the internodes, secondary axillary flowers developed within flowers, indicating partial conversion of flowers into inflorescence, siliques developed without internode elongation, multiple buds

developed in the axil of cauline leaves and fasciated bifurcated stems (Bracha-Drori et al., 2008). Although VIT_12s0028g00210 could lack an anther-specific role, evidence in its ortholog suggest that it could have a putative role in correct flower organ identity and development, as well as VIT_09s0002g06570, the ortholog of RHOMBOID-like protein 5 (RBL5; At1g52580), a regulator of the correct maturation of reproductive organs (Knopf and Adam, 2012). Another intertting HTC gene is VIT_15s0046g03370, which seems to play a specific role in anther and pollen molecular mechanisms. Its *Arabidopsis* ortholog is LIM domain-containing protein (*PLIM2c*; At3g61230) which is mainly expressed in pollen grains. This protein has an actin-stabilizing activity and binds to actin filaments by promoting cross-linking into thick bundles. It associates predominantly with long and dynamic actin bundles in the shank of growing pollen tubes (Papuga et al., 2010). VIT_15s0046g03370 was detected to be extremely expressed in strictly male diclinic inflorescences, but also in hermaphroditic flowers (Bio Project PRJNA593045). This is consistent with what observed from Papuga et al. (2010) in *Arabidopsis*, suggesting that it could have the same pivotal role in pollen tube growth during the fertilization process in grapevine. The same situation for VIT_15s0046g03370 was also observed for VIT_04s0008g02660 with a high transcript accumulation in male and hermaphroditic flowers (Bio Project PRJNA593045). VIT_04s0008g02660 encodes for a Calmodulin-binding protein that in *Arabidopsis* is called *NO POLLEN GERMINATION 1* (*NPG1*; At2g43040). Genetic, histological, and pollen germination studies on the knockout mutant line indicate that *NPG1* is not necessary for microsporogenesis and gametogenesis but is essential for pollen germination (Golovkin and Reddy, 2003), suggesting an analogous function in grapevine. Again, in Bio Project PRJNA593045, a similar situation was found for VIT_18s0122g00910. Even if this gene is involved in plant response against pathogens, there are several pieces of evidence that suggest a more than probable role in pollen growth. At2g33670 is the relative ortholog and encodes for a *MILDEW RESISTANCE LOCUS O 5* protein (*MLO5*) and beyond its already described role in plant-pathogen interaction, some very interesting observations have been made which could bring us to hypothesize a putative role in pollen grain-ovule interaction also in grapevine. In particular, it is expressed in the pollen tube and seems to act as tethering factors for Ca^{2+} channels and by integrating extracellular ovular cues and selective exocytosis, guides pollen tubes grow directionally to the ovule mediating the ovule-derived signals precisely depending on Ca^{2+} dynamics. *Mlo5* mutants show pollen tubes twist and pile up after sensing the ovular cues (Meng et al., 2020). VIT_06s0004g07230 shows evidence about a putative involvement in flowering timing via interaction with *FLOWERING LOCUS C* (*FLC*). The ortholog encodes for a UDP-glycosyltransferase 87A2 (*UGT87A2*; At2g30140) which regulates the correct flowering time by interacting with the repressor *FLC*, indeed knock-out function mutants were described

to exhibit late flowering upon both long- and short-day conditions (Wang et al., 2012). VIT_05s0029g00130 is another gene that could potentially be involved in grapevine pollen tube growth, indeed *Arabidopsis* ortholog knock-out mutants (HMGB15; At1g04880) show delayed pollen tube growth and a significant reduction in the seed set. Moreover, by means of *in vitro* approaches it was demonstrated that *AtHMGB15* binds to DNA and interacts with the transcription factors AGL66 and AGL104, both required for pollen maturation and pollen tube growth (Xia et al., 2014). The last gene we want to highlight amongst MYB108A HCTs is VIT_04s0044g01280. Contrary to these last genes here described, all involved in pollen tube growth, VIT_04s0044g01280 appears to affect pollen development in a former stage. In *Arabidopsis thaliana*, the ortholog is a *CALLOSE SYNTHASE 5* (*CALS5*; At2g13680), which is responsible for the synthesis of callose deposited at the primary cell wall of meiocytes, tetrads and microspores, and the expression of this gene is essential for exine formation in pollen wall, avoiding that the grains remain fused together. Knockout mutations of the *CALS5* gene resulted in a severe reduction in fertility attributed to the degeneration of microspores. Callose deposition in the *cals5* mutant was nearly completely lacking, suggesting that this gene is essential for the synthesis of callose in these tissues. In detail, the pollen exine wall was not formed properly, affecting the baculae and tectum structure and tryphine was deposited randomly as globular structures. These observations suggest that callose synthesis has a vital function in building a properly sculpted exine, the integrity of which is essential for pollen viability (Dong et al., 2005). In addition to all of these observations pointing out a pivotal role of MYB108A in pollen specific mechanisms of genesis, development and germination, the analysis of the most representative binding motif (Fig. 9.3B) gave back further confirmation. It shows a high similarity rate (E value = 3,4648E-08) with POLLEN1LELAT52. This *cis*-element was previously described as one of two co-dependent regulatory elements responsible for pollen specific activation of tomato (*Lycopersicon esculentum*) LAT52 gene (Eyal et al., 1995), namely AGAAA and TCCACCATA, that are required for pollen specific expression (Bate and Twell, 1998). POLLEN1LELAT52 was also found in the promoter of tomato endo-beta-mannanase LeMAN5, a gene involved in weakening of anther wall during pollen development (Filichkin et al., 2004).

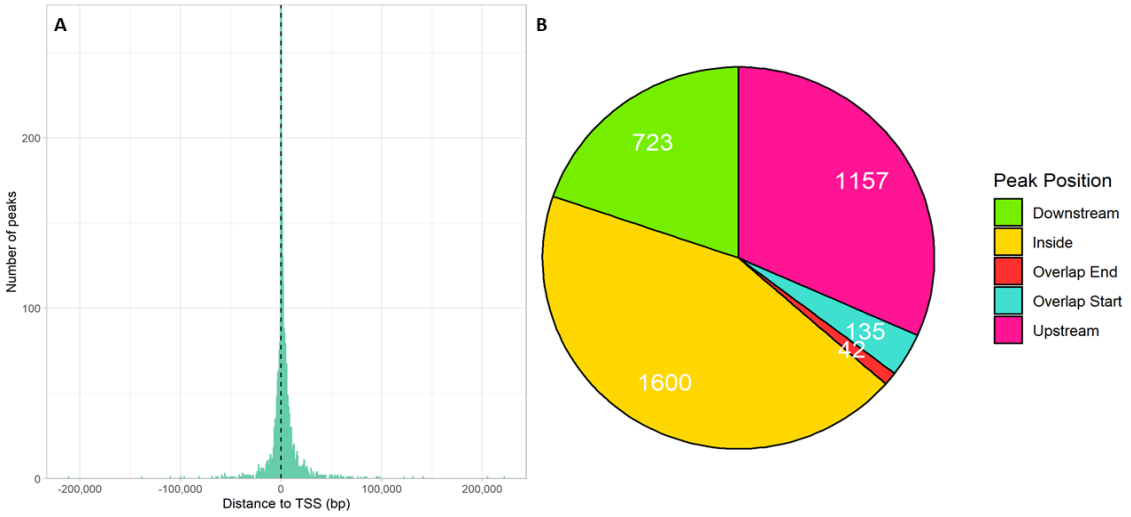


Figure 7.3. MYB108A DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

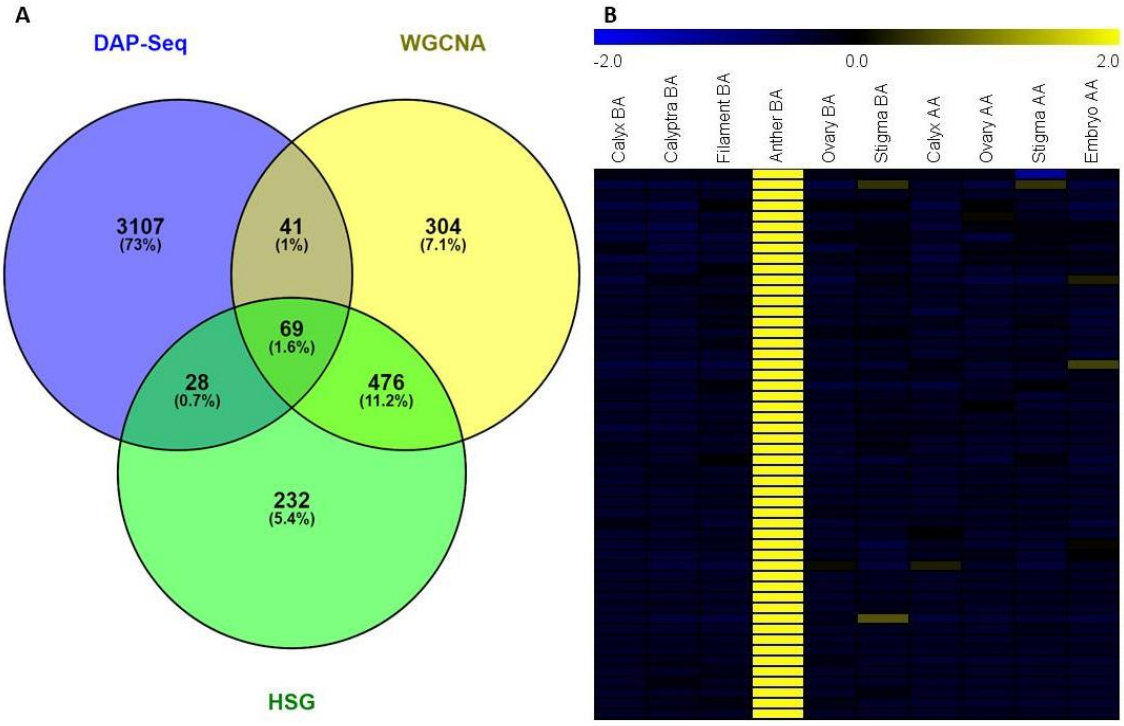


Figure 8.3. Individuation of MYB108A High Confidence Targets. (A) Shared genes between DAP-Seq and Anther_BA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

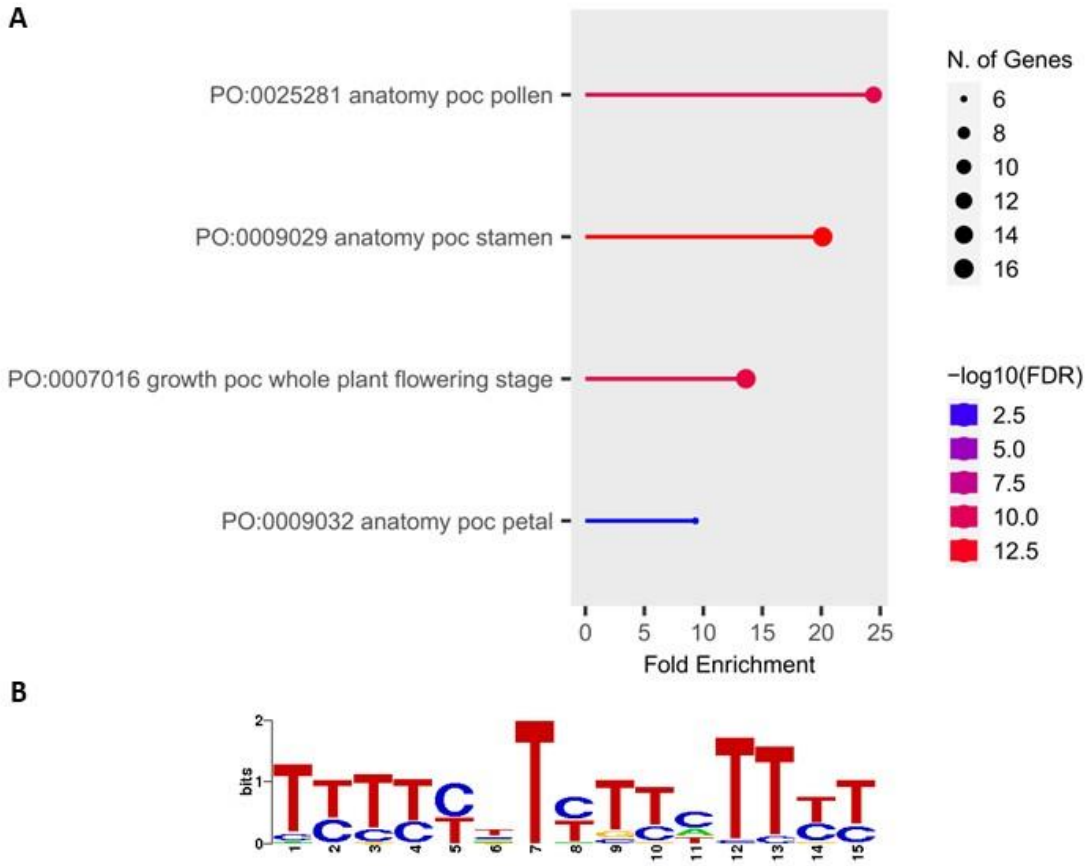


Figure 9.3. (A) GO enrichment analysis performed on MYB108A HCTs: Plant Ontology (PO) categories. (B) *De novo* motif discovery analysis on the whole cistrome of MYB108A discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.3 MYB143

The analysis of MYB143 showed 2948 peaks distributed as follows: 35% in the upstream region, 5% at the start site, 41% within the gene, 1% overlapping gene end and the remaining 18% was localized the downstream region (Fig. 10.3B). A subset of 24 HCTs was achieved by crossing the subset of the DAP-Seq targets of MYB143 with the gene list of WGCNA brown module and HSG relative to the Stigma_AA (Fig. 11.3A). Among them, a subgroup of MYB143 HCTs deserves a separate discussion. The pollen-stigma interaction is one of the most important processes occurring in flowering plants, since it results in fertilization and thus in seed set. This critical cellular dialogue between the haploid pollen (grain and tube) and the diploid cells of the stigma (and style) is one of the most precisely adapted of all activities of the plant, morphologically, physiologically, and biochemically, and became a paradigm for the study of cell recognition and

signaling in plants (Heslop-Harrison and Shivanna, 1978). Because of this peculiarity, there is a delicate hormonal balance involved in the cross-communication process, like reactive oxygen species (ROS), abscisic acid (ABA) and ethylene (McInnis et al., 2006; Kovaleva et al., 2003). In this regard, in addition to pointing out the enrichment of the ontological category named "plant hormone signal transduction" (Fig. 12.3A), it is noteworthy that the HCT VIT_08s0040g00960 is a carboxylesterase (Navarro-Payà et al., 2022) found expressed in Pinot noir inflorescence at anthesis moment (Rossmann et al., 2019), and more specifically in carpel (Bio Project PRJNA307079) and stamens (Fasoli et al., 2012). This could be consistent with the strict interconnection that elapses between stigma and stamen tissues for the peculiar intercommunication which specifically features these two floral whorls. In addition, also the ortholog gene At1g68620 was described having the same pattern, with clear expression in inflorescence (Marshal et al., 2003) and stamens (Winter et al., 2007) at blooming time. A similar trend was also observed for the GDSL-esterase/lipase encoded by VIT_10s0003g02090. Indeed, its transcript was abundantly detected within the carpel and pollen grain (Fasoli et al., 2012), as well as for the putative ortholog (At5g45670) highly accumulated in ovary structure (Klepikova et al., 2016; Winter et al., 2007). The transcript of VIT_08s0032g00710 was relieved in flower of Pinot Meunier (Bio Project PRJNA149111). More in detail, its ortholog in *Arabidopsis* (At5g10500) was found in stamens and specifically in mature pollen grains (Winter et al., 2007). In this last case, its expression was observed to increase with the course of grain development and reaches its maximum during the emission of germinative pollen tube in correspondence of the actively growing tissue penetrating the stigma opening (Qin et al., 2009). These observations are in line with the function of this gene, which encodes for a plant-specific actin-binding protein being part of a membrane-cytoskeletal adapter complex (Deeks et al., 2012). As a matter of fact, it a high cytoskeletal activity in pollen tube growth and development has been well described. For this reason, it is reasonably supposable an analogous function in grapevine where was detected in stigmatic tissues precisely during the after anthesis moment in which pollen and stigma are interacting. A similar expression pattern was highlighted again in Corvina for the EF-hand domain-containing protein VIT_12s0059g00370, which is clearly switched on in anthers and pollen grains (Fasoli et al., 2012), as well as for its ortholog (At4g31240) in *Arabidopsis* stamens (Winter et al., 2007). Still remains to be explained the expression of this gene in ripening berry pericarp as reported for Barbera (Magris et al., 2019), Thompson Seedless (Balic et al., 2018), Sauvignon blanc (Helwi et al., 2016), Cabernet Sauvignon, Riesling (Lu et al., 2022) and Semillon (Blanco-Ulate et al., 2015). Finally, VIT_15s0021g01390 has At3g60270 as ortholog in *Arabidopsis* which expression pattern retraces what we observed, indeed this Cupredoxin turns on in stamens and pollen before anthesis, to then involve ovary and stigma tissues after the

occurred pollination (Winter et al., 2007). The *de novo* motif discovery analysis for MYB143 cistrome highlighted a sequence (Fig. 12.3B) showing an E value equal to 6,7873E-05 for a previously described binding site: the GATA-box. This *cis*-element was widely observed to be conserved among several light-responsive promoters (Gilmartin et al., 1990; Lam and Chua, 1989). In particular, GATA-box was revealed in three sequence repeats positioned between the TATA and CAAT box elements conserved in correspondence of promoter regions of all *LHCII* (light-harvesting complex) Type I Cab genes in *Petunia hybrida* (Gidoni et al., 1989). These genes function as a light receptor, binding at least 14 chlorophylls (8 Chl-a and 6 Chl-b) and carotenoids such as lutein and neoxanthin, capturing and delivering excitation energy to photosystems with which they are closely associated (Wientjes et al., 2011; Wientjes and Croce, 2011; Storf et al., 2005). Moreover, *LHCII* genes expression is induced by low light but repressed by high light (Ganeteg et al., 2004). The biological meaning of an enriched light-responsive binding motif featuring the entire cistrome of a TF stigma-related is not that immediate. The encouraging detection is that for almost all the TFs stigma related it was retrieved at least one light response related motif showing high similarity for the identified best binding sequence, thus defining a nevertheless consistent finding. The explication of this could be the fact that stigma whorl suddenly finds itself in a bright environment after the capfall, later having passed all the before anthesis period in absolute darkness due to the coverage provided form the calyptra, and for this snap change of conditions, the light-responsive elements are triggered. Another explanation could be given by the observation of the specific relationships that exist between the different floral organs. In this regard, the stigma represents the communication interface between pollen and ovule. In fact, it is precisely in this whorl that the pollen grain is recognized and - in the presence of compatibility - left to germinate until the tube reaches the mature ovule ready to be fertilized. Synchronization, which is of fundamental importance in this process, is precisely managed by the stigma in terms of a correct hormonal balance which allows pollen germination. About that, it is likely to think that - in an autogamous species such as grapevine in which fertilization takes place in a period extremely close to the capfall - the low intensity of light that penetrates inside the flower when the abscission line begins to form acts as a trigger for the hormonal changes on the stigmatic surface which will allow the development of the pollen tube only at this point, thus ensuring perfect synchrony with the maturity of the ovule, and subsequently sopping the trigger when calyptra is completely removed and the fertilization has already taken place. Even if this explanation is consistent with the fact that *LHCII Cab* genes, in which GATA-box was observed to be enriched, show induction by low light and repression by high light, of course it is pure theoretical speculation and any light involvement in this sense has yet to be experimentally proven.

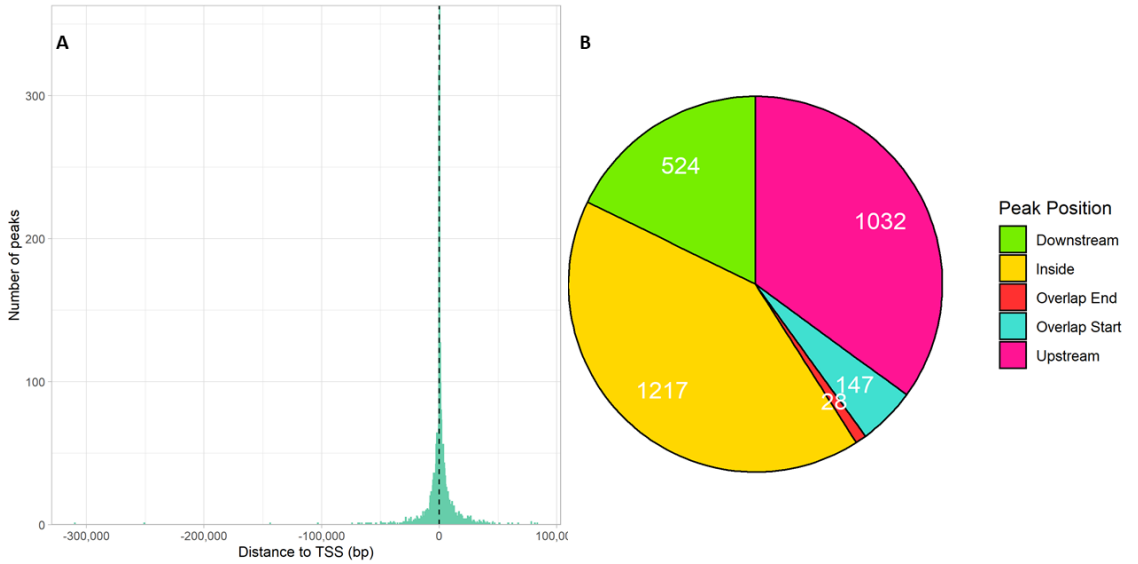


Figure 10.3. MYB143 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

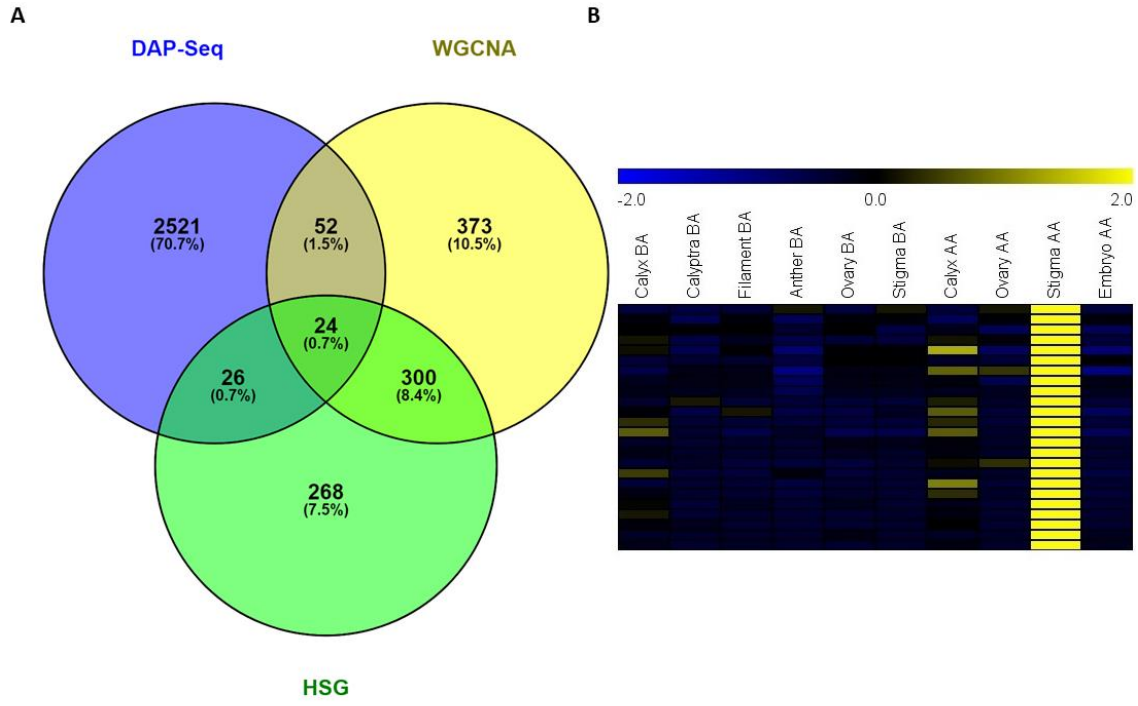


Figure 11.3. Individuation of MYB143 High Confidence Targets. (A) Shared genes between DAP-Seq and Stigma_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

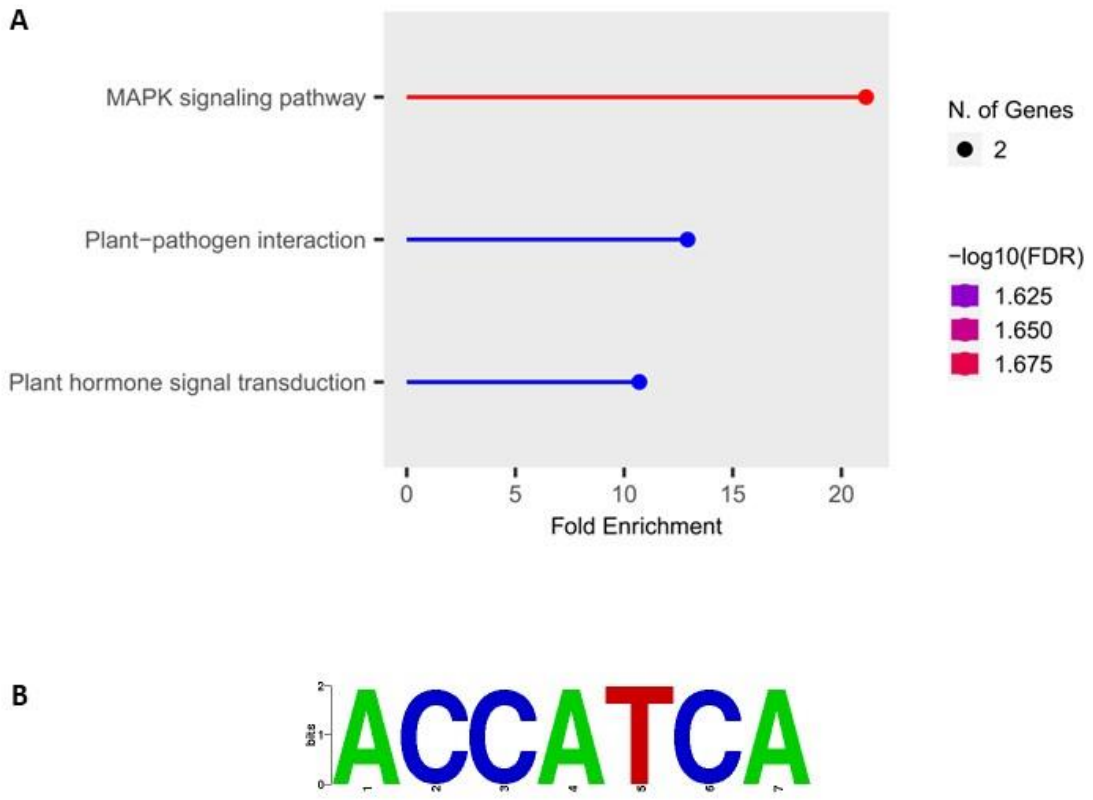


Figure 12.3. (A) GO enrichment analysis performed on MYB143 HCTs. (B) *De novo* motif discovery analysis on the whole cistrome of MYB143 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.4 MYB145

A total number of 3370 binding events was defined with DAP-Seq of MYB145. Regarding peaks distribution, 34% of the total amount was found upstream the relative gene, 3% overlaps with the TSS, 43% defines the set observed inside the gene length, 1% include the transcription end point and the downstream region showed 19% of the total binding events (Fig. 13.3B). By crossing DAP-Seq results with WGCNA brown module and HSG analyses of Sigma_AA, it was possible to individuate 24 distinct HCTs (Fig. 14.3A). A first clear observation was the presence of VIT_15s0021g01390 also in the MYB145 HCTs network as well as for MYB143. The presence of a putative target in the networks of two TFs strictly related with stigma tissue is an additional empowering of a yet strongly consistent regulative relationship. Contrariwise, featuring only MYB145 HCTs, is VIT_08s0007g04680, an extensin (*EXPA12*; Navarro-Payà et al., 2022), which seems to cause loosening and extension of plant cell walls by disrupting non-covalent bonding between cellulose microfibrils and matrix glucans (UniProt). It was observed to be intensely

expressed in pollen grains in correspondence to flowering time (Fasoli et al., 2012), but also in strictly female wild grapevine flowers at the same developmental period (Massonnet et al., 2020). Something similar was reported for the Pectate lyase encoded by VIT_17s0000g05740, which transcript was found highly stored in pollen of Corvina at blooming moment (Fasoli et al., 2012) and in exclusively female flower of monoecious wild *Vitis* (Massonnet et al., 2020). Moreover, also the putative ortholog (At3g01270) is highly expressed in *A. thaliana* pollen grains, showing an increasing trend with the progression of microgametogenesis and remaining constantly intense also during the germination and the penetration of the pollinic tube inside the stigma, and in anthers (Klepikova et al., 2016; Qin et al., 2009; Winter et al., 2007). A similar expression profile was outlined from the ortholog (At1g75900) of VIT_18s0001g14810. This gene encodes for a GDSL esterase/lipase (*EXL3*) which belongs to the gene families of *Arabidopsis* pollen coat proteome and was described to be expressed in floral buds, specifically in carpel and stamen filament of the mature flower (Klepikova et al., 2016; Winter et al., 2007; Swanson et al., 2005; Mayfield et al., 2001). A peculiar expression pattern is highlighted by a Galactinol-raffinose galactosyltransferase (VIT_05s0077g00840), which expression was described to switch on exclusively in stamens and pollen during anthesis, to then go on to define a progressive increase of its expression levels in the developing embryo, reaching the apex in correspondence of mid-ripening phenophase (Fasoli et al., 2012). Also the related *Arabidopsis* ortholog At4g01970 evidenced an expression involving ovules (Winter et al., 2007), embryo tissues in developmental phases (Casson et al., 2005), as well as in the whole ovary and stigma precisely (Swanson et al., 2005). VIT_07s0104g01260 ortholog is the flavin monooxygenase *YUCCA11* and it was described to synthesize auxin essential for embryogenesis (Cheng et al., 2007), but also in this case the expression was encountered in pollen grain at flowering stage (Winter et al., 2007). From a general point of view, *CYP86B1* (At5g23190; ortholog of VIT_01s0011g02060) is involved in very long chain fatty acids (VLCFA) omega-hydroxylation. It is required for the synthesis of saturated VLCFA alpha, omega-bifunctional suberin monomers (Compagnon et al., 2009; Molina et al., 2009) and it shows a conspicuous level of expression in stigma and anthers (Klepikova et al., 2016; Compagnon et al., 2009). On the other hand, VIT_19s0014g04240 shows a specific activity related to the pollen-pistil interaction typically ascribable to the stigma, in fact this gene belongs to the “recognition of pollen” ontology category (GO:0048544), and the same is for the putative ortholog (At4g27300) which encodes for the G-type lectin S-receptor-like serine/threonine-protein kinase SD1-1 (UniProt). In this regard, it is interesting to notice that the analysis on the best representative binding sequence (Fig. 15.3) brings back a conspicuous similarity (E value = 2,2985E-12) with two different motifs. The first one is the same highlighted for MYB108A (see paragraph “3.3.2 MYB108A”), i.e.

POLLEN1LELAT52. As was already reported, this *cis*-element was described to be conserved in pollen development and germination affecting genes, element that is coherent with the several pollen related targets found in MYB145 HCTs network and also with the strict association occurring between pollen and stigma, properly from the anthesis on, as it was defined this TF related to. On the other hand, the most enriched motif of MYB145 cistrome presents the same similarity degree also with another *cis*-element: DOFCOREZM. This oligonucleotide was described as the core site required for binding of DOF proteins in maize (Yanagisawa, 2000). DOF proteins are DNA binding proteins which seem to have a light dependent activity. For example, in *Arabidopsis* DOF2.5 is involved in the maternal control of seed germination and may ensure the activation of a component that would trigger germination as a consequence of red-light perception (Santopolo et al., 2015). Moreover, in maturing siliques DOF2.5 and DOF3.7 were found expressed all through the funiculus connecting the placenta to the ovule (Gualberti et al., 2002), observation that encounters consistency with both the tissue and the phenological stage of our analysis. Also in this case, the main point is the light dependent activity that could bring again to the speculative proposal exposed at the end of paragraph "3.3.3 MYB143".

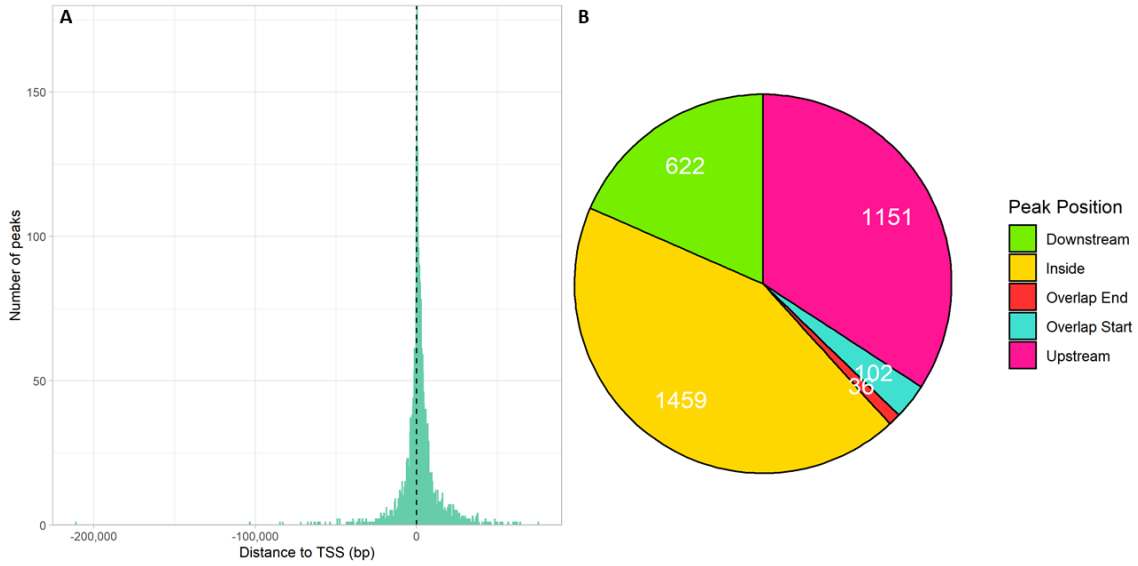


Figure 13.3. MYB145 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

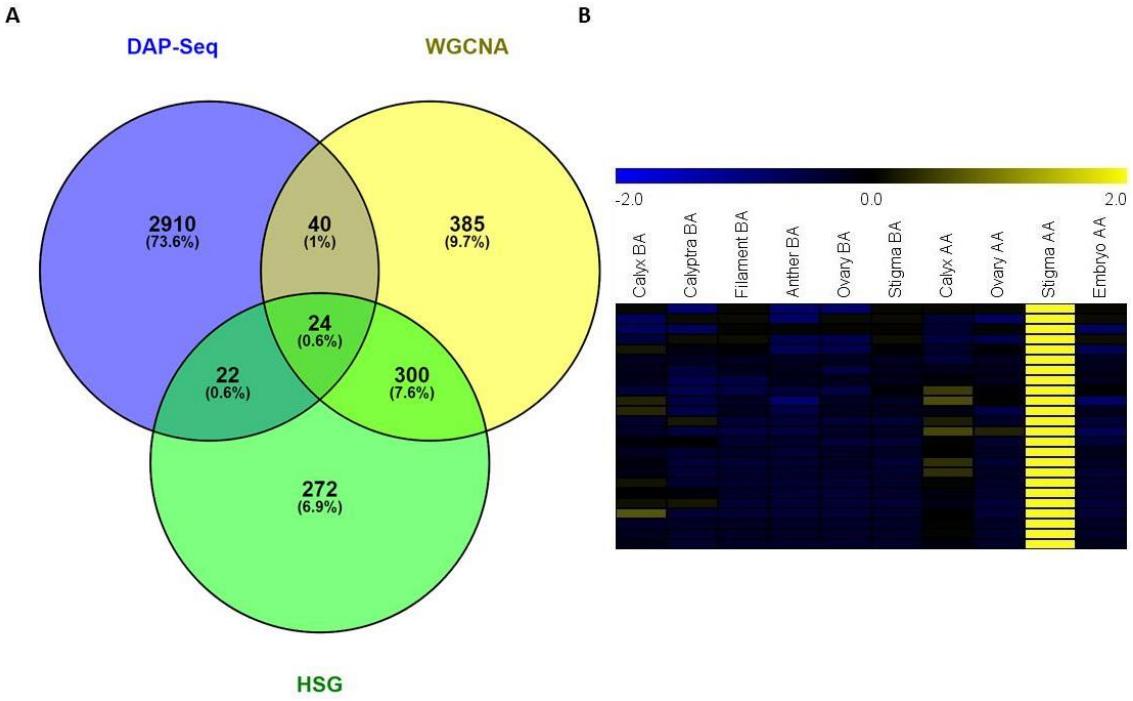


Figure 14.3. Individuation of MYB145 High Confidence Targets. (A) Shared genes between DAP-Seq and Stigma_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

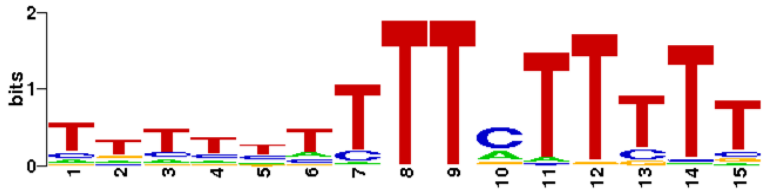


Figure 15.3. *De novo* motif discovery analysis on the whole cistrome of MYB145 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.5 MYB148

To the DAP-Seq of MYB148 were assigned 3988 peaks. In particular, 34% featured the upstream region, 4% overlaps the TSS, 42% falls within the gene length, 1% lays in correspondence with the end site and the final 19% forms the subset of the binding events downstream the relative gene (Fig. 16.3B). Based on our analysis, also MYB148 activity is strictly related to the stigma tissue after the moment of flowering, therefore the DAP-Seq results were intersected with HSG

analysis and WGCNA brown module both related to Stigma_AA whorl, thus obtaining a subset of 37 HCTs (Fig. 17.3A). The enriched categories for this group are reported in Fig. 18A and B. Fist needed specification, is the presence in MYB148 HCTs network of one gene of MYB143, or rather VIT_08s0032g00710, a plant-specific actin binding protein being part of a membrane-cytoskeletal adapter complex (Deeks et al., 2012). As described in paragraph "3.3.3 MYB143", this gene seems to have a specific role in pollen-stigma interaction molecular mechanisms, as well as for what observed by other two genes shared between the HCTs network of MYB148 and MYB145, namely VIT_08s0007g04680 and VIT_19s0014g04240. Even for these two genes, what catches the attention is again the involvement in recognition and interaction processes occurring among pollen and stigma, in which the activity of these two genes would seem to affect. Also the function of VIT_12s0057g00320 may reside in this ambit, indeed it is an Exo-polygalacturonase (Navarro-Payà et al., 2022). The expression of the genes related to this class of enzyme, as described in Chapter II, has been widely described in tapetum, pollen grains, stigmas and pollinated pistils of various species and implies their role in tapetum degradation, pollen maturation, pollen tube growth, and pollination (Yang et al., 2018). The expression of this gene does not deviate from what reported for other polygalacturonase, indeed it was relieved in Tempranillo whole flower (Grimplet et al., 2019), but specifically and exclusively detected in pollen grain and stamens of Corvina at blooming moment (Fasoli et al., 2012). In addition, also the ortholog (At3g07850) was studied as an Exo-polygalacturonase functioning in depolymerizing pectin during pollen development, germination, and tube growth (UniProt), and even in this case the expression was confirmed in *Arabidopsis* stamens and pollen from microgametogenesis up to tube germination (Klepikova et al., 2016; Qin et al., 2009; Winter et al., 2007; Honys and Twell, 2004). This evidence provides us insights outlining a concrete involvement of this gene in grapevine pollen developmental molecular mechanisms, furthermore, inserting itself in a biologically significant way in the MYB148 regulatory network. This point could be reasonably enlarged to the Calcium dependent protein kinase *CPK6* (VIT_06s0009g03150; Chen et al., 2013), which transcript was detected in grapevine anthers and pollen grains (Fasoli et al., 2012). This gene seems to contribute at the same function in which acts VIT_18s0122g00910 in MYB108A HCTs network (see paragraph "3.3.2 MYB108A"), guiding pollen tube to the ovule mediating signals depending on Ca2+ dynamics. Several evidence highlights that the putative ortholog (At5g12180; Calcium-dependent protein kinase 17, *CPK17*) may play a role in signal transduction pathways that involve calcium as a second messenger and regulates polarized tip growth of the pollen tube (Myers et al., 2009), even more so since its expression has been ascertained in *A. thaliana* mature and germinant pollen (Klepikova et al., 2016; Qin et al., 2009; Winter et al., 2007). Interestingly, among HCTs there is a MYB TF, namely

MYB67B (VIT_19s0085g00940; Wong et al., 2016) which expression was encountered in flowering inflorescences of Villard blanc and Touriga Nacional (Kamal et al., 2019; Ramos et al., 2014). By analyzing in detail the function of the related ortholog (At3g13890), in *A. thaliana* called MYB35, it was realized that this TF has a pivotal role in plant fertility, in fact it acts upstream of the lignin biosynthesis pathway playing a role in specifying early endothecial cell development by regulating a number of genes linked to secondary thickening, such as *NST1* and *NST2* (Yang et al., 2007). Specifically, it regulates lignified secondary cell wall thickening of the anther endocethium, which is necessary for anther dehiscence. (Mitsuda et al., 2006; Steiner-Lange et al., 2003). Even if the disruption of the phenotype does not affect the fertility of the pollen, plants are male sterile due to a defect in anther dehiscence (Steiner-Lange et al., 2003). The expression of this TF was assessed in pollen grains and in anthers early during endothecial development, with maximal expression during pollen mitosis I and bicellular stages (Klepikova et al., 2016; Yang et al., 2007; Winter et al., 2007). Because of the reported function, this TF would deserve further insights aimed to study the putative conserved role among different plant species and - based on them - could represent an interesting starting point for breeding programs aimed to individuate and transfer genetic resources, for example, for male sterility featured crops. On the other hand, typical of stigmatic tissues are genes related to hormonal signaling and balance for communication purposes, such as ABA, ROS, and ethylene (McInnis et al., 2006; Kovaleva et al., 2003). In this regard, among MYB148 HCTs were individuated two genes related to ROS (VIT_01s0010g01090 and VIT_08s0058g00970) and one related to ethylene (VIT_06s0004g05240). VIT_08s0058g00970 encodes for the Cationic peroxidase PrxIII02e1 (Navarro-Payà et al., 2022) which ortholog (At5g05340) was reported to be hugely expressed in *Arabidopsis* flower at anthesis (Winter et al., 2007), as well as for VIT_01s0010g01090, which encodes for PrxIII32a, another peroxidase (Navarro-Payà et al., 2022). Finally, VIT_06s0004g05240 is an Ethylene receptor (*ETR2*; Navarro-Payà et al., 2022) highly expressed in mature pollen (Fasoli et al., 2012). At2g40940 is the putative ortholog which in turn was observed to be expressed in almost every floral whorl, but in particular in stamens and pollen, and it was described as an ethylene induced ethylene receptor (Klepikova et al., 2016; Winter et al., 2007; Hua et al., 1998). By matching the most enriched binding sequence (Fig. 18.3C) against the NewPLACE database, it was noticed a high similarity (E value = 1,4654E-12) for three different motifs. The first two are POLLEN1LELAT52 and DOFCOREZM, already mentioned and discussed before. The third *cis*-element is GT1CONSENSUS and was described as specific DNA core binding site of GT-1, a transcription factor described for the first time in *Arabidopsis thaliana* acting as a molecular switch in response to light signals (Ayadi et al., 2004; Merechal et al., 1999; Hiratsuka et al., 1994). In this regard, GT1CONSENSUS was found in a very

huge number of light-regulated genes, with both chloroplast and non-chloroplast function. Just to name a few, *Rubisco* and *Cab* genes from many species, *PHYA* from oat (*Avena sativa* L.) and rice (*Oryza sativa* L.), *RCA* and *PETA* of spinach (*Spinacia oleracea* L.), and bean (*Phaseolus vulgaris* L.) CHS15 (Zhou, 1999; Villain et al., 1996). Very interestingly, GT1CONSENSUS was also detected in maize (*Zea mais* L.) in the pollen-specific gene *LAT52* (Zhou, 1999), the same gene in which was described for the first time the motif POLLEN1LELAT52 in tomato (Filichkin et al., 2004; Eyal et al., 1995). There are two elements that seem to connect the observations made so far. The first is that both GT1CONSENSUS and DOFCOREZM (see paragraph "3.3.4 MYB145") were detected in the *Cab* genes, supporting the hypothesis that light induction indeed plays a central role in these mechanisms. The second element is that GT1CONSENSUS has been found in the pollen-specific gene *LAT52*, thus suggesting that the activating effect operated by light can trigger some processes characterizing the development of pollen, all the more considering the extreme similarity both for POLLEN1LELAT52 and for the two motifs light-related DOFCOREZM and GT1CONSENSUS that distinguishes the binding sequence identified for the MYB148 cistrome. In this sense, this would fit perfectly with the speculation proposed in paragraph "3.3.3 MYB143", suggesting a possible role of synchronization of the development and maturation of the different floral whorls by light.

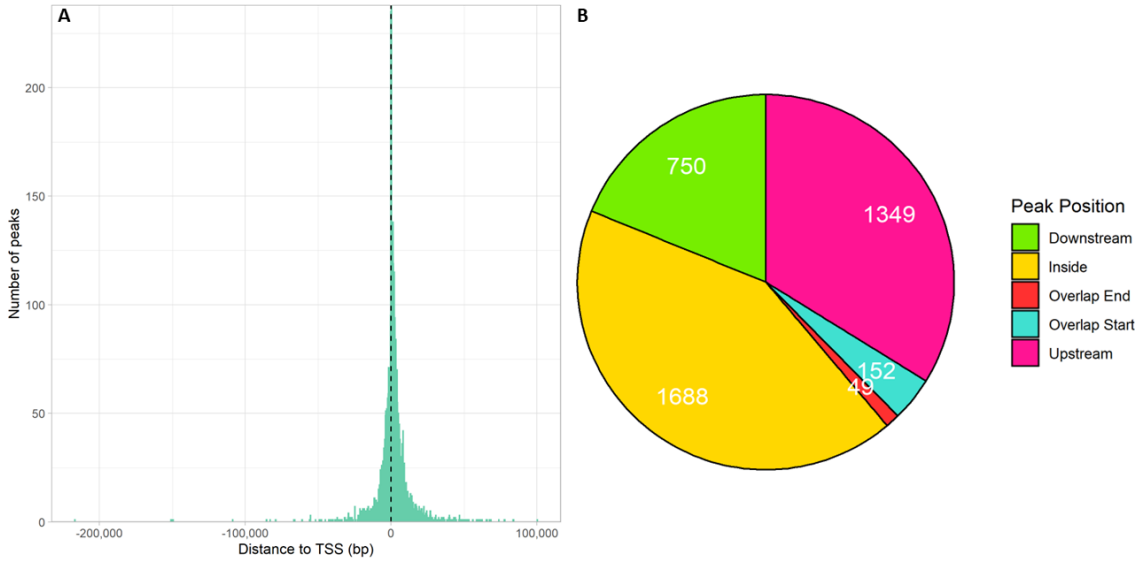


Figure 16.3. MYB148 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

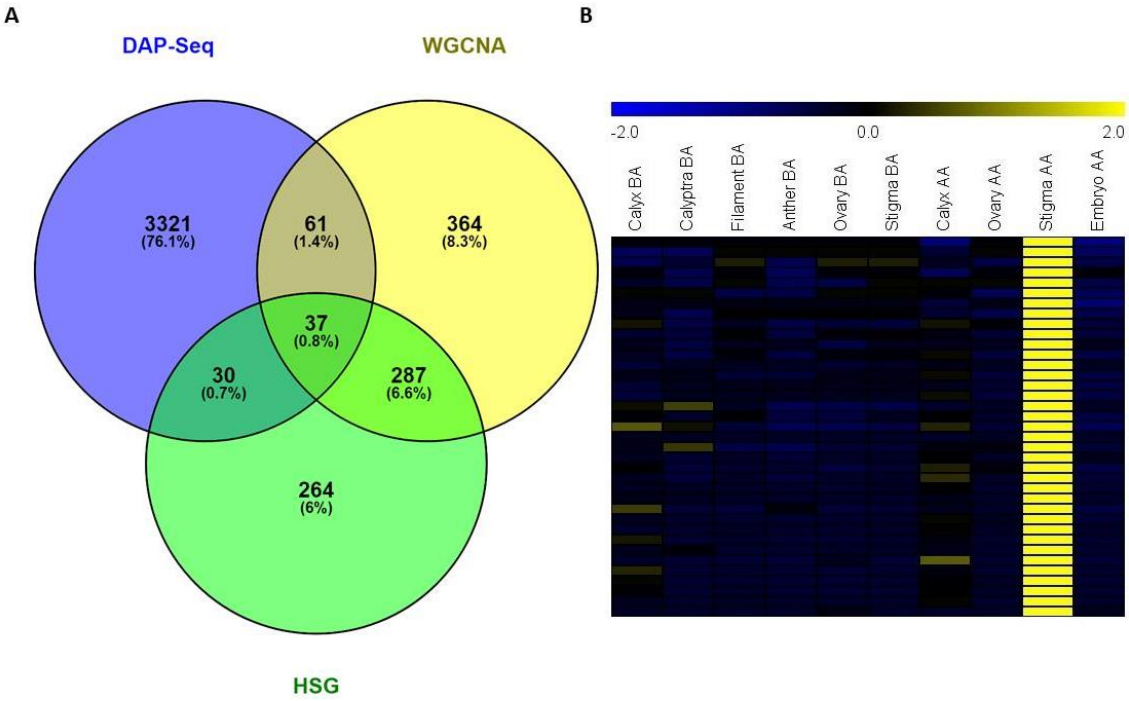


Figure 17.3. Individuation of MYB148 High Confidence Targets. (A) Shared genes between DAP-Seq and Stigma_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

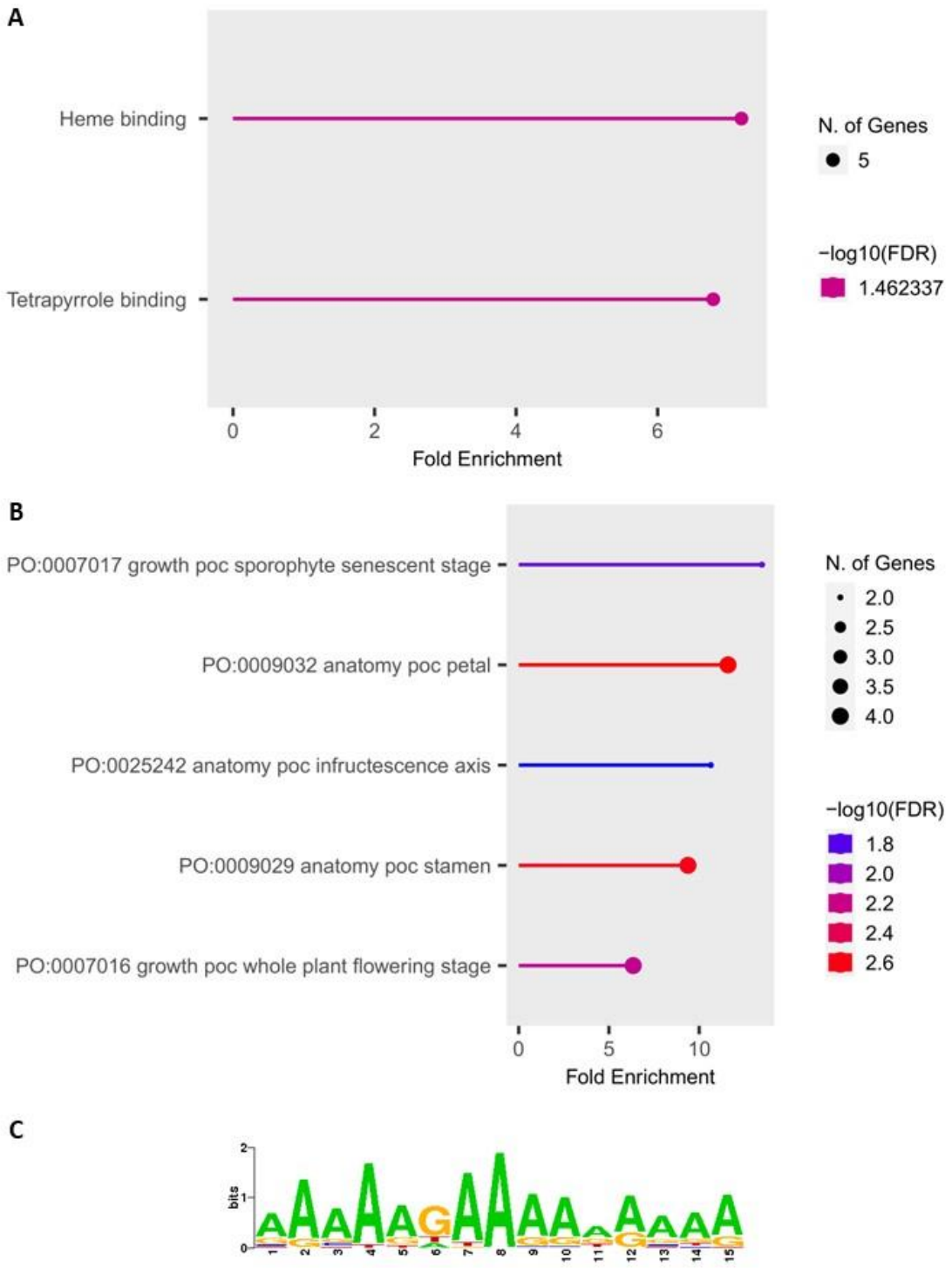


Figure 18.3. GO enrichment analysis performed on MYB148 HCTs: (A) Molecular Function categories and (B) Plant Ontology (PO) categories. (C) *De novo* motif discovery analysis on the whole cistrome of MYB148 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.6 MYB150

5526 are the peaks found to be related to MYB150. The localization is distributed as follows: 32% of the binding events were observed upstream the related gene, 5% comprises the starting site of the transcription, 45% lays inside the gene, 1% was detected in correspondence of gene end and downstream the remaining 17% (Fig. 19.3B). The regulative network of MYB150 consists of 41 HCTs genes retrieved by crossing DAP-Seq results with brown module of WGCNA and HSG of Stigma_AA (Fig. 20.3A). First of all, is noteworthy the enrichment of Plant Ontology categories performed on HCTs, that shows - from general to specific - “growth whole plant flowering stage”, “anatomy petal” and “anatomy stamen” (Fig. 21.3A), in line with the putative involvement of MYB150 in the flowering process, in particular concerning stigma tissues which are strongly interconnected with stamen. Among these 41, 16 are shared with the previous TFs related to Stigma_AA tissue in different combinations, that is 3 with MYB143, 2 with MYB145, 5 with MYB148, 3 with both MYB143 and MYB148, and finally 3 with both MYB145 and MYB148. Within these 16, VIT_08s0007g04680 and VIT_19s0014g04240 are involved in pollen development and in specific interconnection relationships occurring between pollinic grain and stigma, while VIT_06s0004g05240 and VIT_01s0010g01090 are involved in hormonal signaling ROS- and ABA-mediated. The shared genes between different regulative networks make the HCTs approach seem more and more reliable as a method to infer the role of a given TF, by analyzing the molecular function of the putative targets and by framing them in a biological meaning related to the respective tissue. On the other hand, regarding HCTs of MYB150 alone, it was individuated the gene VIT_08s0040g00920. Its expression was detected in grapevine inflorescences (Bio Project PRJNA149115). At2g29420 is the putative ortholog encoding for the Glutathione S-transferase U7 (*GSTU7*; UniProt), which expression level was observed high in *Arabidopsis* stamens at anthesis (Klepikova et al., 2016; Winter et al., 2007). Since it is induced by salicylic acid (Blanco et al., 2005), it is therefore speculable an involvement in stigma intercommunication through hormone signaling. The same speculation could be also extended to the Sinapyl Alcohol Dehydrogenase 2 (VIT_18s0122g00450; Navarro-Payà et al., 2022), whose expression in Pinot noir berry skin was described to be induced by ABA-treatment (Pilati et al., 2017). On the same line, the putative ortholog At4g37990 encodes for the Cinnamyl alcohol dehydrogenase 8 which is involved in lignin biosynthesis by catalyzing the final step specific for the production of lignin monomers (Kim et al., 2004), being expressed in the style, anthers, stamen filaments and stigmatic regions (Kim et al., 2007). Staying in ABA-topic, VIT_18s0041g02410 was also encountered in MYB150 HCTs, described as Aldehyde oxidase 1 (Navarro-Payà et al., 2022). Looking at its putative ortholog At2g27150, it was seen to catalyze the oxidation of abscisic aldehyde into abscisic acid, the last step of ABA biosynthesis (Seo et al.,

2000), in this way entering perfectly in the biological frame of stigma as high ABA-featured tissue. In the hormonal balance topic, it seems to fit also AP2-like ethylene-responsive transcription factor (*AP2-01*, VIT_11s0037g00870; Licausi et al., 2010) which expression was assessed in grapevine anthers at anthesis (Fasoli et al., 2010). Analyzing the related ortholog which was described as ethylene and ABA-responsive (Lee et al., 2009; UniProt), it was realized that At1g16060 encodes for a namesake protein that positively regulates the biosynthetic process of cutin (To et al., 2012) and - by observing the intense expression of the relative transcript properly in stigma (Swanson et al., 2005), anthers (Klepikova et al., 2016) and pollen during microgametogenesis (Honys and Twell, 2004) - it is plausibly inferable an active role in these whorls growth and development. On the other hand, concerning the peculiar cell-recognition mechanisms of stigma, the Receptor-like serine/threonine-protein kinase encoded by VIT_05s0020g03270 was found expressed in grapevine ovary during flowering (Fasoli et al., 2012) and belongs to GO:0048544, the same ontological category of VIT_19s0014g04240, also present in MYB150 HCTs regulative network and shared with ones of MYB145 and MYB148, named “recognition of pollen”, supporting in this way much more a realistic involvement of this gene in grapevine mechanisms in which pollen is recognized and either accepted or rejected by cells in the stigma and the fact that these TFs stigma-related could regulate these molecular interactions. In the subsequent step of the fertilization process, it seems to enter VIT_10s0042g00700 because of the function observed in its putative ortholog in *Arabidopsis*, which is involved in the pollen tube perception of the female signal. In this regard, At4g08850 (*MDIS1-INTERACTING RECEPTOR LIKE KINASE2, MIK2*) encodes a receptor kinase that forms a complex with MDIS1/MIK2 and binds LURE1, the female pollen guidance chemi-attractant (Wang et al., 2016). Properly at this point, there are two other HCTs fitting perfectly, namely VIT_00s0404g00100 and VIT_12s0055g00790. The first encodes the Syntaxin of plants 124, whereas the second for the Clathrin assembly protein 3 (Navarro-Payà et al., 2022). The expression of both is well-superimposable and was relieved in Corvina stamens and pollen grain during the blooming stage (Fasoli et al., 2012), in flowers of Tempranillo (Grimplet et al., 2019), but also in both male and female flowers of *Vitis sylvestris* (Massonnet et al., 2020). The ortholog of VIT_00s0404g00100 (At1g61290) specifically encodes the Syntaxin-124 and the one of VIT_12s0055g00790 for the Phosphatidylinositol binding Clathrin assembly protein 5A (PICALM5A). For both, it was proven a direct control of the growth of pollen tube tip (Muro et al., 2018; Silva et al., 2010). In addition, also for these orthologs it was noticed the same expression profile likely in grapevine, indeed the expression of the two genes was consistently observed in stamens (Klepikova et al., 2016; Winter et al., 2007) and in pollen grain since the latest phase of microgametogenesis (Honys and Twell, 2004), going on during the germination

and extending until the pollinic tube penetration in stigma hole (Qin et al., 2009). Interestingly, VIT_17s0000g04180 expression was detected in Villard Blanc inflorescences (Kamal et al., 2019) and its ortholog (At3g18550; TCP family), even if was largely described to be required for the auxin-induced control of apical dominance by preventing axillary bud outgrowth and delays early axillary bud development (Aguilar-Martìnez et al., 2007), it seems also to have a role in the control of flowering timing. Indeed, beyond the fact that its expression was specifically relieved in stigma (Klepikova et al., 2016), the overexpression of this gene in *Arabidopsis* significantly delayed flowering under both long-day and short-day conditions and dominant repression by this TCP led to various growth defects. The upregulation of this TF led to more accumulated mRNA levels of *FLOWERING LOCUS C* (*FLC*), a central floral repressor of *Arabidopsis*. It seems functioning in an *FLC*-dependent manner, as its overexpression in the flc-6 loss-of-function mutant failed to delay flowering (Wang et al., 2019). At last - but absolutely not at least - it is pivotal to highlight that a member of MYB150 HCTs is MYB148 (VIT_00s1352g00010), TF related to stigma whorl too and with whom shares 11 common targets. This observation suggests the presence of a network of nested and mutually interpenetrating regulatory relationships. Being able to reconstruct the individual actors of these networks allows us to shed light on the individual regulatory mechanisms which, understood individually in the first instance and then summarized in a general perspective, give us the possibility of having an overview of the TF-mediated modulation of biological processes and their interdependence co-occurrence. Perfectly in line with the targets, with the GO enrichment and - primarily - with the other TFs stigma related, also the most representative binding sequence (Fig. 21.3B) of MYB150 cistrome shows similarity (E value = 1,0345E-06) for both the pollen-specific motif POLLEN1LELAT52 and the light trigger related one DOFCOREZM.

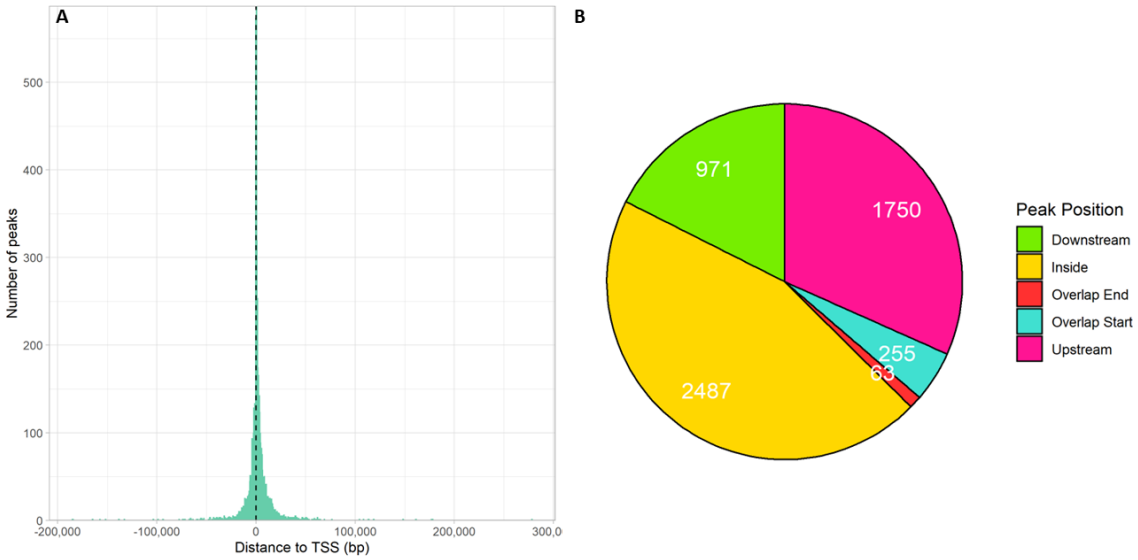


Figure 19.3. MYB150 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

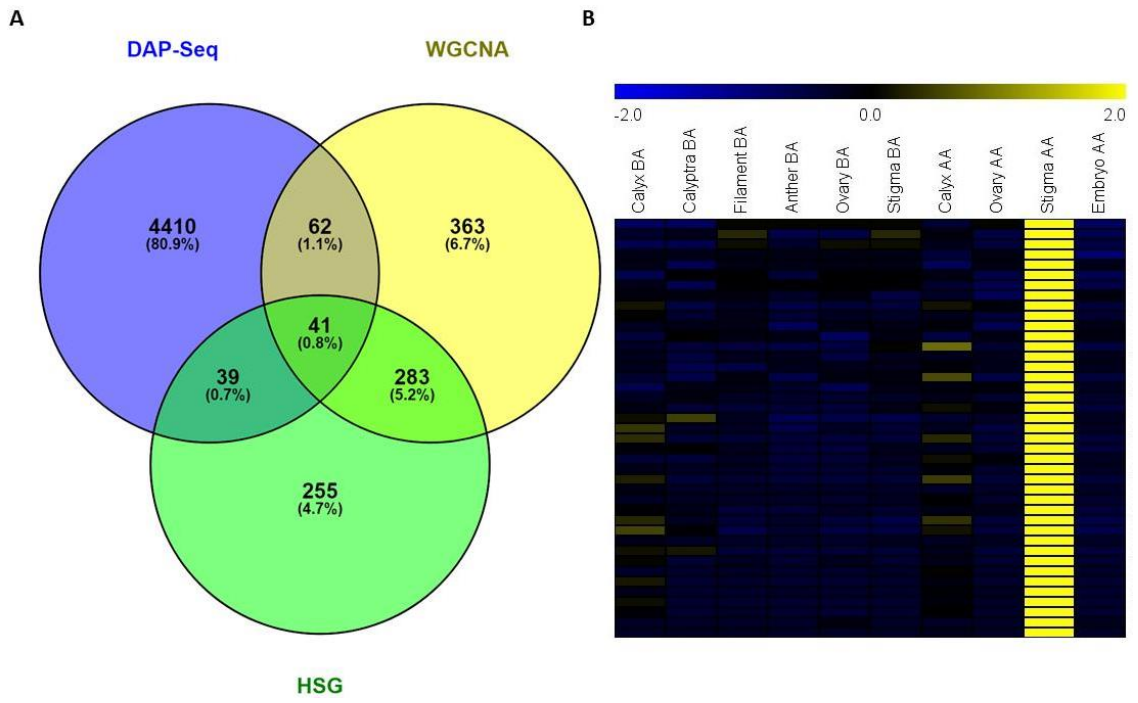


Figure 20.3. Individuation of MYB150 High Confidence Targets. (A) Shared genes between DAP-Seq and Stigma_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

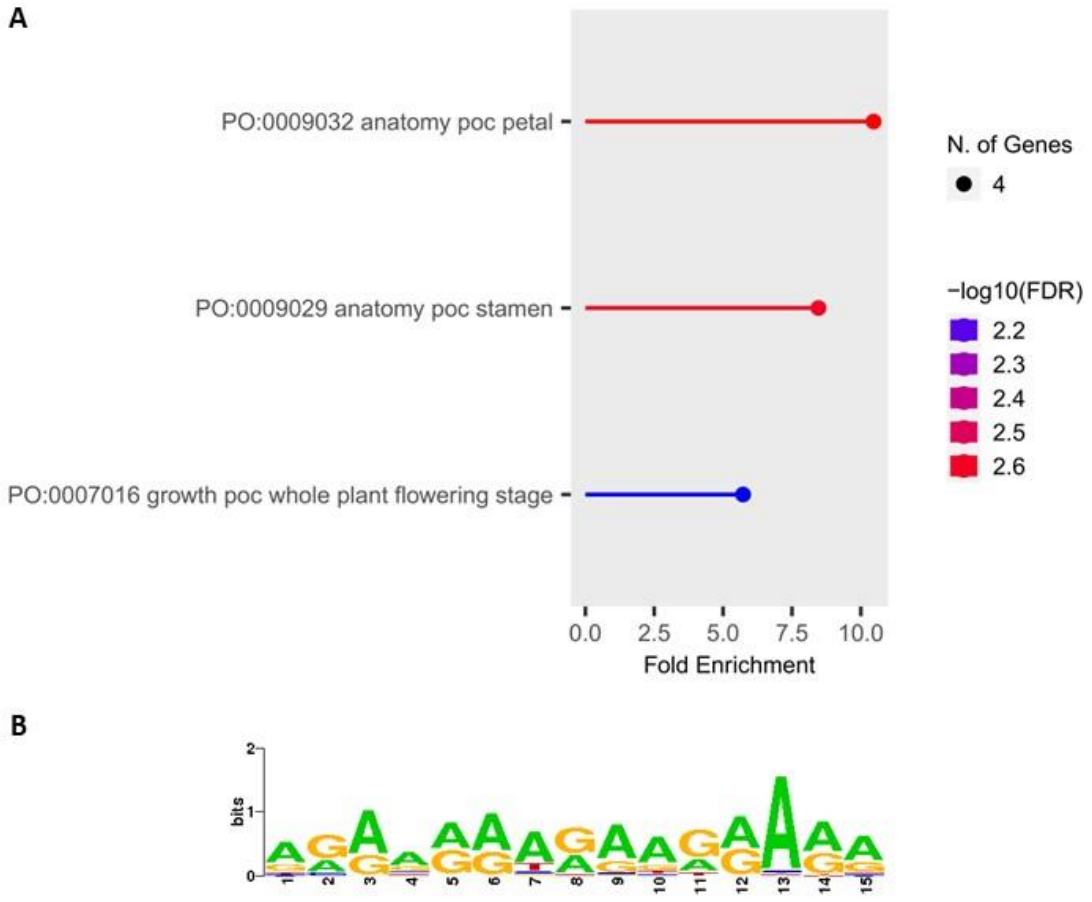


Figure 21.3. (A) GO enrichment analysis performed on MYB150 HCTs: Plant Ontology (PO) categories. (B) *De novo* motif discovery analysis on the whole cistrome of MYB150 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.7 MYB154

Peak calling of analysis of MYB154 resulted in 1498 different binding events. Analyzing the distribution, 34% of the total amount was seen to belong to the upstream region of the relative assigned gene, 6% overlaps the TSS, 42% is referred to the inside of the gene, 1% overlaps the end point and 17% was detected downstream the gene assigned (Fig. 22.3B). The HCTs network related to MYB154 was defined by intersecting DAP-Seq peaks with Stigma_AA of HSG analysis and brown module of WGCNA, obtaining in this way 14 different targets (Fig. 23.3A). In Fig. 24A and B, are reported the enriched categories referred to HCTs list. 8 out of 14 targets are shared with the HCTs network of one or more Stigma_AA related TF. In particular, 1 is shared with MYB145, 2 with MYB148, 1 with MYB150, 1 with both MYB145 and MYB148, 1 with both MYB148 and MYB150 and 2 with MYB143, MYB148 and MYB150. In detail, the latter two

deserve an honorable mention. One is the Ethylene receptor 2 (*ETR2*; VIT_06s0004g05240) and the other is the peroxidase PrxIII32a (VIT_01s0010g01090). Both belong to those genes affecting hormonal balance and signaling, and their shared presence in several regulative networks of stigma related TFs confirms the fundamental importance of these molecular mechanisms in the biological processes characterizing stigma and pollen development and intercommunication. Supporting this, it is interesting to look at HCTs GO enrichment of Molecular Function (Fig. 24.3B): the three most enriched categories are related to hormonal activity, namely “ethylene binding”, “ethylene receptor activity” and “alkene binding”. The communication between floral whorls could conceivably take place not only between stigma and pollen, but also could be extended to all the floral organs that synergically contribute to give rise to the anthesis phenomenon, coordinating the correct development of every single actor to which follows the right timing of flowering. In this regard, it is interesting to point out the presence in MYB154 HCTs list of the gene VIT_07s0151g00700. Its expression was detected in flowers of Villard blanc (Kamal et al., 2019) and it was observed belonging to the ontological category named “megagametogenesis” (GO:0009561). To the same GO category belongs the putative ortholog At2g48140 whose expression was detected in *Arabidopsis* developing embryo (Casson et al., 2005). Moreover, it regulates the ubiquitin-mediated proteolytic pathways and was proven to be essential in female gametophyte development. This last was observed to arrest at the four-nuclear stage in disrupted phenotype mutants (Pagnussat et al., 2005). The presence in stigma tissues of an embryo-related regulator could be precisely explained with the afore proposed intercommunication between the different floral whorls. In this case, it is considerably realistic to think that the stigma coordinates pollination process by receiving signals from pollen and ovule, and that - on the basis of these - it regulates the correct unfolding of those molecular mechanisms which will lead to the growth of the pollen tube until it reaches the ovule in the right moment in which it can be efficiently fertilized. Related to that - and closely related to pollen development - it is noteworthy the presence of VIT_11s0052g01150 in the MYB154 network. This gene is a nicotianamine synthase (Navarr-Payà et al., 2022) as well as for the ortholog At1g56430, the Nicotianamine synthase 4 (*NAS4*), which synthesizes nicotianamine, a polyamine which serves as a sensor for the physiological iron status within the plant, and/or might be involved in the transport of iron. Its activity is involved in pollen tube growth in *Arabidopsis thaliana* (Schuler et al., 2012). Finally, we report the presence of the TF NAC44 (VIT_11s0052g01150), which in *Arabidopsis* was described as ABA-signaling modulator (Jensen et al., 2008). Exactly like MYB145 and MYB150, also the most enriched motif of MYB154 cistrome (Fig. 24.3C) showed similarity for both POLLEN1LELAT52 and DOFCOREZM (E value = 4,3428E-06).

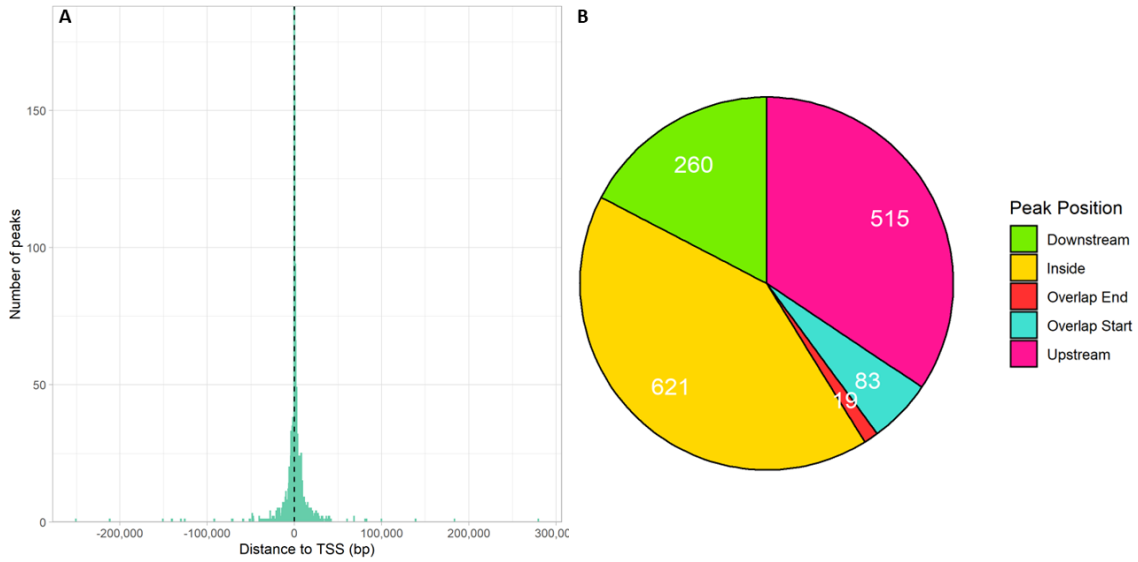


Figure 22.3. MYB154 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

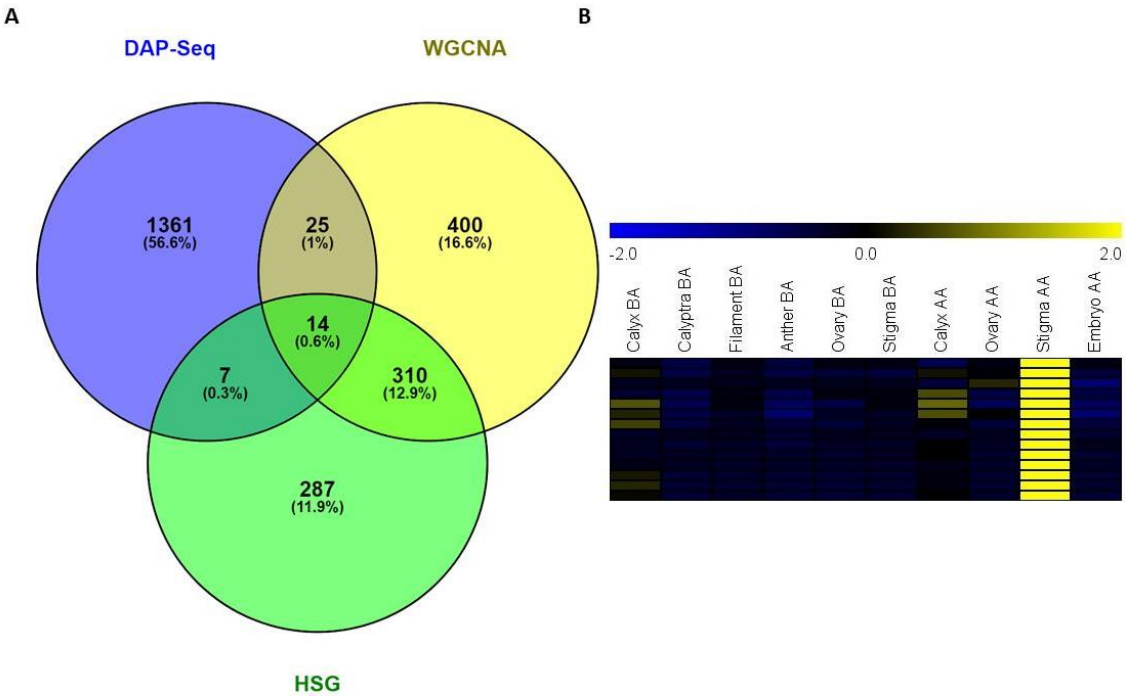


Figure 23.3. Individuation of MYB154 High Confidence Targets. (A) Shared genes between DAP-Seq and Stigma_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

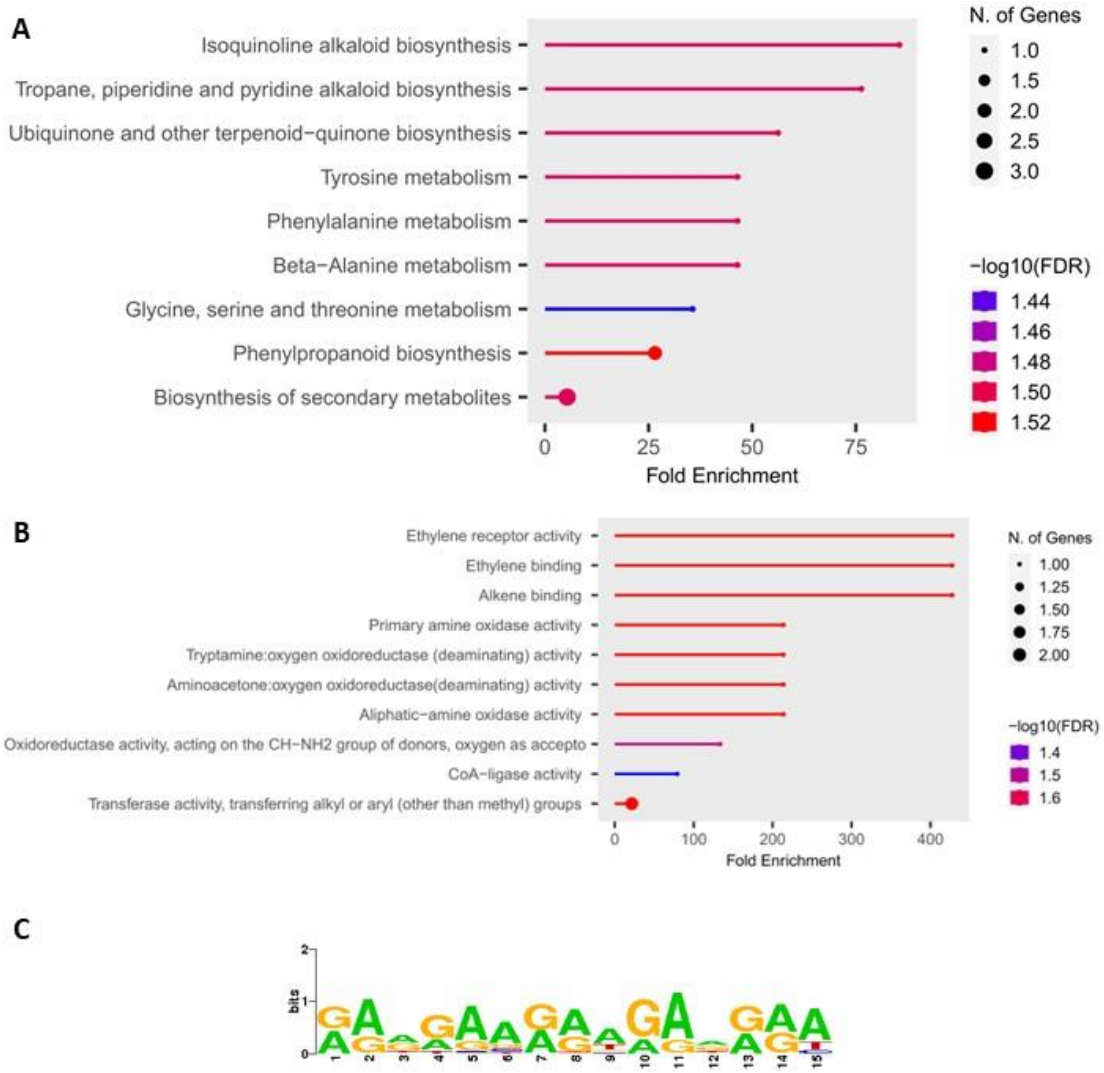


Figure 24.3. GO enrichment analysis performed on MYB154 HCTs: (A) KEGG categories and (B) Molecular Function categories. (C) *De novo* motif discovery analysis on the whole cistrome of MYB154 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.8 MYB192

DAP-Seq identified 268 gene assigned binding events. In particular, 32% of the total peaks were identified in the upstream region, 2% comprehend the transcription starting site, 49% inside the gene structure, 2% in correspondence of the end site and 15% is the fraction of the peaks found downstream the assigned gene (Fig. 25.3B). No HCTs was found from the triple crossing of DAP-Seq targets and Ovary_AA clusters of HGS and WGCNA, but two subsets of 2 and 6 genes were defined by individually intersecting peaks with Ovary_AA HSG and WGCNA lightcyan module respectively (Fig. 26.3A). 1 gene is shared with targets of the other ovary-related TF MYB91A. In

the MYB192 gene set, VIT_13s0067g00990 belongs to the DAP-WGCNA hub genes list and encodes for a Clp R domain-containing protein. This gene was observed to activate its expression in carpel and continue increasing until the phenological phase of post-fruit set in skin, pericarp, and flesh tissues of Corvina berry (Fasoli et al., 2012). A similar trend was detected in Tannat cultivar, in which among different leafy and berry tissue samples across 1-, 3- and 7-weeks post anthesis, it was observed to reach its maximum expression in berry skin 7 weeks after the full flowering (Da Silva et al., 2013). Moreover, VIT_13s0067g00990 was found to be highly expressed in berry in correspondence to the fully veraison stage in three out of the three tested Muscat table grape Italia, Xiangfei and Zaomeiguixiang cultivars (Sun et al., 2019). In our RNA-Seq experiment (Chapter II), VIT_13s0067g00990 is highly expressed in Ovary and Calyx tissues after anthesis time point, the two main whorls which development will give rise to the properly said berry. Based on these observations, it is therefore conceivable an involvement of this gene in early berry development molecular mechanisms, in a period included between ovule fertilization and early ripening, with its maximum exploitation within the phase of starting veraison. VIT_13s0019g00310 transcript was observed to be accumulated in berry skin of Xiangfei Muscat cultivar in correspondence of harvest ripeness phenophase (Sun et al., 2019). Since it encodes for an ABC-transporter protein, its expression is consistent with the molecular function of that family of proteins featured by an import function of molecules inside the cell properly in this time point, inasmuch as this last is typically characterized by an intense activity of production and translocation of secondary metabolites. The DAP-HSG cross found gene VIT_05s0020g01350 has an interesting ortholog in *Arabidopsis thaliana*, At5g47640, which is called *NF-YB2* (*NUCLEAR FACTOR Y, SUBUNIT B*). This gene encodes for a transcription factor which binds CONSTANS (CO/B-BOX PROTEIN1 BBX1), a master flowering regulator, forms a trimer to efficiently bind the CORE element of the *FLOWERING LOCUS T* promoter, and correctly regulates the flowering time depending on long photoperiodism. *nf-yb2* Mutants reveal a mis regulated flowering timing (Gnesutta et al., 2017). A mention of merit is particularly due to VIT_04s0008g05430 whose ortholog in *Arabidopsis* is At3g49500 (*RDR6*). This gene is an RNA-dependent RNA polymerase 6 with some very interesting observations achieved. A mutation in *RDR6*, which functions in trans-acting short interfering RNA (ta-siRNA) production, was found that simultaneously enhances self-incompatibility (SI) and generates an overlong pistil neck which causes the allocation of the stigma above the anthers (Tantikanjana et al., 2009). From an evolutionary point of view, these observations could be emblematic for the explanation on how the changes in floral architecture can occur rapidly upon loss of SI during evolutionary switches from out-crossing to self-fertility as well as the coordination of carpel elongation and SI (Foxe et al., 2009). In this case, instead of discussing the most representative binding sequence of

MYB192 cistrome (Fig. 27.3A), we want to highlight the enriched GO category "CCAAT-binding factor complex" (Fig. 27.3B). In eukaryotic promoters, CCAAT-box motif is bound by HEME ACTIVATOR PROTEIN2 (*HAP2*), which is a subunit of the HAP2/HAP3/HAP5 trimeric complex. On turn, CONSTANS (bound by the ortholog of the HCT VIT_05s0020g01350, as described by Gnesutta et al., 2017 and reported before) promotes *Arabidopsis* flowering and interacts with *AtHAP3* and *AtHAP5* in yeast, *in vitro*, and *in planta*. Mutations in *CONSTANS* delay flowering as well as for the overexpression of *AtHAP2* or *AtHAP3* throughout the plant or in phloem companion cells, where *CONSTANS* is expressed. This phenotype was correlated with reduced abundance of *FLOWERING LOCUS T* mRNA and no change in *CONSTANS* mRNA levels (Wenkel et al., 2006).

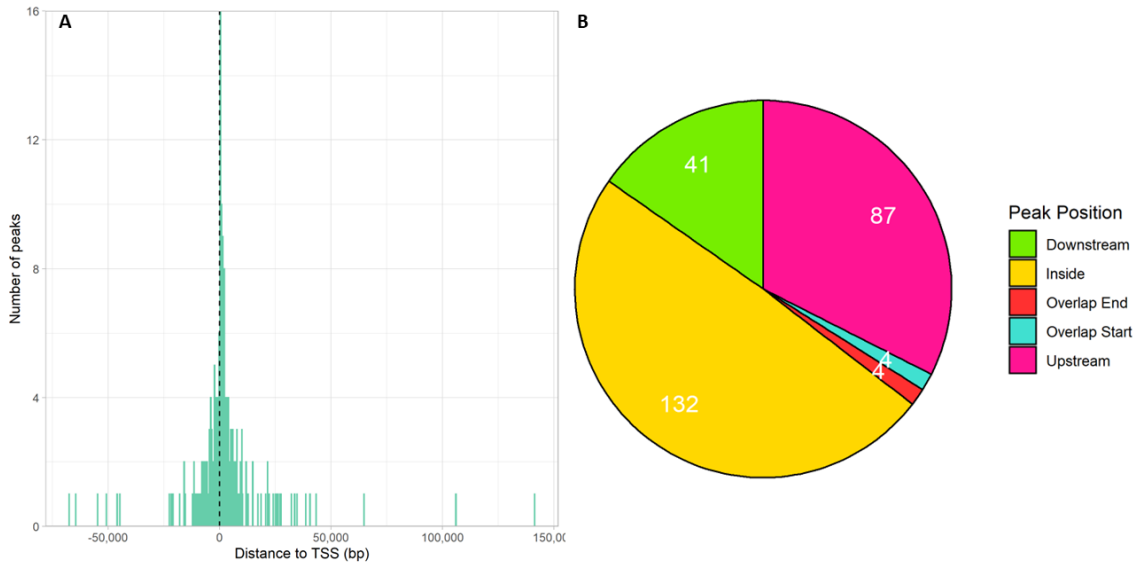


Figure 25.3. MYB192 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

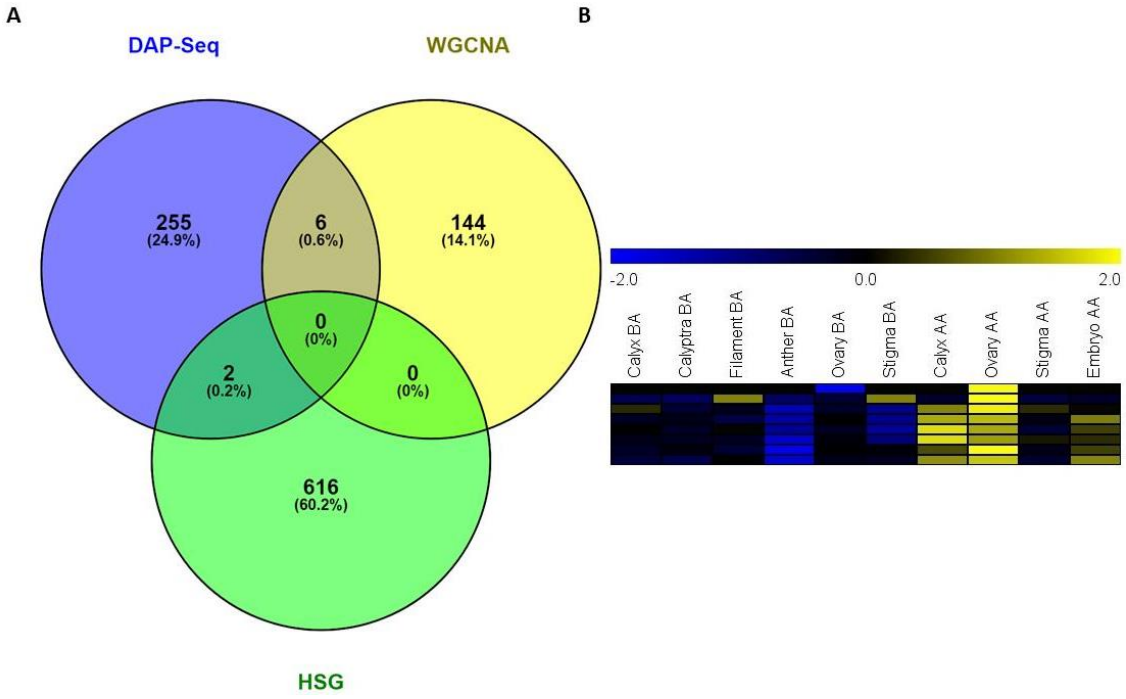


Figure 26.3. Individuation of MYB192 High Confidence Targets. (A) Shared genes between DAP-Seq and Ovary_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

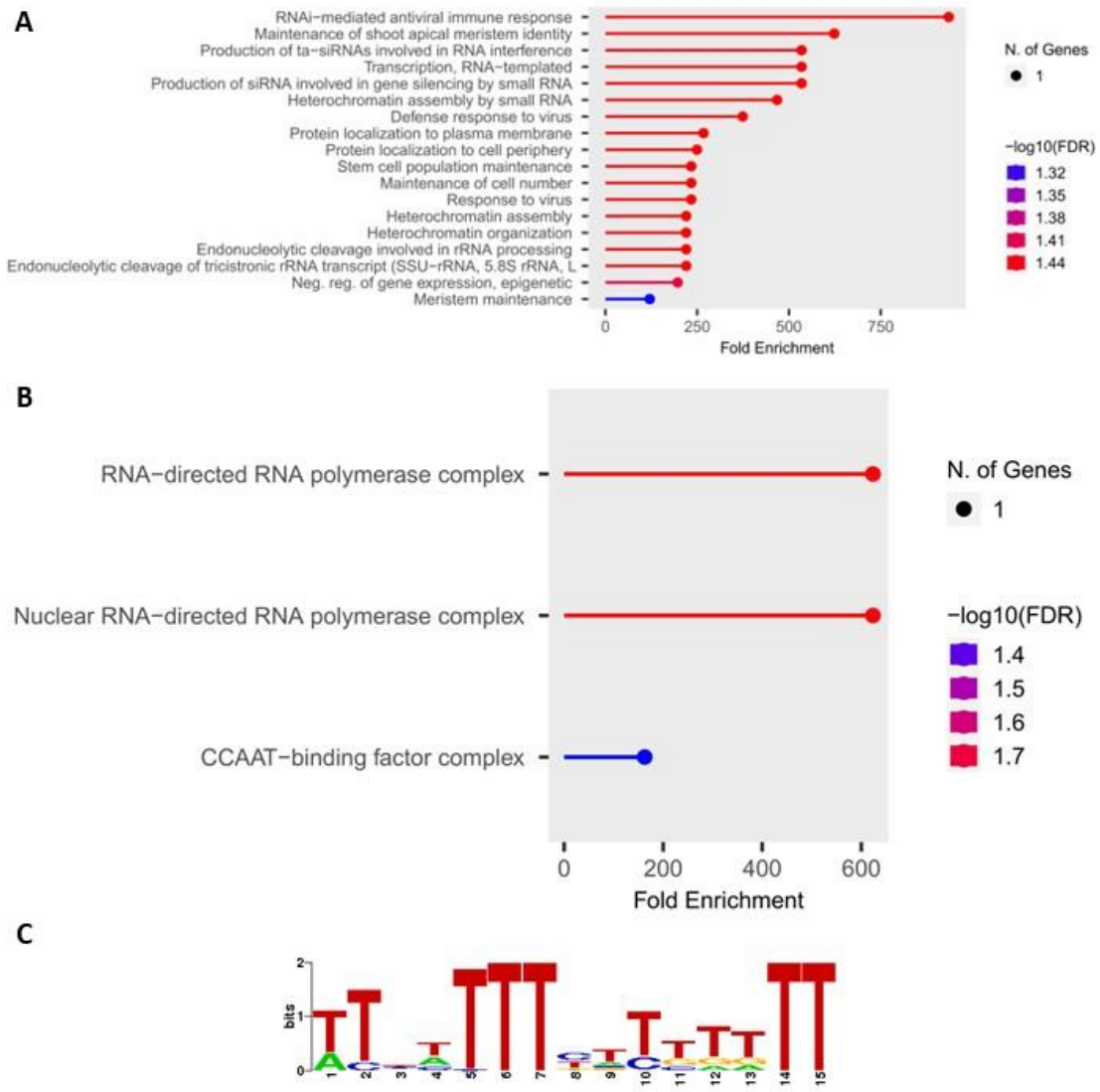


Figure 27.3. GO enrichment analysis performed on MYB192 HCTs: (A) Biological Process categories and (B) Cellular Component categories. (C) *De novo* motif discovery analysis on the whole cistrome of MYB192 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.9 MYBA7

137 peaks were detected for DAP-Seq in MYBA7, distributed with the following percentages: 38% were observed in the upstream region while no binding events were assigned in the zone of the TSS, 40% of the total defined the set falling inside the relative gene, 2% of the peaks overlaps the end and the remaining 20% belongs to the downstream zone of the assigned feature (Fig. 12.3B). No one HCT was retrieved by the intersection of the three datasets namely

DAP-Seq results, WGCNA and HSG therefore the single crossing lists of DAP-Seq with WGCNA and HSG of Stigma_AA cluster were considered. From the DAP-HSG subset it was pointed out one gene and 4 from the DAP-WGCNA brown module.

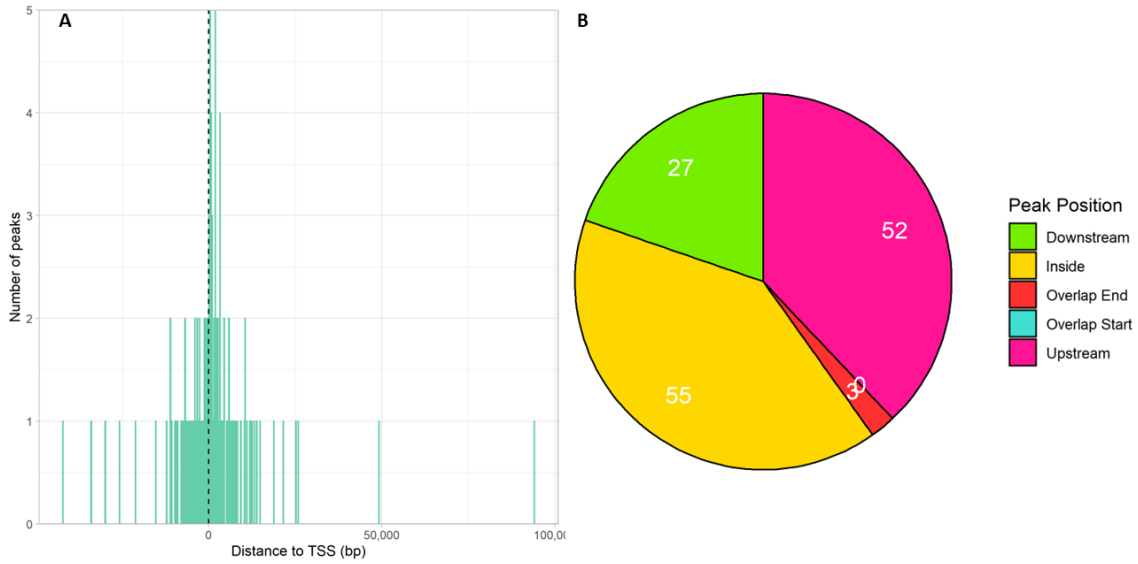


Figure 28.3. MYBA7 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

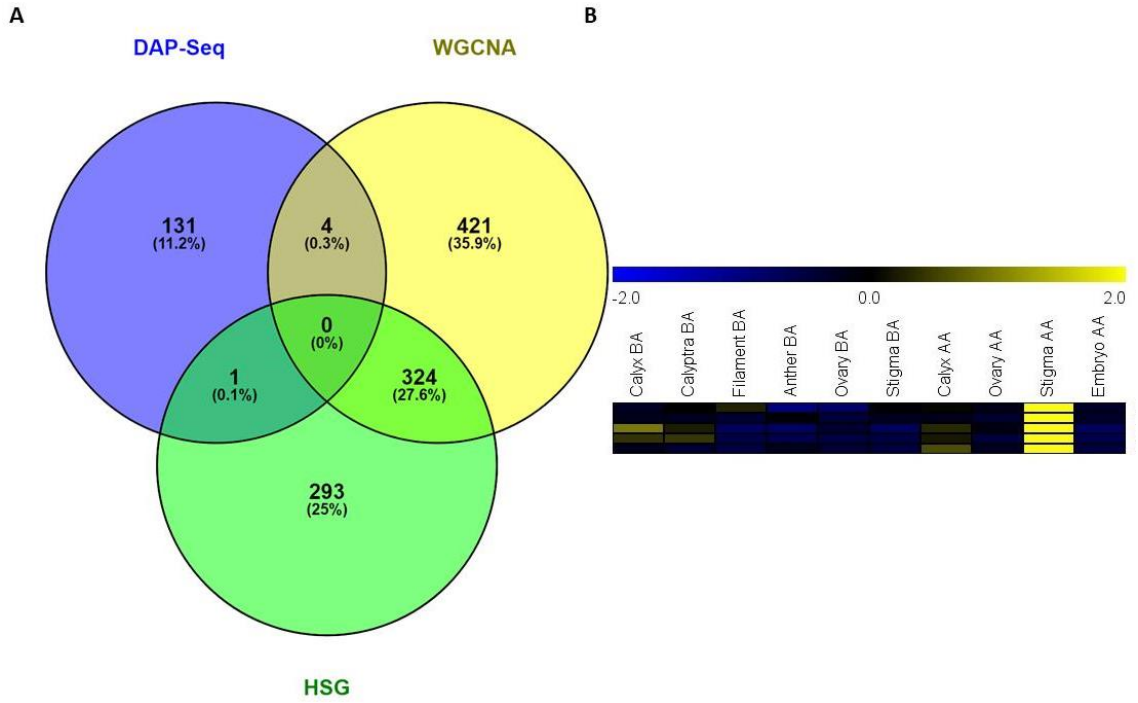


Figure 29.3. Individuation of MYBA7 High Confidence Targets. (A) Shared genes between DAP-Seq and Stigma_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

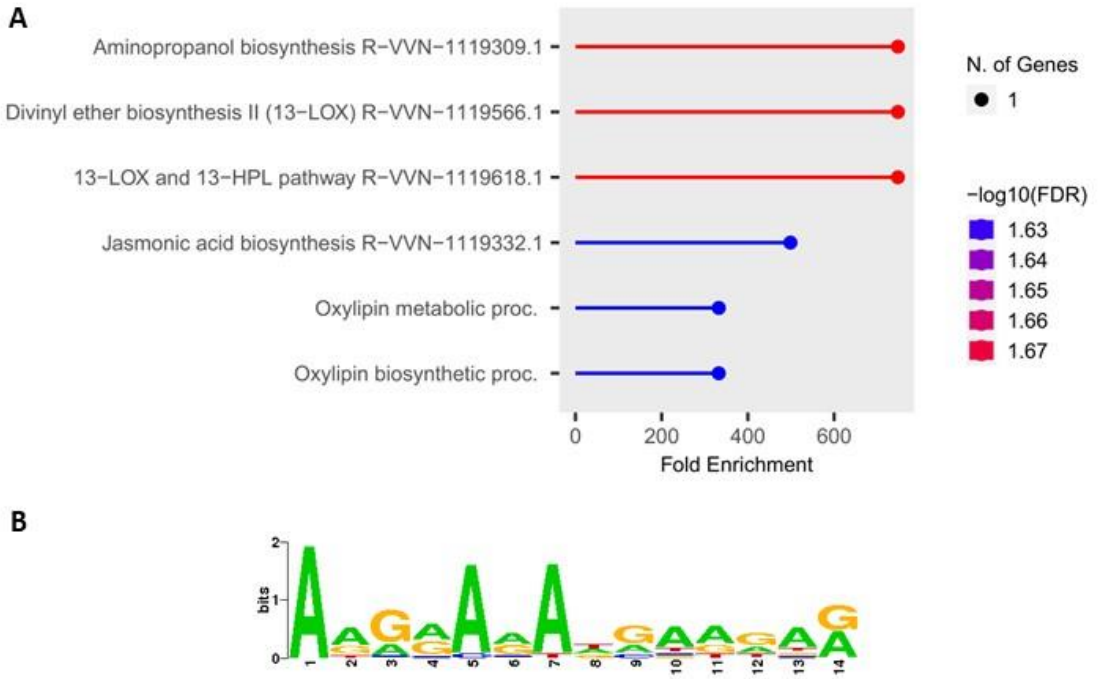


Figure 30.3. (A) GO enrichment analysis performed on MYBA7 HCTs: all available gene sets (B) *De novo* motif discovery analysis on the whole cistrome of MYBA7 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.10 MYBA8

Only 15 binding events resulted in DAP-Seq experiment of MYBA8: 60% of the total number was observed in upstream region, no peaks were detected in correspondence of TSS, 7% falls within the gene length, equally for those observed at gene end, also here 7%, and 26 % downstream the related assigned gene (Fig. 31.3B). Because of the very low number of peaks observed for the cistrome of this transcription factor, the meaning of percentages must be taken very carefully. For this extreme lack, it was individuated only one hub gene from DAP-WGCNA sienna3 module (Stigma_BA specific) list (Fig. 32.3A), namely VIT_06s0004g03840, which ortholog in *Arabidopsis* is a U-box domain-containing protein 35 (*PUB35*; At4g25160) observed to be expressed in all the floral whorl and in pollen (Winter et al., 2007). The observed scarcity in peaks of MYBA8 is therefore consistent with the observation made by Wong et al. (2016) about the possibility that this TF is a pseudogene, since it was encountered expressed but parallelly featured by an incomplete DNA-binding domain.

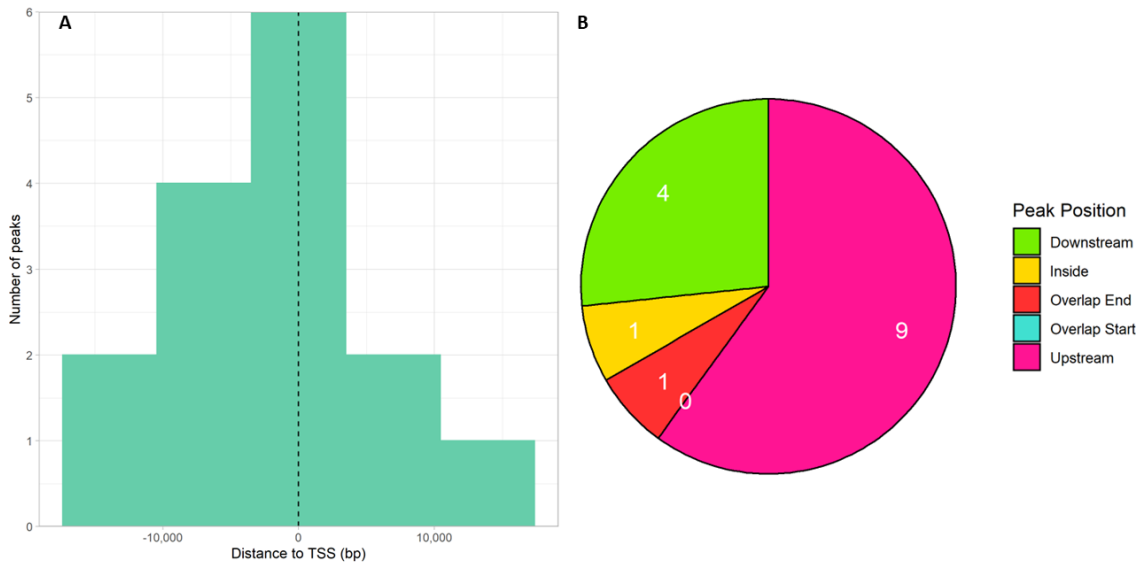


Figure 31.3. MYBA8 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

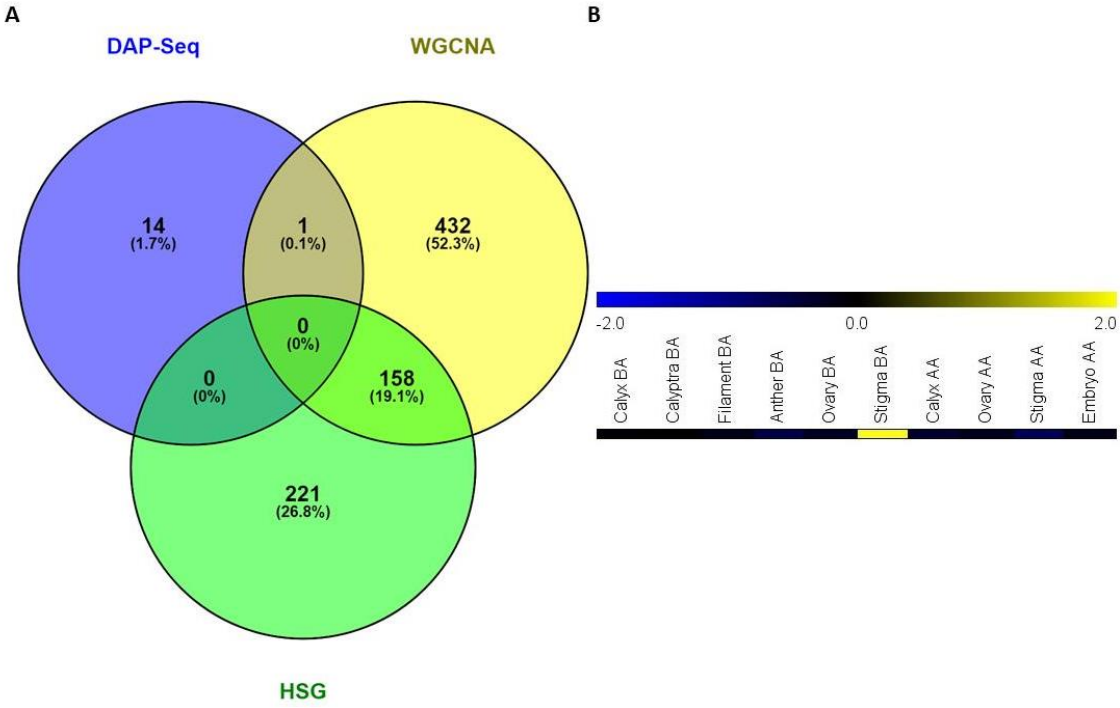


Figure 32.3. Individuation of MYBA8 High Confidence Targets. (A) Shared genes between DAP-Seq and Stigma_BA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

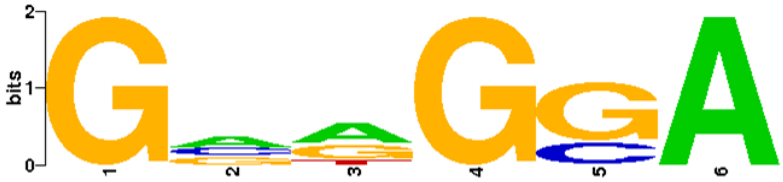


Figure 33.3. *De novo* motif discovery analysis on the whole cistrome of MYBA8 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.11 MYBPA7

Regarding the DAP-Seq analysis performed on MYBPA7, a total amount of 92 binding events was found and peaks repartition showed a fraction of 48% of the global number referred to the upstream region of the assigned gene, 1% comprehend the starting site of the transcription, 29% characterizes the inside of the gene, 3% falls in the region of the gene end and the leftover 19% is localized downstream from the relative gene (Fig. 34.3B). In this case 2 HCTs and 3 genes were retrieved from the single crossing of DAP-Seq data with HSG of Stigma_AA cluster (Fig. 35.3A). VIT_04s0008g02020 belongs to this last achieved subset and encodes for an endoglucanase that was observed to be highly expressed in pollen (Fasoli et al., 2012). The ortholog in *A. thaliana* is At3g43860 and its expression is completely compartmentalized to the pollen grain. Moreover, its transcript was also detected in the whole structure during the phase of pollen germination, until the pollen tube penetrates inside the stigmatic opening and continues its growth through the pistil style, moment at which the expression of this gene is exclusive of the penetrated part of pollen tube (Klepikova et al., 2016). Since in our case the grapevine ortholog belongs to the network of MYBPA7 TF which in turn was observed to be specific to the stigma tissue, it is supposable a putative role of this gene in pollen-stigma interaction molecular mechanisms aimed to dispatch pollen recognition and signaling-mediated tube development. Perfectly in line with the other TFs related to stigma whorls, also for MYBPA7 the analysis of the best matching binding sequence (Fig. 36.3B) gave back interesting and consistent results. With a similarity rate of 2,1506E-06 (E value), they were retrieved 4 different motifs with a coherent biological meaning. The first one is the yet abundantly discussed DOFCOREZM. The second one is TAAAGSTKST1 and as for DOFCOREZM, it was described to be a binding site for the light related DOF TFs proteins (Plesch et al., 2001). The third *cis*-element found is TATABOX5, a TATA-box type motif described for the first time in the promoter of *GS2* of pea (*Pisum sativum* L.), a gene

encoding chloroplast-localized glutamine synthetase regulated by light. Taking advantage from a GUS-assay, the light-regulated expression of this pea gene promoter detected was observed in both tobacco and *Arabidopsis*, suggesting that the regulatory elements are conserved (Tjaden et al., 1995). Finally, the last observation brought back the oligonucleotide motif POLASIG1 which was associated to polyadenylation process in Pea (*Pisum sativum* L.), rice (*O. sativa* L.) and *Arabidopsis thaliana* (Loke et al., 2005; O'Neil et al., 1990; Joshi, 1987). For the explication about the consistency of the first three *cis*-elements, we refer to the other yet discussed light responsive motifs of the stigma related TFs. On the other hand, for POLASIG1, the explanation could be strictly related to the fact that the polyadenylation process follows typically a more intense cellular activity, also consequent to an external stimulus which enhance the transcription. In this case, the light trigger activity which putatively affects the stigma gene regulation could be that trigger factor enabling polyadenylation.

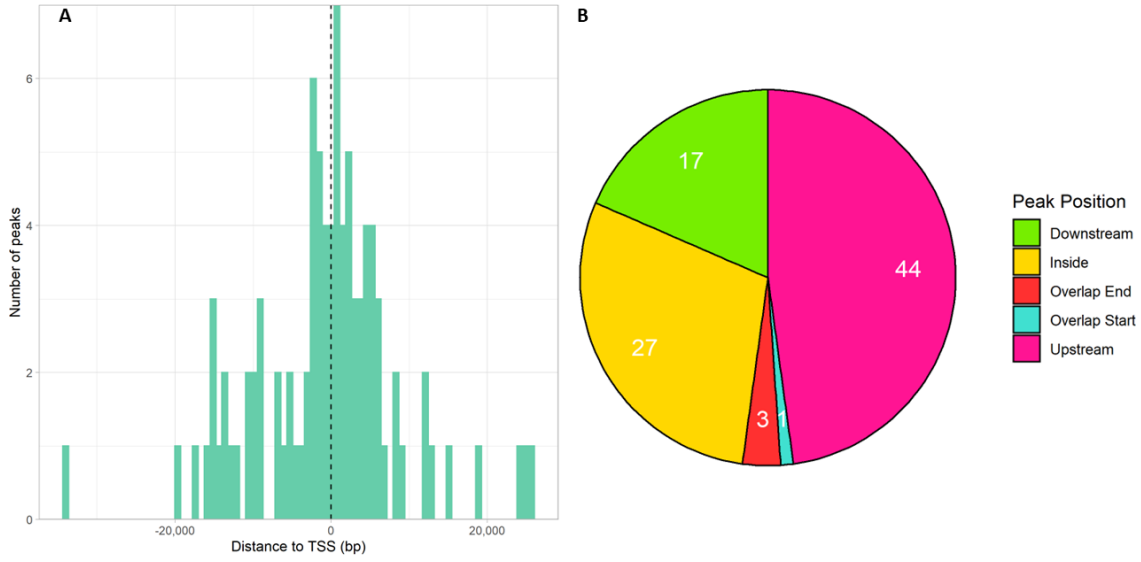


Figure 34.3. MYBPA7 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

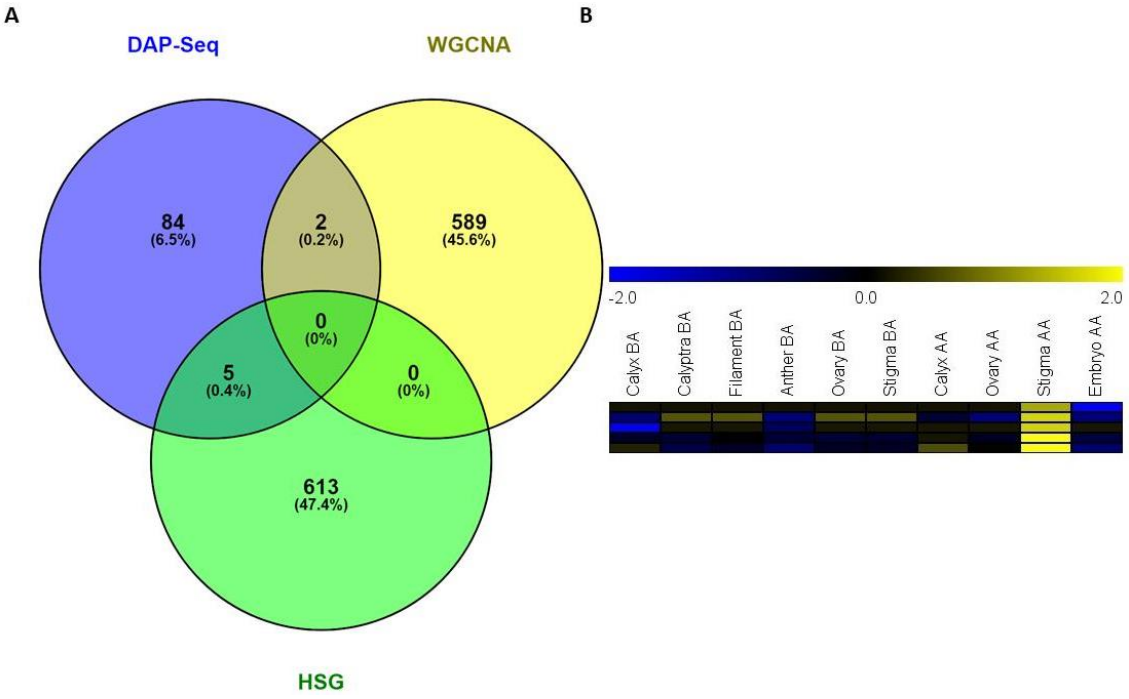


Figure 35.3. Individuation of MYBPA7 High Confidence Targets. (A) Shared genes between DAP-Seq and Stigma_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

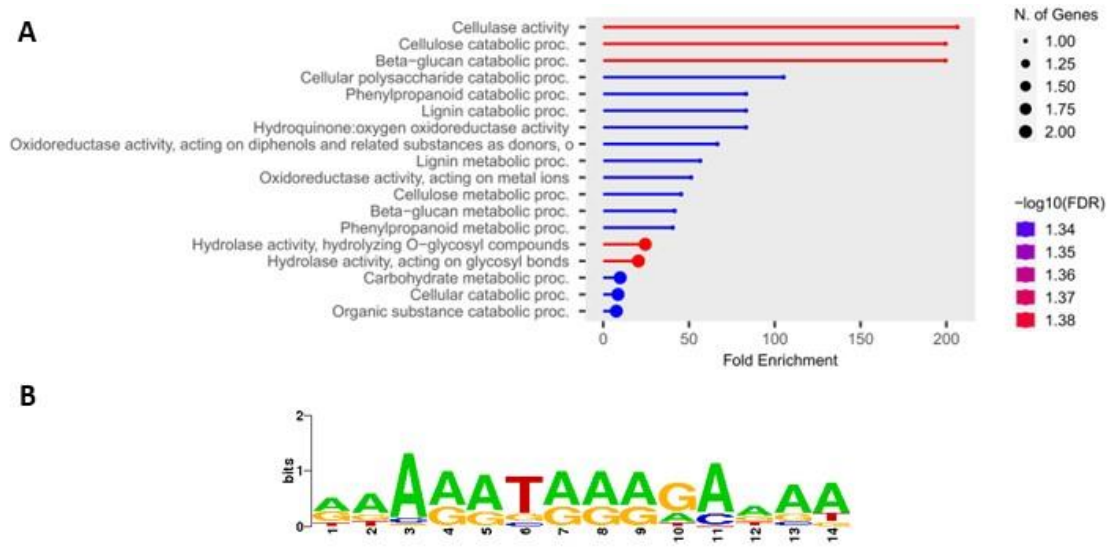


Figure 36.3. (A) GO enrichment analysis performed on MYBPA7 HCTs: all available gene sets. (B) *De novo* motif discovery analysis on the whole cistrome of MYBPA7 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.12 MYBPA9

The peaks assigned to MYBPA9 are 90, 42% of them are referred to the upstream related genetic region, 2% overlaps the TSS, 36% lay inside the gene length, 4% falls in the region of the end and the remaining 16% are located downstream the assigned gene (Fig. 37.3B). In this case, it was relieved only one gene from the intersection of DAP-Seq targets with Filament_BA dataset of HSG (Fig. 38.3A), which surprisingly it is another TF, i.e. MYB186 (VIT_13s0064g00960). This TF is called MYB46 (At5g12870) in *Arabidopsis thaliana* and is involved in the regulation of secondary wall biosynthesis in fibers and vessels (Zhong et al., 2007). This function could be consistent with its presence in the network of a transcription factor related to the filament structure of the anther, tissue that performs a support function for the anther but is also crossed by vessels that ensure the supply of nutrients to the pollen cells and guarantee their correct formation, development, and maturation.

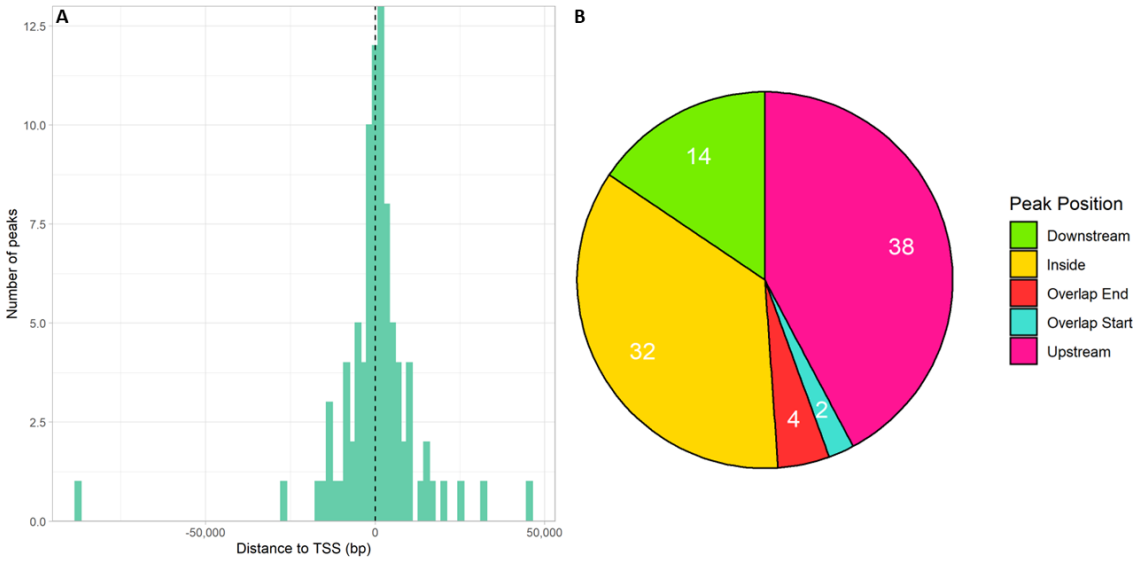


Figure 37.3. MYBPA9 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

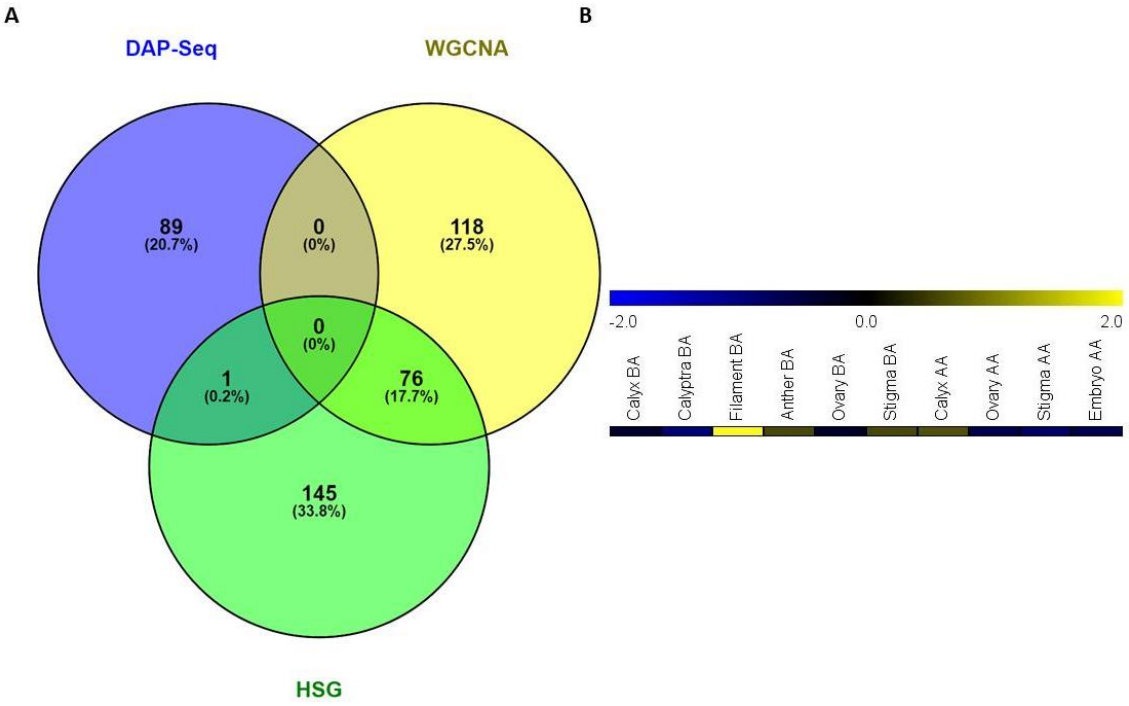


Figure 38.3. Individuation of MYBPA9 High Confidence Targets. (A) Shared genes between DAP-Seq and Filament_BA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

Figure 39.3. *De novo* motif discovery analysis on the whole cistrome of MYBPA9 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.13 MYBPAL3

A total of 55 events of binding defines the cistrome of MYBPAL3 and are spatially distributed as follow: 38% in the upstream region, 4% falls nearby the transcription starting point, 29% is the subset laying within the gene, another 4% was detected in correspondence of the gene end and finally, for the downstream region, 25% was the percentage related to this group (Fig. 40.3B). No HCTs was retrieved from the triple intersection of the three datasets, but a subset of 3 genes was determined by crossing DAP-Seq peaks with WGCNA lightcyan module Ovary_AA specific

(Fig. 41.3A). VIT_10s0116g01670 is an arogenate dehydratase belonging to the shikimate pathway, which transcript was observed to be predominantly accumulated in embryo tissue after anthesis in our expression atlas (Chapter II) but was also reported to be expressed in the seed in the very early stages of berry development, till fruit set (Fasoli et al., 2012), suggesting a possible involvement of this gene in embryo development mechanisms to early seed formation period. The expression of VIT_00s1764g00020 was observed to be induced by gibberellin treatment for berry size enlargement in seedless grapevine cultivar Centennial (Bio Project PRJNA239278) and the expression kinetics of this gene shows that it passes to be highly expressed in seed and slightly less in berry pericarp in correspondence of the phenological phase of fruit set (Fasoli et al., 2012), outlining an inverse situation as from this phase it is reached the post fruit set and veraisòn, moment in which the expression is extremely intense in the pericarp, in the pulp and in the skin, but practically zero in the seed, thus highlighting a gradient which progressively shifts the transcript accumulation outwards from the inside of the berry as the development phases progress. Moreover, it was described to increase in berries of Gewurztraminer and Riesling cultivars passing from the whole green berry phase to the mid-ripening one (Bio Project PRJNA378596). The ortholog in *Arabidopsis* encodes for a LRR receptor-like serine/threonine-protein kinase (At1g12460) and its expression was detected increasing in carpel parallelly to the progress of flower development and in the very first stage of seed formation (Winter et al., 2007). VIT_02s0025g04520 was detected to be accumulated in berry at full maturity phenological stage in Victoria cultivar (Bio Project PRJNA507550), in Muscat Hamburg (Bio Project PRJNA419810), in Muscat Superior (Lu et al., 2022) and in pre-veraison Pinot noir berry pericarp over three different growing seasons (Fasoli et al., 2018), but also in Corvina seed, skin and flesh starting from the fruit-set, through the veraison till the early ripening phases (Fasoli et al., 2012). Moreover, it was detected being highly expressed in Cabernet Sauvignon berry in correspondence of EL-36 phenological stage (early ripening phase characterized by an intermediate Brix value according to Coombe, 1995) after an ABA treatment, highlighting in this way its involvement in grapevine berry ripening mechanisms, featured by the increase of ABA levels and decline of auxin hormones ones (He et al., 2021). For these observations, evidences about an involvement of these genes in the developmental processes of the seed starting from the embryo and also of the tissues of the berry are reasonably conceivable also in grapevines, as is already known, in fact, are of paramount importance the relationships established between seed and surrounding tissues for the purposes of their development, in particular by influencing their growth and quality in terms of accumulation of secondary metabolites which characterize the mechanisms of berry ripening.

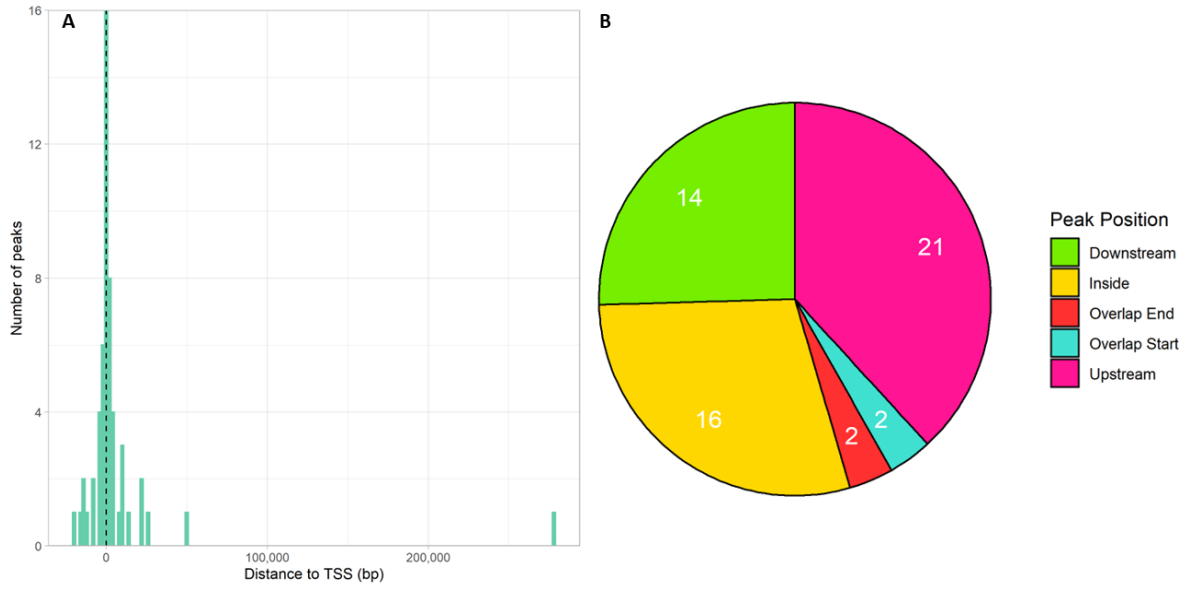


Figure 40.3. MYBPAL3 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

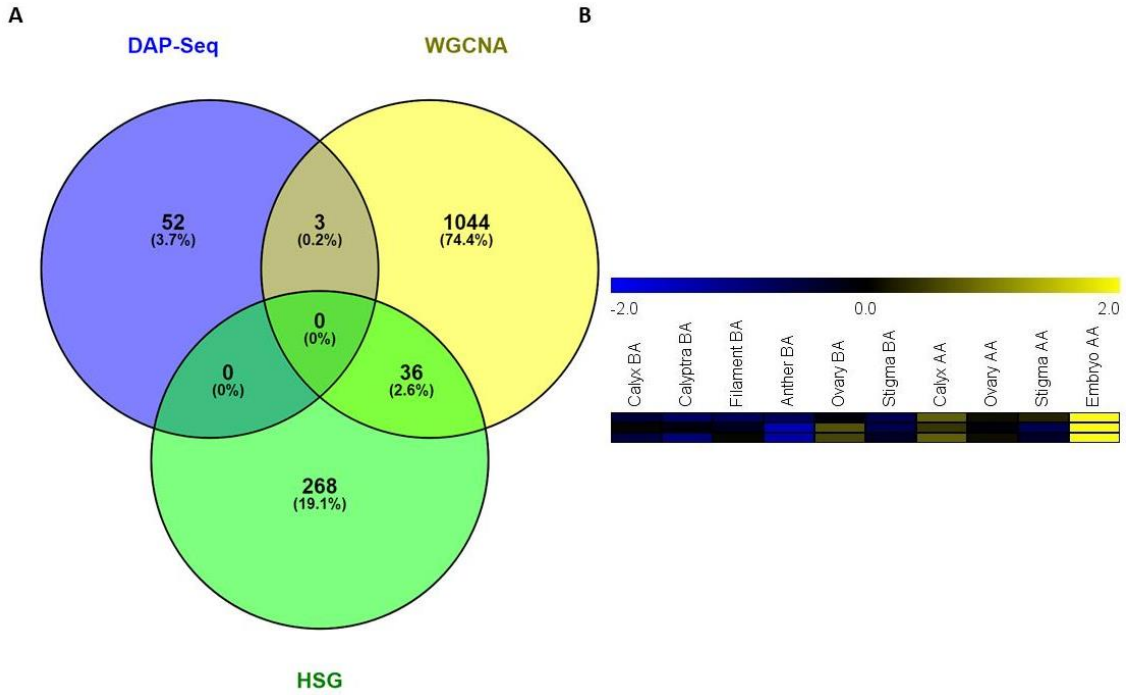


Figure 41.3. Individuation of MYBPAL3 High Confidence Targets. (A) Shared genes between DAP-Seq and Embryo_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

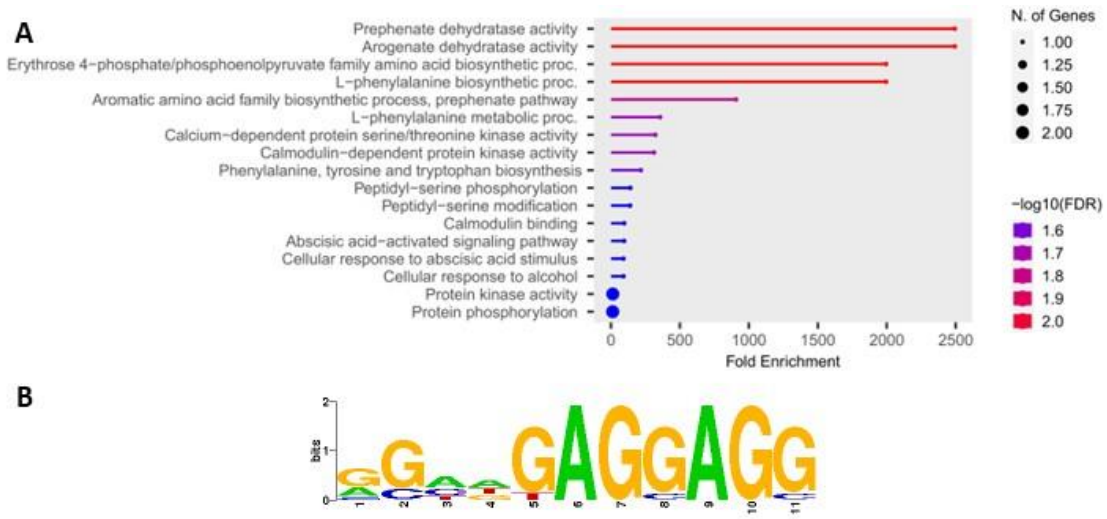


Figure 42.3. (A) GO enrichment analysis performed on MYBPAL3 HCTs: all gene sets available. (B) *De novo* motif discovery analysis on the whole cistrome of MYBPAL3 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.14 NAC62

For the cistrome landscape of NAC62, DAP-Seq analysis identified a global number of 4048 different peaks. More in detail, 32% of the total was observed upstream the assigned gene, 3% comprehends the TSS, 45% was found into the gene structure, 1% overlaps the end and the leftover 45% defines the subset laying downstream the relative gene (Fig. 43.3B). A total number of 14 HCTs was retrieved by the triple intersection of the Ovary_BA specific cluster of WGCNA plum1 module and HSG with DAP-Seq targets, defining in this way the HCTs of the NAC62 network (Fig. 44.3A). Among them, is noteworthy the GDSL esterase/lipase encoded by VIT_08s0007g04300, whose expression was relieved in seed of grapevine (cv. Tannat) 7 weeks after flowering (Da Silva et al., 2013), as well as reported in Fasoli et al., (2012), where the transcript of this gene was detected to be racked up in Corvina seeds increasing with the mid-ripening berry stage incoming. A similar situation was pointed out observing the behavior of the related ortholog (At5g03810) in *Arabidopsis thaliana*, indeed its expression showed high levels at early seed developmental phases, specifically in embryo tissues (Winter et al., 2007). Interestingly, this gene was detected to be overregulated in Kyoho variety berry treated with 5-azaC, demonstrating that the early ripening is promoted by azacytidine treatment, reducing DNA methylation, and suggesting connection with ROS metabolism during the grape berry development (Guo et al., 2019). In addition, the involvement of VIT_08s0007g04300 in early

reproductive tissues developmental processes seems to be supported by the fact that keeping in mind that – in a precise moment – the maturity degree of a cluster in a grapevine branch is directly proportional to the distance of it from the shoot apex, the transcript of this gene is extremely stored in immature inflorescence at first node from the meristematic apex, decrease in the second and is almost insignificant at the third (Bio Project PRJNA149115). These observations appear slightly different from the results of our analysis where it is expressed in pistil, but actually this little discrepancy could be explicable by the fact that since it is precisely the carpel in pre-anthesis time point, the tissues of the ovary proper have not been separated from those of the ovule *de facto* contained in the carpel. VIT_12s0028g03640 is a Bet_v_1 domain-containing protein and shows a similar seed expression pattern to VIT_08s0007g04300 in Corvina expression atlas (Fasoli et al., 2012). In table grape Italia, it reaches a very high expression in berry at the extremely preliminary stages of veraison, where the berry starts to color and enlarge (EL-35; Coombe, 1995), to then disappear completely in the immediately following phenological phases (Sun et al., 2019). This appears to be in line with our findings which see a prominent expression of this gene before the flowering period. Furthermore, the expression of this gene was observed to be induced by azacytidine application on early ripening berries (Guo et al., 2019). Also for the gene VIT_08s0040g03120, it was reported a scalarity in expression depending on inflorescence position in the nodes of Pinot Meunier branches (Bio Project PRNJA149111). Additionally, its expression was exclusively localized in Corvina pistil in the phases immediately preceding and following the anthesis moment (Fasoli et al., 2012), thus corroborating the hypothesized role of NAC62 HCTs in early phases of flowering and berry development. This tissue-temporal definition was also encountered in inflorescence of Cabernet Sauvignon, in which the expression of this gene was at the maximum if compared to berry skin and seed in subsequent timepoints (Koyama et al., 2014). On the other hand, the expression of VIT_08s0040g03120 ortholog (At5g02810) seems to be more prolonged, indeed it covers all the periods of seed formation and development (Winter et al., 2007). Again, VIT_17s0053g00070 is another NAC62-related HCT observed showing exactly the same expression trend of the previously described ones regarding the node position of the inflorescence in shoot, another time describing a reduction of the expression level by gradually moving away from the meristematic apex (Bio Project PRNJA149111). Superimposable results compared to our expression study were verified by Fasoli et al., (2012) in Corvina grapevine cultivar, in which the expression of VIT_17s0053g00070 was totally restricted to carpel close to flowering time. In *A. thaliana*, its ortholog (At1g48370) maintains a high level of transcript accumulation in all the floral whorls even if it remains preeminent in pistil tissues (Klepikova et al., 2016; Winter et al., 2007). On the other side, VIT_10s0042g00940 shows an even earlier expression timing, indeed

it is expressed starting from the floral bud swelling stage, raises unto the bud burst and from that moment it goes on slightly decreasing progressively until the inflorescence well developed phenological phase, moment starting from which on, the transcript levels collapse. The same trend was evidenced in Touriga Nacional (Ramos et al., 2014) and Corvina (Fasoli et al., 2012) and Pinot noir (Rossmann et al., 2020), being in line with our results that propose this gene as early reproductive processes involved. VIT_16s0022g02310 is a Gibberellin 3-beta-dioxygenase whose function was predicted to complete in flower development inferred from biological aspects based on phylogenetic analysis of ancestors (Gaudet et al., 2011). In support of this assertion, there are in fact some studies achieved in *Arabidopsis* on the related ortholog At4g25420 which represents a key oxidase enzyme in the biosynthesis of gibberellin that catalyzes the conversion of GA12 to GA9, via a three-step oxidation at C-20 of the GA skeleton (Phillips et al., 1995). It is highly expressed in stems and inflorescence tissues, and in these lasts, it was described to be involved in the promotion of the floral transition, fertility, and silique elongation (Rieu et al., 2008; Phillips et al., 1995). A noteworthy HCT in the NAC62 network is VIT_17s0053g00740 for the reason why it encodes for another grapevine NAC transcription factor, namely NAC57 (Navarro-Payà et al., 2022). Its expression was surveyed starting from the floral bud burst, continuing in ovary during the anthesis stages and carrying on in flesh and pericarp during berry veraison and ripening, up to the early phase of post-harvest withering, in correspondence of which the expression reaches its apotheosis in the same tissues (Fasoli et al., 2012). An analogous expression pattern was appreciated by analyzing the time course of the *Arabidopsis* GDSL esterase/lipase At5g03610, indeed its transcript accumulation starts at meristematic apex transition from vegetative to reproductive and it goes on along all the flowering periods being totally predominant in pistil among all the floral whorls. In seed, it is intensely expressed at late stages of development (Winter et al., 2007). Its ortholog in grapevine is the HCT VIT_13s0106g00440 and it is expressed in developing embryo from the post- fruit set increasing gradually up to the mid-ripening phase, at which the expression passes at pulp and skin tissues, going on till the middle phases of post-harvest withering (Fasoli et al., 2012). VIT_18s0164g00050 and VIT_11s0016g01250 are both expressed starting from floral bud burst in Corvina cultivar, but if on one side the first gene continues to be expressed in carpel at anthesis and then in the developing seed till the mid-ripening berry stage also confirming our transcriptome profile that sees this gene conspicuously expressed also in Embryo_AA tissue (Fig. 44.3B), on the other hand, VIT_11s0016g01250 stops its transcript accumulation at young inflorescence phenological stage (Fasoli et al., 2012), outlining a precocious involvement in reproductive process. This seems to be confirmed by its ortholog function, indeed At3g13960 is a growth-regulating factor (GRF5) strongly expressed in actively growing and developing tissues,

among which the shoot apical meristem and flower buds, but also in mature flowers (Horiguchi et al., 2005; Kim et al., 2003). Continuing, VIT_06s0004g00150 and VIT_06s0009g00320 are plenty expressed in ovary close to flowering time. The first is still present in berry flesh, increasing from fruit set up to veraison, moment in which is maximally expressed, and the second in developing seed in fruit set and post-fruit set time. Exclusively in tissues of embryo evolving in seed is racked up the transcript of VIT_12s0028g03640 during veraison and mid-ripening stages (Fasoli et al., 2012), finding support also from our observation of its expression in embryo after the anthesis.

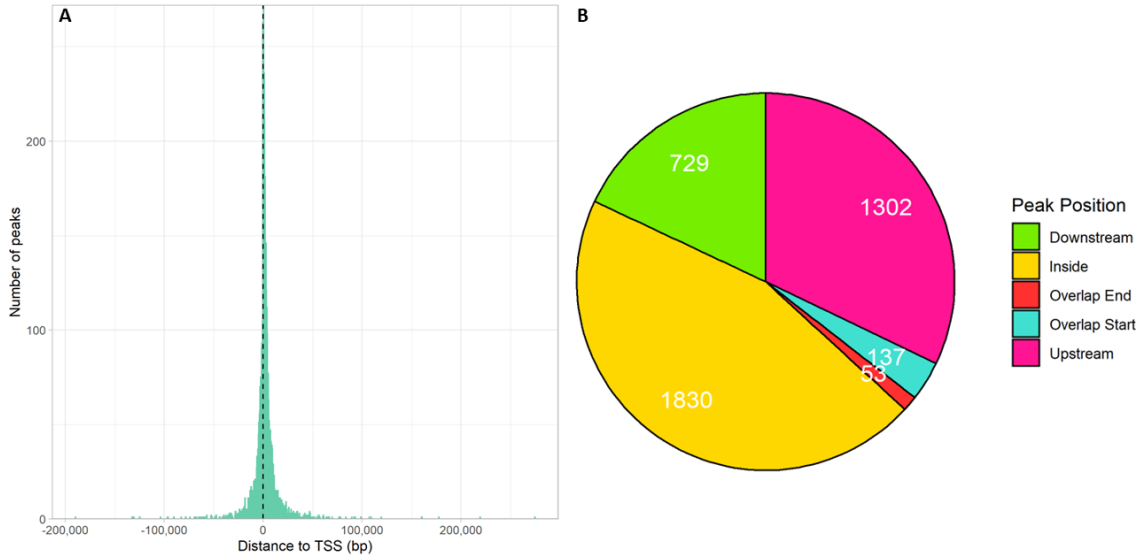


Figure 43.3. NAC62 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

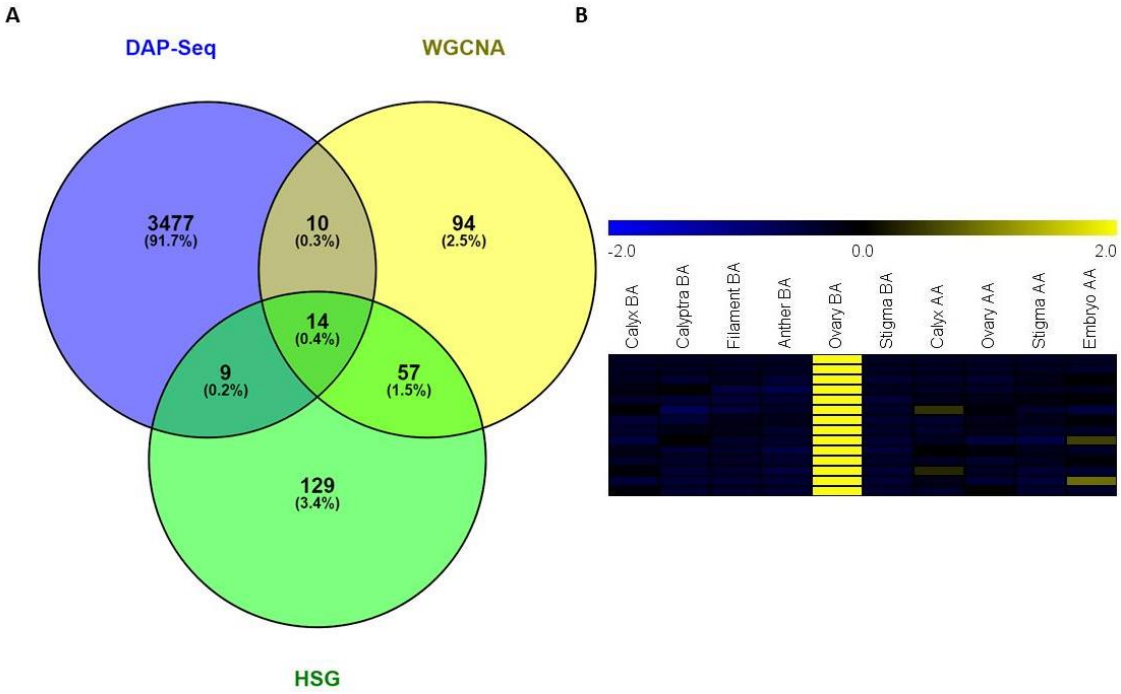


Figure 44.3. Individuation of NAC62 High Confidence Targets. (A) Shared genes between DAP-Seq and Ovary_BA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

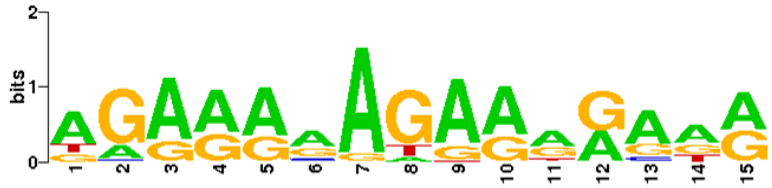


Figure 45.3. *De novo* motif discovery analysis on the whole cistrome of NAC62 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.15 WRKY22

DAP-Seq experiment on WRKY22 evidenced 11097 peaks distributed with the following repartition: 34% in the upstream genic region, 4% around the TSS, 43% was observed laying inside the gene structure, 1% overlaps the stop codon and the remaining 18% is attributed to the part downstream the assigned gene (Fig. 46.3B). It was profiled as a dataset of 24 HCTs belonging to the regulative web of WRKY22 by isolating the shared genes of DAP-Seq results and the clusters of Calyptra_BA of HSG and green module of WGCNA (Fig. 47.3A). In this list, VIT_12s0134g00390 is a protein kinase domain-containing protein (UniProt) and its expression

was seen to be high during the veraison in berry by analyzing its accumulation trend in Muscat Xiangfei (Sun et al., 2019), in Corvina (Fasoli et al., 2012) and Muscat Blanc (Wen et al., 2015). The hypothetical involvement of this gene in calyptra tissues still remains to be explained since in scientific literature there is no evidence in this regard. The same situation was detected for VIT_18s0001g06060, which transcript abundance was observed to be high since veraison moment on for all the ripening time in many grapevine cultivars (Lu et al., 2022; Magris et al., 2019; Fasoli et al., 2012). Interestingly, its expression seems to be light repressed in Sauvignon blanc (du Plessis et al., 2017) and in Zinfandel berry totally reset by Grapevine red blotch-associated virus (GRBaV) infection, by demonstrating the capacity of this pathogen to deeply alter the transcriptional and hormonal regulation of ripening in grapevine (Blanco-Ulate et al., 2017). Moreover, again Sauvignon blanc, it was observed differentially expressed following a nitrogen supply treatment and it was indicated as candidate gene which might be implicated in the biosynthetic pathway of 3SH precursors, since N supply increased Glut-3SH levels in grape berries at late berry ripening stages (Helwi et al., 2016). VIT_01s0127g00470 is an hexosyl transferase and shows the same accumulation pattern of VIT_18s0001g06060 in ripening berry (Lu et al., 2022; Magris et al., 2019), but contrary to the latter, its transcript accumulation was also reported in pre-anthesis flower button of Pinot noir (Rossmann et al., 2020) and specifically in petals of Corvina (Fasoli et al., 2012). Also VIT_18s0001g13770 (Cytochrome P450, family 83, subfamily B, polypeptide 1; Navarro-Payà et al., 2022) seems to have a soft expression in calyptra of Corvina at blooming time, even if its most evident period is the post-harvest withering in berry pericarp and skin (Fasoli et al., 2012). In addition, the ortholog At1g13080 was reported to be expressed in *Arabidopsis* petals and inflorescence stem (Winter et al., 2007; Mizutani et al., 1998). Another CYP protein (Cytochrome P450 87A3-like) is encoded by VIT_18s0122g01480, whose expression finds full correspondence in the before flowering cap of Corvina atlas (Fasoli et al., 2012). VIT_06s0004g01450 is a lipoxygenase (LOX2) whose expression in calyptra finds match in what surveyed in Corvina (Fasoli et al., 2012), but also observing the molecular trend of its ortholog At3g45140, which was detected increasing in a progressive fashion getting closer to the anthesis moment in petals of *A. thaliana* (Winter et al., 2007; Vellosillo et al., 1997). The function of this protein was associated with organ wounding in plants (Moran and Thompson, 2001), for this reason its accumulation in calyptra of grapevine could be supposable as consequences of self-induced abscission line at which follows cap fall event after autogamous fertilization. VIT_08s0040g00790 transcript was detected in Corvina cap before the anthesis as well (Fasoli et al., 2012), likewise very intensely in petals of *Arabidopsis* referring to At5g48570. This last is a peptidyl-prolyl cis-trans isomerase (FKBP65; Meiri et al., 2010) which shows high aminoacidic sequence similarity with another peptidyl-prolyl cis-trans isomerase, namely

FKBP62, which was described to characterize the flower, in particular stigma, sepals, and anthers (Aviezer-Hagai et al., 2007; Vucich and Gasser, 1996), but above all in petals in pre-anthesis (Winter et al., 2007). Peculiar is the case of the calcium-binding protein (CML) encoded by VIT_18s0001g01630 which finds in Calmodulin-like protein 2 (At4g12860) its putative ortholog in *Arabidopsis.* At4g12860 was described to act as potential calcium sensor required for pollen tube attraction in ovule fertilization process (Pagnussat et al., 2005) and indeed its expression was relieved before flowering in pistil and specifically in ovules, but also in petals (Winter et al., 2007). Surprisingly, an analogous situation was described for VIT_05s0029g01410 whose expression was confirmed in calyptra at pre-anthesis stage (Fasoli et al., 2012). Indeed, also its ortholog (At5g61230) was observed regulating interaction relationships between male and male gametophytes. It encodes for the phytochrome-interacting ankyrin-repeat protein 2, which levels in leaves diminishes after transition from the vegetative to the reproductive phase. It accumulates strongly in developmental tissues, is highly expressed in the male (e.g. pollen grains and pollen tubes) and female (e.g. synergids, egg cell and central cell) gametophytes before and during, but not after fertilization. In fertilized ovules, levels decrease rapidly to become undetectable at the stage before the first division of the endosperm. Required for gametophytes development as well as male-female gamete recognition during fertilization, possibly by regulating mitochondrial gene expression (Yu et al., 2010). These shared expressions could also suggest other specific functions of floral interaction mechanisms - in this case at the level of the petal tissue - but the function of this gene is still not clear in this sense. This situation in which two genes expressed in male and female gametophytes, which are involved in the genetic and molecular processes of interaction and recognition between the two, and which are simultaneously reflected in the expression in calyptra, could suggest a hypothetical third actor, the calyptra precisely, which is strictly interconnected to the fertilization mechanisms due to the containing function it plays, which in many cases ensures cleistogamy in grapevines. For this reason, a synchronization at the molecular level of the cap fall is reasonably conceivable, which is chronologically subordinated to the occurred fertilization of the ovule by the pollen grain, and that consequently the ovule-pollen grain communication actually could also include petal tissues.

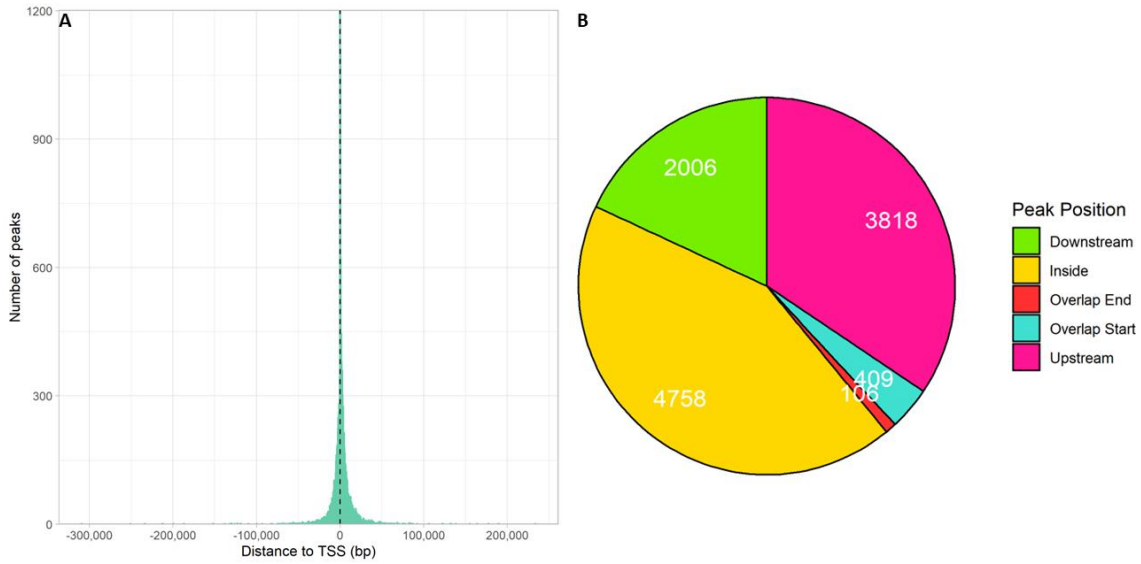


Figure 46.3. WRKY22 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

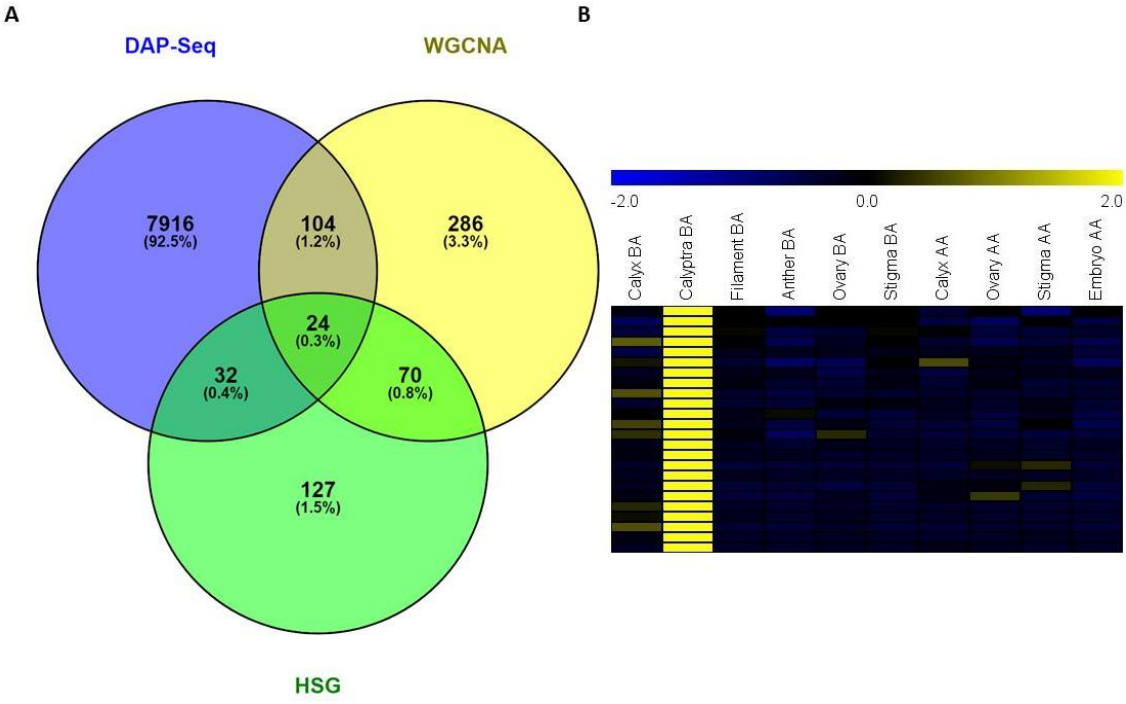


Figure 47.3. Individuation of WRKY22 High Confidence Targets. (A) Shared genes between DAP-Seq and Calyptra_BA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

Figure 48.3. *De novo* motif discovery analysis on the whole cistrome of WRKY22 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.16 WRKY15

WRKY15 highlighted 25096 distinct binding events, 35% of them falls upstream the assigned gene, 3% showed a localization nearby the transcription initiation point, 41% was observed along the gene length, 1% in correspondence of gene end and downstream the related gene 20% of the total peaks (Fig. 49.3B). The related HCTs network outlined 39 members and it was compiled by crossing the cluster of Calyptra_BA of HSG and WGCNA green module with WRKY15 DAP-Seq (Fig. 50.3A). Interestingly, 14 out of 39 are shared with WRKY22, TF Calyptra_BA linked too, empowering the consistency of our analysis workflow. Among these shared targets, some genes deserve to be pointed out, such as *CYP87A3-like* (VIT_18s0122g01480) clearly expressed in calyptra of Corvina (Fasoli et al., 2012). Moreover, since VIT_12s0134g00390 and VIT_01s0127g00470 – a protein-kinase and a hexosyl transferase, respectively – appear in the regulative webs of two different TFs calyptra-related and, as before reported, are turned on during berry veraison and ripening, it therefore seems clear that these findings are not the result of chance and their involvement in petals identity definition appear to be more realistic. Anyway, further insights are needed better to understand this putative involvement in the specific floral whorl molecular mechanisms. Finally, the last shared target of pivotal mention is VIT_18s0001g01630. It is the calcium-binding protein which ortholog was found expressed in pre-anthesis petals and ovules of *Arabidopsis* (Winter et al., 2007), and was described to act as potential calcium sensor required for pollen tube attraction in ovule fertilization process (Pagnussat et al., 2005). The presence of this gene in WRKY22 and WRKY15 networks corroborates the proposed hypothesis of molecular communication of grapevine calyptra with male and female gametophytes in order to establish a correct timing of the cap fall subordinated to the occurred fertilization (as proposed at the end of "3.3.15 WRKY22'' paragraph). In addition, this theory appears to be consistent if we consider the WRKY15 HCT VIT_17s0000g08760 that clearly appears accumulated in grapevine cap (Fasoli et al., 2012). *GSO2* (LRR receptor-like

serine/threonine-protein kinase; At5g44700) is the putative ortholog for which - during embryogenesis - the expression was observed to be uniform from the globular embryo to the mature cotyledonary embryo stages, except in the suspensor tissue. It was proposed an involvement in the nuclear division phase of megagametogenesis, specifically in epidermal surface development in embryos, even if its expression -again- was detected in pollen grain (Tsuwamoto et al., 2007; Pagnussat et al., 2005). Similar expression pattern was highlighted in VIT_18s0001g15650, being intensely expressed in calyptra and in developing embryos (Fasoli et al., 2012). Its ortholog At1g78780 is hugely accumulated in ovules, ovary, and stigma of *Arabidopsis* (Winter et al., 2007). Curiously, among the HCTs there are two WRKY proteins: VIT_04s0069g00980, namely WRKY12, and VIT_06s0004g00230 that is WRKY15 itself. The analysis of the most enriched binding motif of WRKY15 cistrome (Fig. 51.3B) pointed out consistency with the strict molecular biology of WRKYs family, indeed it was evidenced high similarity (E value = 1,2309E-06) with the W-boxes of other previously described TF of this group in several plant species, like WBBOXPCWRKY1, WRKY71OS, WBOXATNPR1, WBOXNTERF3 and ELRECOREPCRP1 (Nishiuchi et al., 2004; Zhang et al., 2004; Chen et al., 2002; Eulgem et al., 2000; Ishiguro and Nakamura, 1994).

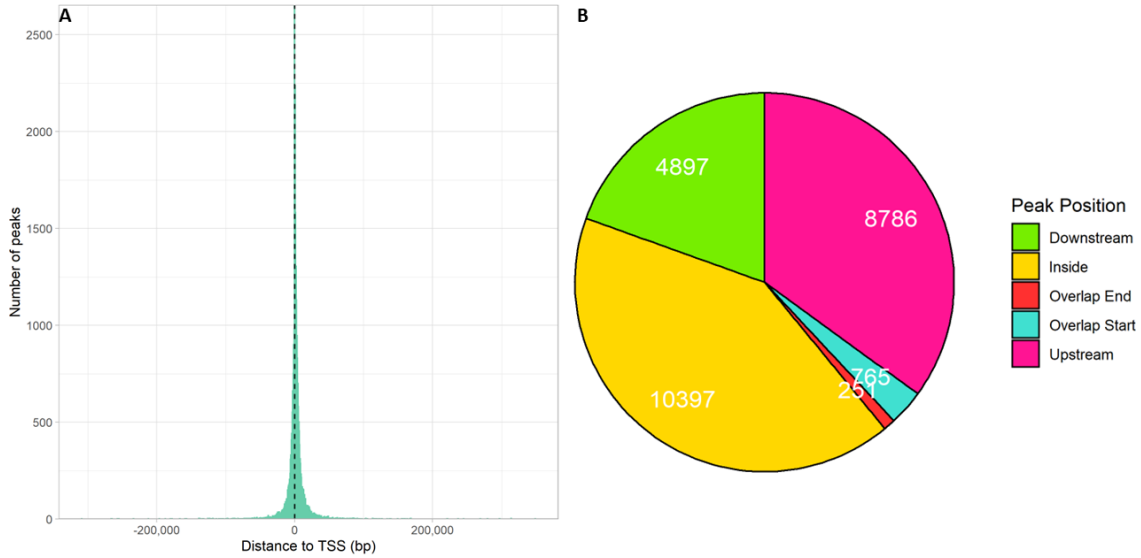


Figure 49.3. WRKY15 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

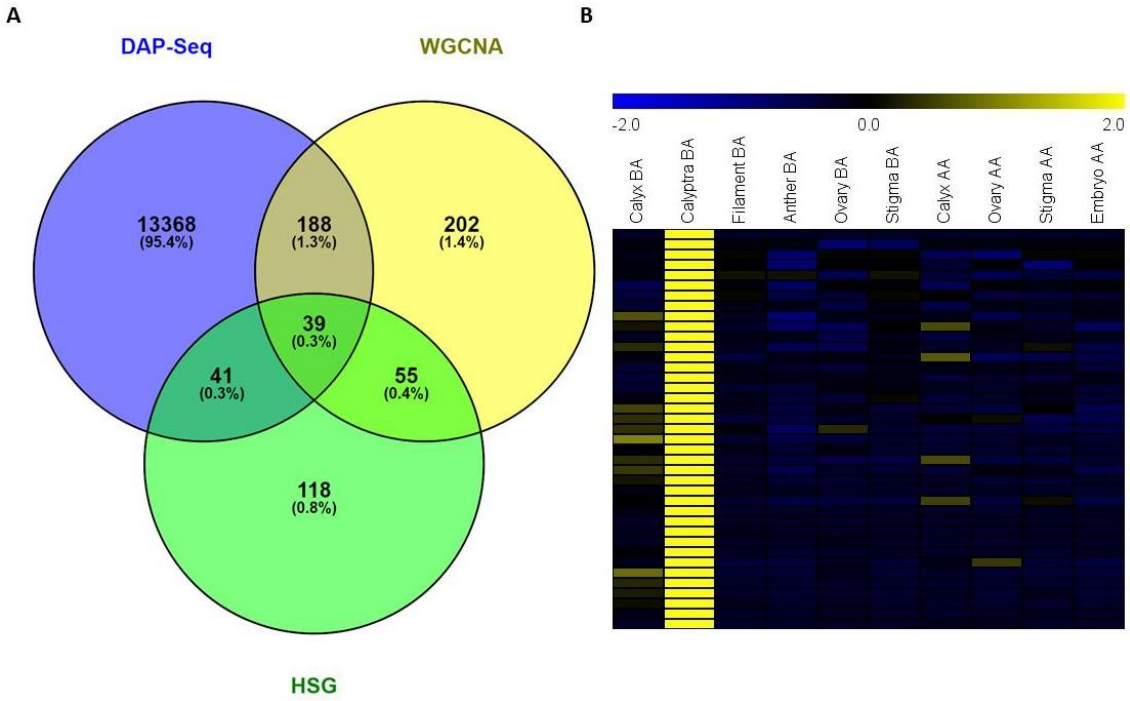


Figure 50.3. Individuation of WRKY15 High Confidence Targets. (A) Shared genes between DAP-Seq and Calyptra_BA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

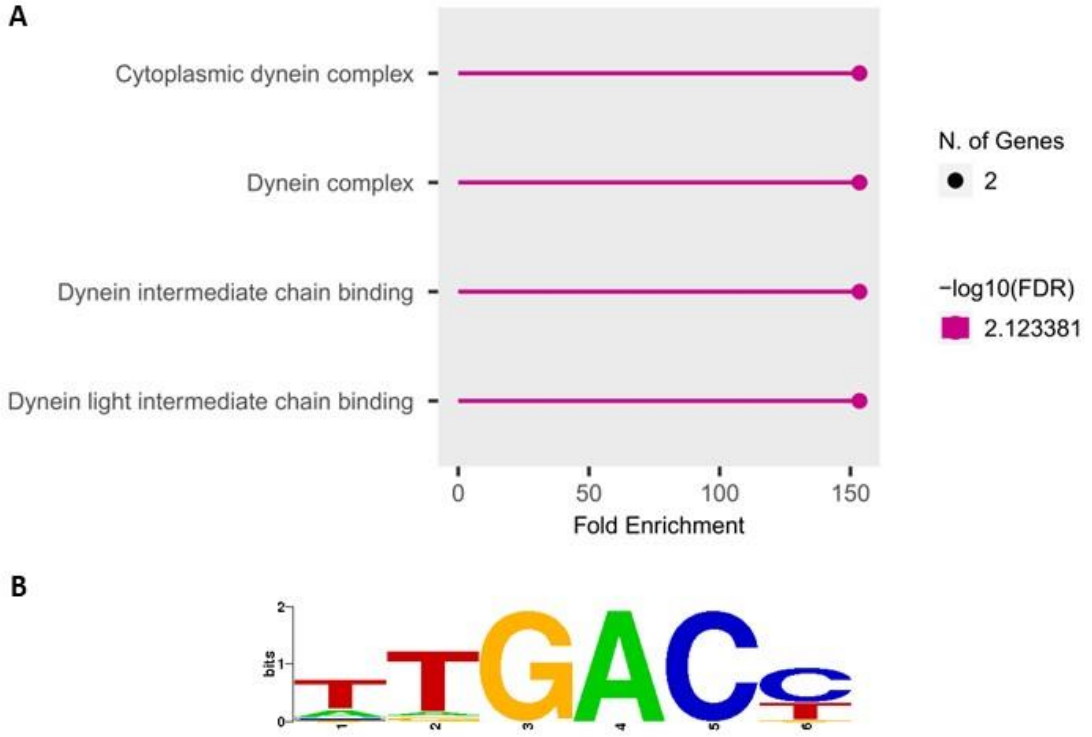


Figure 51.3. (A) GO enrichment analysis performed on WRKY15 HCTs: all available gene sets. (B) *De novo* motif discovery analysis on the whole cistrome of WRKY15 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence

conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.3.17 WRKY42

A total amount of 51 independent peaks were called following the DAP-Seq analysis of WRKY42, with 35% of them belonging to the upstream region of the gene assigned, 43% rests within the gene, 22% stands downstream the related gene and no binding events were individuated neither in the TSS nor at the gene end (Fig. 51.3B). Since the scarcity in DAP-Seq detected peaks did not allowed to identify a subset of HCTs, the attention was focused on the 3 genes obtained with the single crossing of DAP-Seq results with Embryo_AA clusters, defining 1 gene through the intersection with HSG and 2 with that of midnightblue module of WGCNA (Fig. 52.3A). VIT_07s0005g01430 encodes for the DNA replication licensing factor MCM7 that acts as component of the mcm2-7 complex (mcm complex), the putative replicative helicase essential for 'once per cell cycle' DNA replication initiation and elongation in eukaryotic cells (UniProt). In grapevine variety Touriga Nacional, it was seen to be highly expressed from the floral bud swelling up to the moment in which inflorescence is well developed and the single flowers are compact in group (Ramos et al., 2014; Coombe, 1995). Moreover, the same trend was observed in Corvina, where indeed the expression improves from the bud swelling to the young flower featured inflorescence stage – as well as in Pinot noir (Rossmann et al., 2020), then stop in flowering moment and resumes it course in embryo and pericarp during the fruit-set (Fasoli et al., 2012). This last, seems to find support from what reported by Bian et al. (2022), study in which seems to be involved in molecular mechanisms of cytokinin interaction with other hormones in the early berry development, being induced from N-(2-chloro-4-pyridyl)-N'-phenylurea CPPU treatment and more softly also by hydrogen peroxide treatment (Guo et al., 2019). The stop in flowering moment is interestingly coherent with what observed in Pinot Meunier, where the expression was very intense in immature inflorescences positioned at first node under the meristematic apex and progressively decreases moving to the inflorescences of the inferior nodes (Bio Project PRJNA 149111), thus defining a gradual shutdown of the expression of this gene as the actual flowering approaches. This relieved expression pattern in grapevine is consistent with the studied function of its ortholog At4g02060 which is involved in DNA replication initiation. The related transcript is abundant in proliferating and endocycling tissues, indeed it was specifically described featuring shoot apex and floral bud tissues (Schultz et al., 2004). Moreover, it is inferable - basing on sequence similarity – a requirement of this protein for megagametophyte and embryo development (UniProt) and its specific expression in

Arabidopsis carpel during anthesis and in ovules after that moment (Klepikova et al., 2016) could confirm this assertion. Very interestingly, VIT_07s0005g01430 was reported to be expressed in E-L 27 (young berries enlarging; Coombe, 1995) both in Sangiovese grapevine variety and in its spontaneous seedless somatic variant (Nwafor et al., 2014), opening the possibility that this gene is not dependent on the presence of the seed as sink for the storage in berry of nutrients and metabolites during ripening process. Finally, the expression of VIT_17s0053g00360 – which encodes for a DYW-deaminase domain-containing protein – finds match with the pattern of Corvina, in which the transcript levels are high in developing seed from fruit set to mid-ripening stages and in berry pulp and skin, featuring these tissues from the fruit-set up to the very last phase of post-harvest withering (Fasoli et al., 2012).

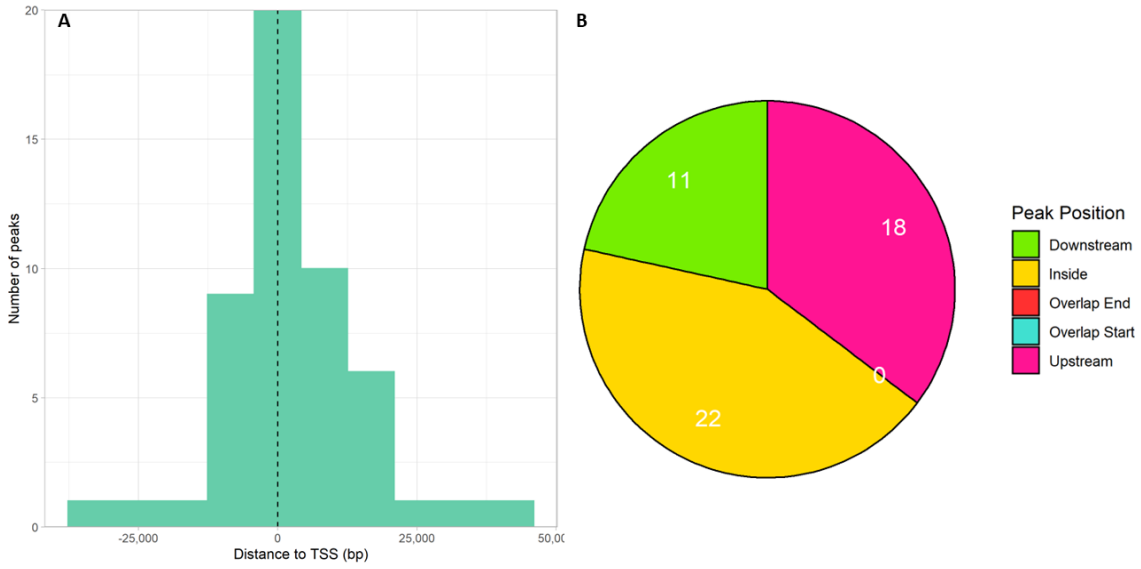


Figure 52.3. WRKY42 DAP-Seq derived cistrome landscapes in *V. vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites (TSS) of assigned genes. (B) The proportions of binding peaks represented within the pie-charts.

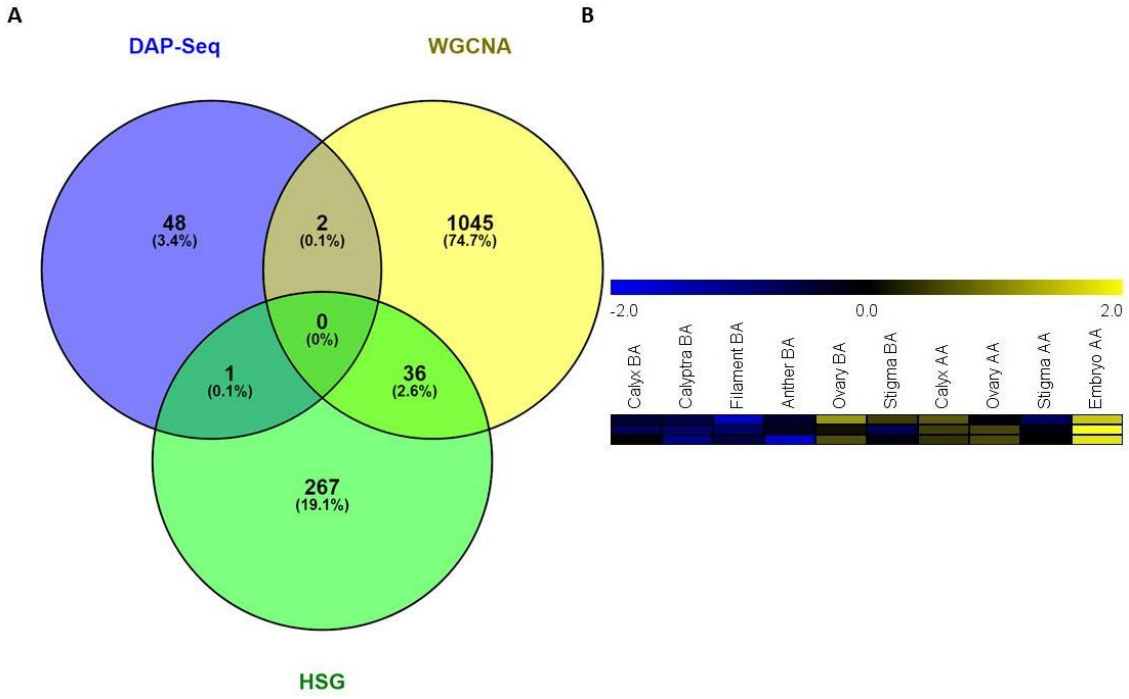


Figure 53.3. Individuation of WRKY42 High Confidence Targets. (A) Shared genes between DAP-Seq and Embryo_AA cluster of WGCNA and HSG analysis. (B) HCTs expression across all the samples of floral expression atlas defined in Chapter II.

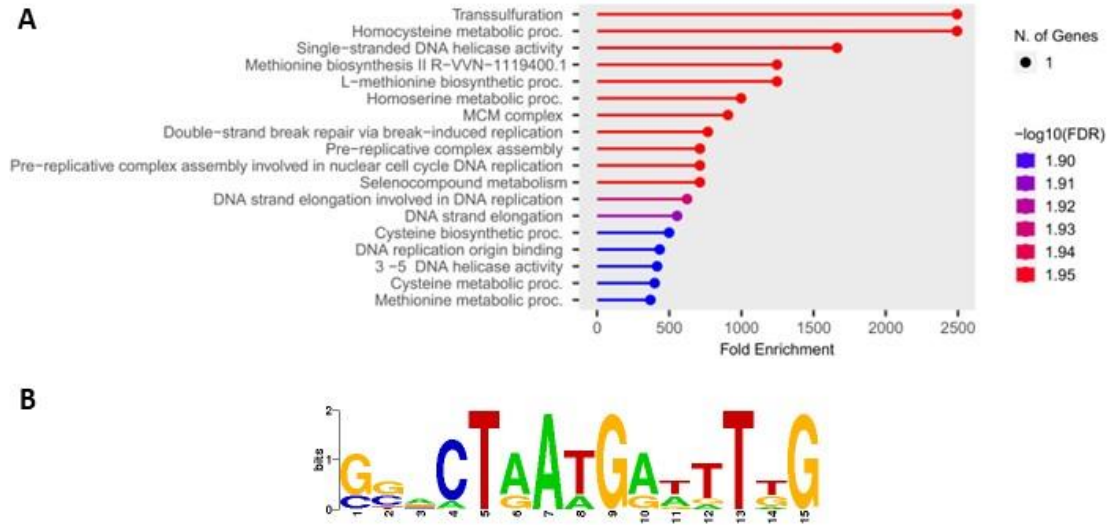


Figure 54.3. (A) GO enrichment analysis performed on WRKY42HCTs: all available gene sets. (B) *De novo* motif discovery analysis on the whole cistrome of WRKY42 discovered with DAP-Seq. The sequence represents the top-ranking TF binding motif identified based on the detection of overrepresented oligonucleotides. The overall height of each letter stack indicates the sequence conservation at that position, and the height of symbols within the stack reflects the relative frequency of the corresponding nucleic acid at that position.

3.4 Conclusions

In the present study, we took advantage of this approach to shed light on the role of 17 whorl specific TFs selected based on a grapevine floral expression atlas reported in Chapter II. Based on the proximity to the peaks we identified candidate gene targets for each of the selected TFs, and, to give strength to our investigation and to results, we crossed the lists of putative targets with results obtained from the WGCNA and tau analyses reported in Chapter II. The ultimate aim was to define a subset of target genes characterized by a similar expression profile, highly specificity for a determined whorl and, putatively regulated by the same TF. We defined these genes as High Confidence Targets (HCTs). Altogether, results indicated that this approach is effective and can be extremely useful to clarify the genetic relationships characterizing the various organs, indeed the biological meaning encountered analysing one by one the HCTs within the different networks is unquestionably relevant. The reading key for the HCTs regulative webs is that the diverse function defined by the targets are mutually interconnected – being in the network of the same TF - and at the same time are highly related to the peculiar mechanisms of a whorl. At this regard, it is pivotal to evidence the case of the TFs related to the stigma tissue, namely MYB143, MYB145, MYB148, MYB150, MYB154, MYBA7 and MYBPA7. These regulators share a conspicuous number of targets. This point suggests a remarkable reliability of the proposed method to investigate the tissue-specific regulome which is evidenced by the conservation of some HCTs among TFs specific for the same organ. Finally, the regulative network of each TF was enriched with a *de novo* motif discovery analysis, to individuate the most representative sequences within the peaks population for a given TF that can be considered a sort of molecular signature for a particular TF. Emblematic is the case of *MYB108A*, which was described strictly related to the anther (Chapter II). The regulative HCTs network of this TF highlighted a considerable number of genes showing a specific role in pollen and anther development at different stages of growth, suggesting its regulative control in various molecular mechanisms, like PCD of tapetum cells, the formation of exine in pollen wall, the correct germination of the grain, the synthesis of the pollen tube components, its development and its efficient direction of the growth to the ovule (Meng et al., 2020; Zhang et al., 2014; Xia et al., 2014; Papuga et al., 2010; Dong et al., 2005; Golovkin and Reddy, 2003). Moreover, the most representative binding motif of its cistrome revealed a previously characterized sequence, which showed specificity for processes characterizing pollen and anther development (Filichkin et al., 2004; Bate and Twell, 1998; Eyal et al., 1995). Wrapping up, we deem that the workflow used in this analysis could be useful not only to draw TF-centred regulative maps, but also to combine each of these to achieve a unique and integrate regulative network describing the molecular relationships occurring among different TFs that characterize a given tissue as in our case, but

also describing other situations, like phenological phases, different organs and specific treatments.

ID	Annotation	Tissue	TF	DAP-Seq p-value	WGCNA module	tau	Tissue Expression (TPM)
VIT_02s0025g03170	ERF052	Ovary_AA	MYB91A	6,01	NA	0,889	2,28
VIT_09s0002g03820	RPS5	Ovary_AA	MYB91A	4,04	NA	0,889	1,2
VIT_14s0068g00900	NA	Ovary_AA	MYB91A	3,68	NA	0,889	2,58
VIT_13s0084g00630	NA	Ovary_AA	MYB91A	4,25	lightcyan	NA	26,6
VIT_05s0029g01080	Ribosome biogenesis protein Bms1	Ovary_AA	MYB192	5,18	lightcyan	NA	18,18
VIT_13s0067g00990	Heat shock protein-related	Ovary_AA	MYB192	4,83	lightcyan	NA	24,15
VIT_13s0084g00630	NA	Ovary_AA	MYB192	4,61	lightcyan	NA	26,6
VIT_04s0008g05430	RDR6 (RNA-dependent RNA polymerase 6)	Ovary_AA	MYB192	4,24	lightcyan	NA	18,39
VIT_13s0019g00310	ABC transporter B member 20	Ovary_AA	MYB192	4,02	lightcyan	NA	21,24
VIT_11s0016g00100	Adapter protein SPIKE1 (SPK1)	Ovary_AA	MYB192	3,87	lightcyan	NA	32,47
VIT_15s0107g00090	Gag-pol polyprotein	Ovary_AA	MYB192	4,02	NA	1	1,5
VIT_05s0020g01350	Nuclear transcription factor Y subunit B-5	Ovary_AA	MYB192	3,57	NA	1	1,24
VIT_16s0013g02160	NA	Sigma_AA	MYBA7	15,68	brown	NA	95,71
VIT_06s0004g01480	Lipoxygenase LOX1	Sigma_AA	MYBA7	4,29	brown	NA	33,77
VIT_07s0151g00260	NA	Sigma_AA	MYBA7	3,83	brown	NA	65,33
VIT_13s0019g01710	Scarecrow transcription factor 14 (SCL14)	Sigma_AA	MYBA7	3,79	brown	NA	3,51
VIT_11s0016g01060	Avr9/Cf-9 induced kinase 1	Sigma_AA	MYBA7	4,53	NA	0,905	24,61
VIT_06s0004g03840	Chitinase 2	Sigma_BA	MYBA8	5	sienna3	NA	121,42
VIT_04s0008g02020	Cellulase	Sigma_AA	MYBPA7	5,32	NA	1	1,57
VIT_00s0203g00010	NA	Sigma_AA	MYBPA7	5,12	NA	1	1,23
VIT_18s0001g00080	LAC33	Sigma_AA	MYBPA7	3,91	NA	1	1,41
VIT_19s0027g01630	Disease resistance protein (NBS-LRR class)	Sigma_AA	MYBPA7	5,72	brown	0,852	5,33
VIT_13s0019g00390	F-box family protein	Sigma_AA	MYBPA7	4,33	brown	0,963	5,73
VIT_13s0064g00960	MYB46	Filament_BA	MYBPA9	5	NA	0,889	1,93
VIT_10s0116g01670	AD3	Embryo_AA	MYBPAL3	4	midnightblue	NA	19,33

VIT_02s0025g04520	Calcium-dependent protein kinase 1 CDPK protein kinase	Embryo_AA	MYBPAL3	4	midnightblue	NA	13,41
VIT_00s1764g00020	Leucine-rich repeat transmembrane protein kinase	Embryo_AA	MYBPAL3	4	midnightblue	NA	24,26
VIT_08s0007g05410	Cystathionine beta-lyase	Embryo_AA	WRKY42	5	midnightblue	NA	20,17
VIT_07s0005g01430	PROLIFERA protein	Embryo_AA	WRKY42	4	midnightblue	NA	24,29
VIT_17s0053g00360	Pentatricopeptide (PPR) repeat	Embryo_AA	WRKY42	4	NA	0,889	1,15
VIT_15s0048g01940	Generic methyltransferase	Anther_BA	MYB108A	5,34	blue	0,873	23,05
VIT_04s0023g01390	Cupin family protein	Anther_BA	MYB108A	5,19	blue	0,889	841,51
VIT_08s0056g01350		Anther_BA	MYB108A	4,36	blue	1	20,29
VIT_12s0028g00210	Prenylcysteine alpha-carboxyl methyltransferase	Anther_BA	MYB108A	4,25	blue	0,938	58,64
VIT_08s0007g04650	Thaumatin	Anther_BA	MYB108A	4,25	blue	1	13,22
VIT_18s0001g06780	Protein kinase CRK1	Anther_BA	MYB108A	4,06	blue	0,933	9,57
VIT_09s0002g06570	Rhomboid ATRBL1	Anther_BA	MYB108A	4,04	blue	0,963	88,35
VIT_09s0002g06790	Small heat shock protein	Anther_BA	MYB108A	4,02	blue	0,901	67,9
VIT_09s0002g02260	H(+)-ATPase 9 AHA9	Anther_BA	MYB108A	3,83	blue	1	15,45
VIT_17s0000g04290	Phosphoprotein NtEPc	Anther_BA	MYB108A	3,81	blue	0,956	449,06
VIT_18s0001g06450	NA	Stigma_AA	MYB143	9,1	brown	0,963	7,18
VIT_17s0000g03660	NAC06	Stigma_AA	MYB143	7,14	brown	0,968	26,44
VIT_06s0004g05240	ETR2	Stigma_AA	MYB143	6,66	brown	1	3,73
VIT_15s0048g01080	NA	Stigma_AA	MYB143	5,15	brown	0,889	35,76
VIT_08s0032g00710	CXE carboxylesterase	Stigma_AA	MYB143	4,98	brown	1	4,64
VIT_12s0059g00370	EF hand	Stigma_AA	MYB143	4,94	brown	0,889	3,23
VIT_01s0137g00680	NA	Stigma_AA	MYB143	4,48	brown	0,852	6,33
VIT_08s0040g00960	CXE carboxylesterase	Stigma_AA	MYB143	4,11	brown	0,852	6,99
VIT_06s0009g03380	ABC transporter G member 14	Stigma_AA	MYB143	4,01	brown	0,963	5,72
VIT_01s0146g00320	NA	Stigma_AA	MYB143	4,01	brown	0,878	141,37
VIT_19s0027g01630	Disease resistance protein (NBS-LRR class)	Stigma_AA	MYB145	8,34	brown	0,852	5,33

VIT_04s0008g02070	Avr9/Cf-9 induced kinase 1	Stigma_AA	MYB145	4,02	brown	0,926	17,16
VIT_18s0075g01030	Laccase	Stigma_AA	MYB145	3,81	brown	1	7,27
VIT_14s0066g02470	NA	Stigma_AA	MYB145	3,62	brown	0,889	5,01
VIT_01s0137g00680	NA	Stigma_AA	MYB145	3,51	brown	0,852	6,33
VIT_15s0021g01390	Uclacyanin 3	Stigma_AA	MYB145	3,39	brown	0,856	139,8
VIT_08s0007g04680	EXPA9	Stigma_AA	MYB145	3,31	brown	0,944	48,67
VIT_01s0026g00010	NA	Stigma_AA	MYB145	2,95	brown	0,926	18,1
VIT_18s0001g14810	Lipase 3 (EXL3) family II extracellular	Stigma_AA	MYB145	2,78	brown	0,917	8,4
VIT_17s0000g00200	ERF114	Stigma_AA	MYB145	2,77	brown	0,873	30,85
VIT_05s0020g03280	Copper amine oxidase	Stigma_AA	MYB148	4,95	brown	0,968	29,41
VIT_08s0007g04680	EXPA9	Stigma_AA	MYB148	3,87	brown	0,944	48,67
VIT_08s0032g00710	CXE carboxylesterase	Stigma_AA	MYB148	3,69	brown	1	4,64
VIT_17s0000g04360	WAK2	Stigma_AA	MYB148	3,52	brown	1	6,71
VIT_08s0058g00970	Cationic peroxidase	Stigma_AA	MYB148	3,33	brown	0,861	51,2
VIT_16s0013g00590	Plastocyanin domain-containing protein	Stigma_AA	MYB148	3,18	brown	0,864	86,01
VIT_15s0045g01130	DNA-binding protein	Stigma_AA	MYB148	2,98	brown	1	4,39
VIT_12s0055g00250	UDP-glucose glucosyltransferase	Stigma_AA	MYB148	2,98	brown	0,889	7,52
VIT_12s0057g00320	Exopolygalacturonase	Stigma_AA	MYB148	2,92	brown	0,972	9,51
VIT_01s0146g00320	NA	Stigma_AA	MYB148	2,77	brown	0,878	141,37
VIT_16s0013g00590	Plastocyanin domain-containing protein	Stigma_AA	MYB150	7,79	brown	0,864	86,01
VIT_08s0007g04680	EXPA9	Stigma_AA	MYB150	7,43	brown	0,944	48,67
VIT_15s0048g01080	NA	Stigma_AA	MYB150	6,01	brown	0,889	35,76
VIT_18s0166g00090	VQ motif-containing protein	Stigma_AA	MYB150	5,91	brown	0,944	4,53
VIT_10s0003g05550	Carbohydrate oxidase	Stigma_AA	MYB150	5,09	brown	0,986	37,28
VIT_12s0055g00250	UDP-glucose glucosyltransferase	Stigma_AA	MYB150	4,68	brown	0,889	7,57
VIT_04s0008g06210	Nodulin	Stigma_AA	MYB150	4,46	brown	0,944	11,48
VIT_05s0020g03270	S-receptor kinase	Stigma_AA	MYB150	4,15	brown	0,852	17,42
VIT_18s0001g14420	NA	Stigma_AA	MYB150	4,11	brown	1	4,03

VIT_19s0014g04240	S-locus protein kinase	Stigma_AA	MYB150	2,93	brown	1	3,86
VIT_19s0027g01630	Disease resistance protein (NBS-LRR class)	Stigma_AA	MYB154	10,84	brown	0,852	5,33
VIT_06s0004g05240	ETR2	Stigma_AA	MYB154	6,22	brown	1	3,37
VIT_07s0151g00700	EDA4	Stigma_AA	MYB154	4,53	brown	0,852	78,4
VIT_03s0088g01240	bHLH	Stigma_AA	MYB154	4,46	brown	0,963	6,2
VIT_05s0051g00110	GSTU8	Stigma_AA	MYB154	4,35	brown	0,963	4,89
VIT_14s0081g00550	Metal ion binding	Stigma_AA	MYB154	4,04	brown	1	5,1
VIT_02s0025g01060	Phytocyanin	Stigma_AA	MYB154	3,75	brown	0,968	27,14
VIT_01s0010g01090	Peroxidase ATP23a	Stigma_AA	MYB154	3,63	brown	0,903	51,18
VIT_11s0052g01110	4-coumarate-CoA ligase 1	Stigma_AA	MYB154	3,57	brown	1	18,78
VIT_01s0026g00010	NA	Stigma_AA	MYB154	3,44	brown	0,926	18,01
VIT_08s0007g04300	Lipase GDSL	Ovary_BA	NAC62	2,89	plum1	1	14,85
VIT_08s0040g03120	APRR7	Ovary_BA	NAC62	2,48	plum1	0,856	125,74
VIT_18s0164g00050	NA	Ovary_BA	NAC62	2,44	plum1	0,864	54,9
VIT_10s0042g00940	Hydroxyproline-rich glycoprotein family protein	Ovary_BA	NAC62	2,42	plum1	0,978	13,18
VIT_06s0004g00150	MYR1	Ovary_BA	NAC62	2,23	plum1	0,861	39,97
VIT_06s0009g00320	NA	Ovary_BA	NAC62	2,07	plum1	0,889	15,21
VIT_12s0028g03640	Ripening induced protein	Ovary_BA	NAC62	1,87	plum1	0,861	11,23
VIT_11s0016g01250	Growth-regulating factor 6	Ovary_BA	NAC62	1,58	plum1	0,963	6,86
VIT_17s0053g00070	Metal-nicotianamine transporter YSL5	Ovary_BA	NAC62	1,5	plum1	0,944	9,58
VIT_04s0023g01590	NA	Ovary_BA	NAC62	1,45	plum1	1	4,08
VIT_18s0001g08840	Serine carboxypeptidase S10	Calyptra_BA	WRKY15	199	cyan	0,984	27,92
VIT_04s0023g00560	Auxin responsive SAUR protein	Calyptra_BA	WRKY15	74	cyan	0,889	10,85
VIT_18s0122g01480	CYP87A2	Calyptra_BA	WRKY15	48	cyan	1	8,35
VIT_15s0048g01530	Geraniol 10-hydroxylase	Calyptra_BA	WRKY15	35	cyan	0,889	5,2
VIT_08s0007g01270	Dynein light chain LC6, flagellar outer arm	Calyptra_BA	WRKY15	19	cyan	0,968	30,56
VIT_09s0002g03730	NA	Calyptra_BA	WRKY15	18	cyan	0,944	10,3

VIT_17s0000g08760	Receptor kinase TRKa	Calyptra_BA	WRKY15	17	cyan	0,861	8,69
VIT_06s0004g00230	WRKY48	Calyptra_BA	WRKY15	16	cyan	0,944	3,76
VIT_06s0061g00120	Beta-1,3-glucanase [Vitis riparia]	Calyptra_BA	WRKY15	13	cyan	0,889	12,7
VIT_08s0007g04580	UGT73C2	Calyptra_BA	WRKY15	12	cyan	0,956	15,39
VIT_06s0004g01450	LOX2	Calyptra_BA	WRKY22	17	cyan	0,864	66,23
VIT_18s0001g08840	Serine carboxypeptidase S10	Calyptra_BA	WRKY22	6	cyan	0,984	27,92
VIT_08s0040g00790	FK506-binding protein 4/5	Calyptra_BA	WRKY22	5	cyan	0,944	8,1
VIT_18s0001g01630	Calcium-binding protein CML	Calyptra_BA	WRKY22	5	cyan	0,914	98,22
VIT_04s0008g00130	Avr9/Cf-9 rapidly elicited protein 146	Calyptra_BA	WRKY22	4	cyan	1	8,24
VIT_14s0068g02190	Chloride channel protein B	Calyptra_BA	WRKY22	4	cyan	0,861	7,65
VIT_18s0122g01480	CYP87A2	Calyptra_BA	WRKY22	4	cyan	1	8,35
VIT_19s0015g01720	fructose-bisphosphate aldolase, cytoplasmic isozyme 1	Calyptra_BA	WRKY22	3	cyan	1	9,35
VIT_12s0134g00390	S-locus lectin protein kinase family	Calyptra_BA	WRKY22	3	cyan	0,889	3,86
VIT_13s0019g03530	NA	Calyptra_BA	WRKY22	3	cyan	0,963	18,96

Table S1.3. Top 10 HCTs for each TF, ranked basing on DAP-Seq peak p value

BIBLIOGRAPHIC REFERENCES

Aguilar-Martínez J. A., Poza-Carrión C., Cubas P. Arabidopsis BRANCHED1 acts as an integrator of branching signals within axillary buds. Plant Cell. 2007 Feb;19(2):458-72. doi: 10.1105/tpc.106.048934. Epub 2007 Feb 16. PMID: 17307924; PMCID: PMC1867329.

Arenas-Huertero F., Arroyo A., Zhou L., Sheen J., León P. Analysis of Arabidopsis glucose insensitive mutants, gin5 and gin6, reveals a central role of the plant hormone ABA in the regulation of plant vegetative development by sugar. Genes Dev. 2000 Aug 15;14(16):2085-96. PMID: 10950871; PMCID: PMC316855.

Aviezer-Hagai K., Skovorodnikova J., Galigniana M., Farchi-Pisanty O., Maayan E., Bocovza S., Efrat Y., von Koskull-Döring P., Ohad N., Breiman A. Arabidopsis immunophilins ROF1 (AtFKBP62) and ROF2 (AtFKBP65) exhibit tissue specificity, are heat-stress induced, and bind HSP90. Plant Mol Biol. 2007 Jan;63(2):237-55. doi: 10.1007/s11103-006-9085-z. Epub 2006 Nov 2. PMID: 17080288.

Ayadi M., Delaporte V., Li Y. F., Zhou D. X. Analysis of GT-3a identifies a distinct subgroup of trihelix DNA-binding transcription factors in Arabidopsis. FEBS Lett. 2004 Mar 26;562(1-3):147-54. doi: 10.1016/S0014-5793(04)00222-4. PMID: 15044016.

Balic I., Vizoso P., Nilo-Poyanco R., Sanhueza D., Olmedo P., Sepúlveda P., ... & Campos-Vargas R. (2018). Transcriptome analysis during ripening of table grape berry cv. Thompson Seedless. *PloS one*, *13*(1), e0190087.

Bartlett A., O'Malley R., Huang Ss. *et al.* Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat Protoc* **12**, 1659–1672 (2017). https://doi.org/10.1038/nprot.2017.055

Bate N., Twell D. Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. Plant Mol Biol. 1998 Jul;37(5):859-69. doi: 10.1023/a:1006095023050. PMID: 9678581.

Bian L., Shi B.X., Yu K.K. *et al.* Genome-Wide Characterization of Cytokinin Response Regulator in Grape and Expression Analyses during Berry Set Process. *Russ J Plant Physiol* **69**, 46 (2022). https://doi.org/10.1134/S1021443722030049

Blanco F., Garretón V., Frey N., Dominguez C., Pérez-Acle T., Van der Straeten D., Jordana X., Holuigue L. Identification of NPR1-dependent and independent genes early induced by salicylic acid treatment in Arabidopsis. Plant Mol Biol. 2005 Dec;59(6):927-44. doi: 10.1007/s11103-005-2227-x. PMID: 16307367.

Blanco-Ulate B., Hopfer H., Figueroa-Balderas R., Ye Z., Rivero R. M., Albacete A., Pérez-Alfocea F., Koyama R., Anderson M. M., Smith R. J., Ebeler S. E., Cantu D. Red blotch disease alters grape berry development and metabolism by interfering with the transcriptional and hormonal regulation of ripening, *Journal of Experimental Botany*, Volume 68, Issue 5, 1 February 2017, Pages 1225–1238, https://doi.org/10.1093/jxb/erw506

Boss, P., Thomas, M. Association of dwarfism and floral induction with a grape 'green revolution' mutation. *Nature* 416, 847–850 (2002). https://doi.org/10.1038/416847a

Bracha-Drori K., Shichrur K., Cohen Lubetzky T., Yalovsky S., Functional Analysis of Arabidopsis Post Prenylation CaaX Processing Enzymes and Their Function in Subcellular Protein Targeting , *Plant Physiology*, Volume 148, Issue 1, September 2008, Pages 119–131, https://doi.org/10.1104/pp.108.120477

Canaguier A., Grimplet J., Di Gaspero G., Scalabrin S., Duchêne E., Choisne N., ... & Adam-Blondon A. F. (2017). A new version of the grapevine reference genome assembly (12X. v2) and of its annotation (VCost. v3). *Genomics data*, *14*, 56.

Casson, S., Spencer, M., Walker, K., & Lindsey, K. (2005). Laser capture microdissection for the analysis of gene expression during embryogenesis of Arabidopsis. The Plant Journal, 42(1), 111-123.

Cerise M., Giaume F., Galli M., Khahani B., Lucas J., Podico F., ... & Fornara F. (2021). OsFD4 promotes the rice floral transition via florigen activation complex formation in the shoot apical meristem. New Phytologist, 229(1), 429-443.

Chen, F., Fasoli, M., Tornielli, G. B., Dal Santo, S., Pezzotti, M., Zhang, L., ... & Cheng, Z. M. (2013). The evolutionary history and diverse physiological roles of the grapevine calcium-dependent protein kinase gene family. PLoS One, 8(12), e80818.

Cheng Y., Dai X., Zhao Y. Auxin synthesized by the YUCCA flavin monooxygenases is essential for embryogenesis and leaf formation in Arabidopsis. Plant Cell. 2007 Aug;19(8):2430-9. doi: 10.1105/tpc.107.053009. Epub 2007 Aug 17. PMID: 17704214; PMCID: PMC2002601.

Chen W., Provart N. J., Glazebrook J., Katagiri F., Chang H. S., Eulgem T., Mauch F., Luan S., Zou G., Whitham S. A., Budworth P. R., Tao Y., Xie Z., Chen X., Lam S., Kreps J. A., Harper J. F., Si-Ammour A., Mauch-Mani B., Heinlein M., Kobayashi K., Hohn T., Dangl J. L., Wang X., Zhu T. Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell. 2002 Mar;14(3):559-74. doi: 10.1105/tpc.010410. PMID: 11910004; PMCID: PMC150579.

Chin C. S., Peluso P., Sedlazeck F. J., Nattestad M., Concepcion G. T., Clum A., ... & Schatz M. C. (2016). Phased diploid genome assembly with single-molecule real-time sequencing. *Nature methods*, *13*(12), 1050-1054.

Compagnon V., Diehl .P, Benveniste I., Meyer D., Schaller H., Schreiber L., Franke R., Pinot F. CYP86B1 is required for very long chain omega-hydroxyacid and alpha, omega -dicarboxylic acid synthesis in root and seed suberin polyester. Plant Physiol. 2009 Aug;150(4):1831-43. doi: 10.1104/pp.109.141408. Epub 2009 Jun 12. PMID: 19525321; PMCID: PMC2719127.

Da Silva C., Zamperin G., Ferrarini A., Minio A., Dal Molin A., Venturini L., Buson G., Tononi P., Avanzato C., Zago E., Boido E., Dellacassa E., Gaggero C., Pezzotti M., Carrau F., Delledonne M. The High Polyphenol Content of Grapevine Cultivar Tannat Berries Is Conferred Primarily by Genes That Are Not Shared with the Reference Genome, *The Plant Cell*, Volume 25, Issue 12, December 2013, Pages 4777–4788, https://doi.org/10.1105/tpc.113.118810

Dal Santo S., Tornielli G. B., Zenoni S., Fasoli M., Farina L., Anesi A., ... & Pezzotti M. (2013). The plasticity of the grapevine berry transcriptome. *Genome biology*, *14*(6), 1-18.

Dagert M., & Ehrlich S. D. (1979). Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene*, *6*(1), 23-28.

Dong X., Hong Z., Sivaramakrishnan M., Mahfouz M., Verma DP. Callose synthase (CalS5) is required for exine formation during microgametogenesis and for pollen viability in Arabidopsis. Plant J. 2005 May;42(3):315-28. doi: 10.1111/j.1365-313X.2005.02379.x. PMID: 15842618.

du Plessis K., Young P.R., Eyéghé-Bickong H.A., Vivier M.A. The Transcriptional Responses and Metabolic Consequences of Acclimation to Elevated Light Exposure in Grapevine Berries. Front Plant Sci. 2017 Jul 20;8:1261. doi: 10.3389/fpls.2017.01261. PMID: 28775728; PMCID: PMC5518647.

Eyal Y., Curie C., McCormick S. Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes., The Plant Cell, Volume 7, Issue 3, March 1995, Pages 373–384, https://doi.org/10.1105/tpc.7.3.373

Eulgem T., Rushton P. J., Robatzek S., Somssich I. E. The WRKY superfamily of plant transcription factors. Trends Plant Sci. 2000 May;5(5):199-206. doi: 10.1016/s1360-1385(00)01600-9. PMID: 10785665.

Fasoli M., Dal Santo S., Zenoni S., Tornielli G. B., Farina L., Zamboni A., Porceddu A., Venturini L., Bicego M., Murino V., Ferrarini A., Delledonne M., Pezzotti M. The Grapevine Expression Atlas Reveals a Deep Transcriptome Shift Driving the Entire Plant into a Maturation Program , *The Plant Cell*, Volume 24, Issue 9, September 2012, Pages 3489–3505, https://doi.org/10.1105/tpc.112.100230

Fasoli M., Richter C.L., Zenoni S., Bertini E., Vitulo N., Dal Santo S., Dokoozlian N., Pezzotti M., Tornielli G.B. Timing and Order of the Molecular Events Marking the Onset of Berry Ripening in Grapevine. Plant Physiol. 2018 Nov;178(3):1187-1206. doi: 10.1104/pp.18.00559. Epub 2018 Sep 17. PMID: 30224433; PMCID: PMC6236592.

Filichkin S.A., Leonard J.M., Monteros A., Liu P.P., Nonogaki H. A novel endo-beta-mannanase gene in tomato LeMAN5 is associated with anther and pollen development. Plant Physiol. 2004 Mar;134(3):1080-7. doi: 10.1104/pp.103.035998. Epub 2004 Feb 19. PMID: 14976239; PMCID: PMC389932.

Finkelstein R. R., Wang M. L., Lynch T. J., Rao S., & Goodman H. M. (1998). The Arabidopsis Abscisic Acid Response Locus ABI4 Encodes an APETALA2 Domain Protein. *The Plant Cell*, *10*(6), 1043–1054. https://doi.org/10.2307/3870689

Foxe J. P., Slotte T., Stahl E. A., Neuffer B., Hurka H., & Wright S. I. (2009). Recent speciation associated with the evolution of selfing in Capsella. *Proceedings of the National Academy of Sciences*, *106*(13), 5241-5245.

Galli M., Khakhar A., Lu Z. et al. The DNA binding landscape of the maize AUXIN RESPONSE FACTOR family. Nat Commun 9, 4526 (2018). https://doi.org/10.1038/s41467-018-06977-6

Ganeteg U., Klimmek F. & Jansson S. Lhca5 – an LHC-Type Protein Associated with Photosystem I. Plant Mol Biol 54, 641–651 (2004). https://doi.org/10.1023/B:PLAN.0000040813.05224.94

Gaudet P., Livstone M. S., Lewis S. E., Thomas P. D. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium, *Briefings in Bioinformatics*, Volume 12, Issue 5, September 2011, Pages 449–462, https://doi.org/10.1093/bib/bbr042

Ge S. X., Jung D., Yao R., ShinyGO: a graphical gene-set enrichment tool for animals and plants, Bioinformatics, Volume 36, Issue 8, 15 April 2020, Pages 2628–2629, https://doi.org/10.1093/bioinformatics/btz931

Ge S.X., Jung D., Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 2020 Apr 15;36(8):2628-2629. doi: 10.1093/bioinformatics/btz931. PMID: 31882993; PMCID: PMC7178415.

Gilmartin P. M., Sarokin L., Memelink J., Chua N. H. Molecular light switches for plant genes. Plant Cell. 1990 May;2(5):369-78. doi: 10.1105/tpc.2.5.369. PMID: 2152164; PMCID: PMC159894.

Gidoni D., Brosio P., Bond-Nutter D., Bedbrook J., Dunsmuir P. Novel cis-acting elements in Petunia Cab gene promoters. Mol Gen Genet. 1989 Jan;215(2):337-44. doi: 10.1007/BF00339739. PMID: 2651885.

Gnesutta N., Kumimoto R.W., Swain S., Chiara M., Siriwardana C., Horner D.S., Holt B.F. 3rd, Mantovani R. CONSTANS Imparts DNA Sequence Specificity to the Histone Fold NF-YB/NF-YC Dimer. Plant Cell. 2017 Jun;29(6):1516-1532. doi: 10.1105/tpc.16.00864. Epub 2017 May 19. PMID: 28526714; PMCID: PMC5502446.

Goes da Silva F., Iandolino A., Al-Kayal F., Bohlmann M. C., Cushman M. A., Lim, H., ... & Cook D. R. (2005). Characterizing the grape transcriptome. Analysis of expressed sequence tags from multiple Vitis species and development of a compendium of gene expression during berry development. *Plant Physiology*, *139*(2), 574-597.

Golovkin M. and Reddy A.S. A calmodulin-binding protein from Arabidopsis has an essential role in pollen germination. Proc Natl Acad Sci U S A. 2003 Sep 2;100(18):10558-63. doi: 10.1073/pnas.1734110100. Epub 2003 Aug 19. PMID: 12928497; PMCID: PMC193600.

Grimplet J., Ibáñez S., Baroja E., Tello J., & Ibáñez J. (2019). Phenotypic, hormonal, and genomic variation among Vitis vinifera clones with different cluster compactness and reproductive performance. Frontiers in plant science, 9, 1917.

Gualberti G., Papi M., Bellucci L., Ricci I., Bouchez D., Camilleri C., Costantino P., Vittorioso P. Mutations in the Dof zinc finger genes DAG2 and DAG1 influence with opposite effects the germination of Arabidopsis seeds. Plant Cell. 2002 Jun;14(6):1253-63. doi: 10.1105/tpc.010491. PMID: 12084825; PMCID: PMC150778.

Guo D. L., Li Q., Zhao H. L., Wang Z. G., Zhang G. H., & Yu, Y. H. (2019). The variation of berry development and DNA methylation after treatment with 5-azaC on 'Kyoho' grape. Scientia Horticulturae, 246, 265-271.

Guo D. L., Wang Z. G., Li Q., Gu S. C., Zhang G. H., & Yu Y. H. (2019). Hydrogen peroxide treatment promotes early ripening of Kyoho grape. *Australian journal of grape and wine research*, *25*(3), 357-362.

Guo Y., Mahony S. & Gifford D.K. (2012) High resolution genome wide binding event finding and motif discovery reveals transcription factor spatial binding constraints. *PLoS Computational Biology*, **8**, 1–14. https://doi.org/10.1371/journal.pcbi.1002638.

He L., Meng N., Castellarin S.D., Wang Y., Sun Q., Li X.Y., Dong Z.G., Tang X.P., Duan C.Q., Pan Q.H. Combined Metabolite and Transcriptome Profiling Reveals the Norisoprenoid Responses in Grape Berries to Abscisic Acid and Synthetic Auxin. Int J Mol Sci. 2021 Jan 31;22(3):1420. doi: 10.3390/ijms22031420. PMID: 33572582; PMCID: PMC7867017.

Helm M., Schmid M., Hierl G., Terneus K., Tan L., Lottspeich F., Kieliszewski M.J., Gietl C. KDEL-tailed cysteine endopeptidases involved in programmed cell death, intercalation of new cells, and dismantling of extensin scaffolds. Am J Bot. 2008 Sep;95(9):1049-62. doi: 10.3732/ajb.2007404. PMID: 21632425.

Helwi P., Guillaumie S., Thibon C. *et al.* Vine nitrogen status and volatile thiols and their precursors from plot to transcriptome level. *BMC Plant Biol* **16**, 173 (2016). https://doi.org/10.1186/s12870-016-0836-y

Heslop-Harrison Y. and Shivanna K.R. 1977. The receptive surface of the angiosperm stigma. *Annals of Botany* **41**: 1233–1258.

Higo K., Ugawa Y., Iwamoto M. and Korenaga T. "Plant cis-acting regulatory DNA elements (PLACE) database: 1999" Nucleic Acids Research, Volume 27, Issue 1, 1999, Pages 297-300.

Hiratsuka K., Wu X., Fukuzawa H., Chua N. H. Molecular dissection of GT-1 from Arabidopsis. Plant Cell. 1994 Dec;6(12):1805-13. doi: 10.1105/tpc.6.12.1805. PMID: 7866025; PMCID: PMC160563.

Honys D. and Twell D. Transcriptome analysis of haploid male gametophyte development in Arabidopsis. Genome Biol 5, R85 (2004). https://doi.org/10.1186/gb-2004-5-11-r85

Horiguchi G., Kim G.T., Tsukaya H. The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of Arabidopsis thaliana. Plant J. 2005 Jul;43(1):68-78. doi: 10.1111/j.1365-313X.2005.02429.x. PMID: 15960617.

Hua J., Sakai H., Nourizadeh S., Chen Q.G., Bleecker A.B., Ecker J.R., Meyerowitz E.M. EIN4 and ERS2 are members of the putative ethylene receptor gene family in Arabidopsis. Plant Cell. 1998 Aug;10(8):1321-32. doi: 10.1105/tpc.10.8.1321. PMID: 9707532; PMCID: PMC144061.

Ishiguro S. and Nakamura K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. Mol Gen Genet. 1994 Sep 28;244(6):563-71. doi: 10.1007/BF00282746. PMID: 7969025.

Jaillon O., Aury J. M., Noel B., Policriti A., Clepet C., Casagrande A., ... & Wincker P. (2007). French-Italian Public Consortium for Grapevine Genome Characterization The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, *449*(7161), 463-467.

Jensen M.K., Hagedorn P.H., de Torres-Zabala M., Grant M.R., Rung J.H., Collinge D.B., Lyngkjaer M.F. Transcriptional regulation by an NAC (NAM-ATAF1,2-CUC2) transcription factor attenuates ABA signalling for efficient basal defence towards Blumeria graminis f. sp. hordei in Arabidopsis. Plant J. 2008 Dec;56(6):867-80. doi: 10.1111/j.1365-313X.2008.03646.x. Epub 2008 Oct 4. PMID: 18694460.

Joshi C. P. Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. Nucleic Acids Res. 1987 Dec 10;15(23):9627-40. doi: 10.1093/nar/15.23.9627. PMID: 3697078; PMCID: PMC306520.

Kamal N., Ochßner I., Schwandner A., Viehöver P., Hausmann L., Töpfer R., Weisshaar B., Holtgräwe D. Characterization of genes and alleles involved in the control of flowering time in grapevine. PLoS One. 2019 Jul 3;14(7):e0214703. doi: 10.1371/journal.pone.0214703. PMID: 31269026; PMCID: PMC6608932.

Kagaya Y., Ohmiya K., Hattori T. RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. Nucleic Acids Res. 1999 Jan 15;27(2):470-8. doi: 10.1093/nar/27.2.470. PMID: 9862967; PMCID: PMC148202.

Kim S.J., Kim M.R., Bedgar D.L., Moinuddin S.G., Cardenas C.L., Davin L.B., Kang C., Lewis N.G. Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in Arabidopsis. Proc Natl Acad Sci U S A. 2004 Feb 10;101(6):1455-60. doi: 10.1073/pnas.0307987100. Epub 2004 Jan 26. PMID: 14745009; PMCID: PMC341741.

Kim S.J., Kim K.W., Cho M.H., Franceschi V.R., Davin L.B., Lewis N.G. Expression of cinnamyl alcohol dehydrogenases and their putative homologues during Arabidopsis thaliana growth and development: lessons for database annotations? Phytochemistry. 2007 Jul;68(14):1957-74. doi: 10.1016/j.phytochem.2007.02.032. Epub 2007 Apr 27. PMID: 17467016.

Kim J.H., Choi D., Kende H. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in Arabidopsis. Plant J. 2003 Oct;36(1):94-104. doi: 10.1046/j.1365-313x.2003.01862.x. PMID: 12974814.

Klepikova A. V., Kasianov A. S., Gerasimov E. S., Logacheva M. D., & Penin A. A. (2016). A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *The Plant Journal*, *88*(6), 1058-1070.

Knopf R.R. and Adam Z. Rhomboid proteases in plants - still in square one? Physiol Plant. 2012 May;145(1):41-51. doi: 10.1111/j.1399-3054.2011.01532.x. Epub 2011 Nov 21. PMID: 22007993.

Koyama K., Numata M., Nakajima I., Goto-Yamamoto N., Matsumura H., Tanaka N. Functional characterization of a new grapevine MYB transcription factor and regulation of proanthocyanidin biosynthesis in grapes. J Exp Bot. 2014 Aug;65(15):4433-49. doi: 10.1093/jxb/eru213. Epub 2014 May 23. PMID: 24860184.

Kovaleva L. and Zakharova E. Hormonal status of the pollen-pistil system at the progamic phase of fertilization after compatible and incompatible pollination in *Petunia hybrida* L.. *Sex Plant Reprod* **16**, 191–196 (2003). https://doi.org/10.1007/s00497-003-0189-1

Lam E. and Chua N. H. ASF-2: a factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in Cab promoters. Plant Cell. 1989 Dec;1(12):1147-56. doi: 10.1105/tpc.1.12.1147. PMID: 2535536; PMCID: PMC159850.

Langmead B. & Salzberg S. (2012) Langmead b, salzberg sl.. fast gapped-read alignment with bowtie 2. *Nature Methods*, **9**, 357–359. https://doi.org/10.1038/nmeth.1923.

Lee S., Cho D., Kang J. et al. An ARIA-interacting AP2 domain protein is a novel component of ABA signaling. Mol Cells 27, 409–416 (2009). https://doi.org/10.1007/s10059-009-0058-3

Licausi F., Giorgi F.M., Zenoni S. et al. Genomic and transcriptomic analysis of the AP2/ERF superfamily in Vitis vinifera. BMC Genomics 11, 719 (2010). https://doi.org/10.1186/1471-2164-11-719

Loke J. C., Stahlberg E. A., Strenski D. G., Haas B. J., Wood P. C., Li Q. Q. Compilation of mRNA polyadenylation signals in Arabidopsis revealed a new signal element and potential secondary structures. Plant Physiol. 2005 Jul;138(3):1457-68. doi: 10.1104/pp.105.060541. Epub 2005 Jun 17. PMID: 15965016; PMCID: PMC1176417.

Los G. V., Encell L. P., McDougall M. G., Hartzell D. D., Karassina N., Zimprich C., ... & Wood K. V. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS chemical biology*, *3*(6), 373-382.

Lu H.C., Chen W.K., Wang Y., Bai X.J., Cheng G., Duan C.Q., Wang J., He F. Effect of the Seasonal Climatic Variations on the Accumulation of Fruit Volatiles in Four Grape Varieties Under the Double Cropping System. Front Plant Sci. 2022 Jan 27;12:809558. doi: 10.3389/fpls.2021.809558. PMID: 35154206; PMCID: PMC8829325.

Ma, Q., & Yang, J. (2019). Transcriptome profiling and identification of the functional genes involved in berry development and ripening in Vitis vinifera. Gene, 680, 84-96.

Machanick P. & Bailey T. (2011) Meme-chip: motif analysis of large dna datasets. *Bioinformatics (Oxford, England)*, **27**, 1696–1697. https://doi.org/10.1093/bioinformatics/btr189.

Magris G., Di Gaspero G., Marroni F., Zenoni S., Tornielli G.B., Celii M., De Paoli E., Pezzotti M., Conte F., Paci P., Morgante M. Genetic, epigenetic and genomic effects on variation of gene expression among grape varieties. Plant J. 2019 Sep;99(5):895-909. doi: 10.1111/tpj.14370. Epub 2019 Jun 7. PMID: 31034726.

Magris G., Jurman I., Fornasiero A. *et al.* The genomes of 204 *Vitis vinifera* accessions reveal the origin of European wine grapes. *Nat Commun* **12**, 7240 (2021). https://doi.org/10.1038/s41467-021-27487-y

Maréchal E., Hiratsuka K., Delgado J., Nairn A., Qin J., Chait B. T., Chua N. H. Modulation of GT-1 DNA-binding activity by calcium-dependent phosphorylation. Plant Mol Biol. 1999 Jun;40(3):373-86. doi: 10.1023/a:1006131330930. PMID: 10437822.

Marshall S.D., Putterill J.J., Plummer K.M., Newcomb R.D. The carboxylesterase gene family from Arabidopsis thaliana. J Mol Evol. 2003 Nov;57(5):487-500. doi: 10.1007/s00239-003-2492-8. PMID: 14738307.

Massonnet M., Cochetel N., Minio A. *et al.* The genetic basis of sex determination in grapes. *Nat Commun* **11**, 2902 (2020). https://doi.org/10.1038/s41467-020-16700-z

Mayfield J. A., Fiebig A., Johnstone S. E., Preuss D. Gene families from the Arabidopsis thaliana pollen coat proteome. Science. 2001 Jun 29;292(5526):2482-5. doi: 10.1126/science.1060972. PMID: 11431566.

McInnis S. M., Desikan R., Hancock J. T., & Hiscock S. J. (2006). Production of reactive oxygen species and reactive nitrogen species by angiosperm stigmas and pollen: potential signalling crosstalk?. *New Phytologist*, *172*(2), 221-228.

Meiri D., Tazat K., Cohen-Peer R., Farchi-Pisanty O., Aviezer-Hagai K., Avni A., Breiman A. Involvement of Arabidopsis ROF2 (FKBP65) in thermotolerance. Plant Mol Biol. 2010 Jan;72(1-2):191-203. doi: 10.1007/s11103-009-9561-3. Epub 2009 Oct 29. PMID: 19876748.

Meng J. G., Liang L., Jia P. F., Wang Y. C., Li H. J., Yang W .C. Integration of ovular signals and exocytosis of a Ca^{2+} channel by MLOs in pollen tube guidance. Nat Plants. 2020 Feb;6(2):143-153. doi: 10.1038/s41477-020-0599-1. Epub 2020 Feb 13. PMID: 32055051.

Mitsuda N., Hiratsu K., Todaka D., Nakashima K., Yamaguchi-Shinozaki K., Ohme-Takagi M. Efficient production of male and female sterile plants by expression of a chimeric repressor in Arabidopsis and rice. Plant Biotechnol J. 2006 May;4(3):325-32. doi: 10.1111/j.1467-7652.2006.00184.x. PMID: 17147638.

Mizutani M., Ward E., Ohta D. Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of cDNAs, differential expression, and RFLP mapping of multiple cytochromes P450. Plant Mol Biol. 1998 May;37(1):39-52. doi: 10.1023/a:1005921406884. PMID: 9620263.

Molina I., Li-Beisson Y., Beisson F., Ohlrogge J. B., Pollard M. Identification of an Arabidopsis feruloyl-coenzyme A transferase required for suberin synthesis. Plant Physiol. 2009 Nov;151(3):1317-28. doi: 10.1104/pp.109.144907. Epub 2009 Sep 16. PMID: 19759341; PMCID: PMC2773081.

Moran P. J. and Thompson G. A. Molecular responses to aphid feeding in Arabidopsis in relation to plant defense pathways. Plant Physiol. 2001 Feb;125(2):1074-85. doi: 10.1104/pp.125.2.1074. PMID: 11161062; PMCID: PMC64906.

Moser C., Segala C., Fontana P., Salakhudtinov I., Gatto P., Pindo M., ... & Velasco R. (2005). Comparative analysis of expressed sequence tags from different organs of Vitis vinifera L. *Functional & integrative genomics*, *5*(4), 208-217.

Myers C., Romanowsky S. M., Barron Y. D., Garg S., Azuse C. L., Curran A., Davis R. M., Hatton J., Harmon A. C., Harper J. F. Calcium-dependent protein kinases regulate polarized tip growth in pollen tubes. Plant J. 2009 Aug;59(4):528-39. doi: 10.1111/j.1365-313X.2009.03894.x. Epub 2009 Apr 11. PMID: 19392698.

Muro K., Matsuura-Tokita K., Tsukamoto R. et al. ANTH domain-containing proteins are required for the pollen tube plasma membrane integrity via recycling ANXUR kinases. Commun Biol 1, 152 (2018). https://doi.org/10.1038/s42003-018-0158-8

Navarro-Payá D., Santiago A., Orduña L., Zhang C., Amato A., D'Inca E., Fattorini C., Pezzotti M., Tornielli G. B., Zenoni S., Rustenholz C., & Matus J. T. (2022). The Grape Gene Reference Catalogue as a Standard Resource for Gene Selection and Genetic Improvement. *Frontiers in Plant Science*, 12, 803977. https://doi.org/10.3389/fpls.2021.803977.

Nishiuchi T., Shinshi H., Suzuki K. Rapid and transient activation of transcription of the ERF3 gene by wounding in tobacco leaves: possible involvement of NtWRKYs and autorepression. J Biol Chem. 2004 Dec 31;279(53):55355-61. doi: 10.1074/jbc.M409674200. Epub 2004 Oct 27. PMID: 15509567.

Nwafor C. C., Gribaudo I., Schneider A. *et al.* Transcriptome analysis during berry development provides insights into co-regulated and altered gene expression between a seeded wine grape variety and its seedless somatic variant. *BMC Genomics* **15**, 1030 (2014). https://doi.org/10.1186/1471-2164-15-1030

O'Neill S. D., Kumagai M. H., Majumdar A., Huang N., Sutliff T. D., Rodriguez R. L. The alpha-amylase genes in Oryza sativa: characterization of cDNA clones and mRNA expression during seed germination. Mol Gen Genet. 1990 Apr;221(2):235-44. doi: 10.1007/BF00261726. PMID: 2370848.

Pagnussat G. C., Yu H. J., Ngo Q. A., Rajani S., Mayalagu S., Johnson C. S., Capron A., Xie L. F., Ye D., Sundaresan V. Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. Development. 2005 Feb;132(3):603-14. doi: 10.1242/dev.01595. Epub 2005 Jan 5. Erratum in: Development. 2005 Mar;132(5):1161. PMID: 15634699.

Phillips A. L., Ward D. A., Uknes S., Appleford N. E., Lange T., Huttly A. K., Gaskin P., Graebe J. E., Hedden P. Isolation and expression of three gibberellin 20-oxidase cDNA clones from Arabidopsis. Plant Physiol. 1995 Jul;108(3):1049-57. doi: 10.1104/pp.108.3.1049. PMID: 7630935; PMCID: PMC157456.

Pilati S., Bagagli G., Sonego P., Moretto M., Brazzale D., Castorina G., ... & Moser C. (2017). Abscisic acid is a major regulator of grape berry ripening onset: new insights into ABA signaling network. Frontiers in plant science, 8, 1093.

Plesch G., Ehrhardt T., Mueller-Roeber B. Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression. Plant J. 2001 Nov;28(4):455-64. doi: 10.1046/j.1365-313x.2001.01166.x. PMID: 11737782.

Qin Y., Leydon A. R., Manziello A., Pandey R., Mount D., Denic S., ... & Palanivelu R. (2009). Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. *PLoS genetics*, *5*(8), e1000621.

Ramon M., Rolland F., Thevelein J.M. *et al.* ABI4 mediates the effects of exogenous trehalose on *Arabidopsis* growth and starch breakdown. *Plant Mol Biol* **63**, 195–206 (2007). https://doi.org/10.1007/s11103-006-9082-2

Ramos M. J. N., Coito J. L., Silva H. G. *et al.* Flower development and sex specification in wild grapevine. *BMC Genomics* **15**, 1095 (2014). https://doi.org/10.1186/1471-2164-15-1095

Richter R., Rossmann S., Gabriel D., Töpfer R., Theres K., & Zyprian E. (2020). Differential expression of transcription factor-and further growth-related genes correlates with contrasting

cluster architecture in Vitis vinifera 'Pinot Noir'and Vitis spp. genotypes. Theoretical and Applied Genetics, 133(12), 3249-3272.

Rieu I., Ruiz-Rivero O., Fernandez-Garcia N., Griffiths J., Powers S. J., Gong F., Linhartova T., Eriksson S., Nilsson O., Thomas S. G., Phillips A. L., Hedden P. The gibberellin biosynthetic genes AtGA20ox1 and AtGA20ox2 act, partially redundantly, to promote growth and development throughout the Arabidopsis life cycle. Plant J. 2008 Feb;53(3):488-504. doi: 10.1111/j.1365-313X.2007.03356.x. Epub 2007 Dec 6. PMID: 18069939.

Rossmann S., Richter R., Sun H., Schneeberger K., Töpfer R., Zyprian E., & Theres K. (2020). Mutations in the miR396 binding site of the growth-regulating factor gene VVGRF4 modulate inflorescence architecture in grapevine. *The Plant Journal*, *101*(5), 1234-1248.

Royo C., Carbonell-Bejerano P., Torres-Pérez R., Nebish A., Martínez O., Rey M., Aroutiounian R., Ibáñez J., Martínez-Zapater J. M. Developmental, transcriptome, and genetic alterations associated with parthenocarpy in the grapevine seedless somatic variant Corinto bianco, *Journal of Experimental Botany*, Volume 67, Issue 1, January 2016, Pages 259–273, https://doi.org/10.1093/jxb/erv452

Santopolo S., Boccaccini A., Lorrai R., Ruta V., Capauto D., Minutello E., Serino G., Costantino P., Vittorioso P. DOF AFFECTING GERMINATION 2 is a positive regulator of light-mediated seed germination and is repressed by DOF AFFECTING GERMINATION 1. BMC Plant Biol. 2015 Mar 4;15:72. doi: 10.1186/s12870-015-0453-1. PMID: 25850831; PMCID: PMC4355143.

Schuler M., Rellán-Álvarez R., Fink-Straube C., Abadía J., Bauer P. Nicotianamine functions in the Phloem-based transport of iron to sink organs, in pollen development and pollen tube growth in Arabidopsis. Plant Cell. 2012 Jun;24(6):2380-400. doi: 10.1105/tpc.112.099077. Epub 2012 Jun 15. PMID: 22706286; PMCID: PMC3406910.

Seo M., Koiwai H., Akaba S., Komano T., Oritani T., Kamiya Y., & Koshiba T. (2000). Abscisic aldehyde oxidase in leaves of Arabidopsis thaliana. The Plant Journal, 23(4), 481-488.

Silva P.Â., Ul-Rehman R., Rato C. et al. Asymmetric localization of Arabidopsis SYP124 syntaxin at the pollen tube apical and sub-apical zones is involved in tip growth. BMC Plant Biol 10, 179 (2010). https://doi.org/10.1186/1471-2229-10-179

Shultz R. W., Lee T. J., Allen G. C., Thompson W. F., Hanley-Bowdoin L. Dynamic localization of the DNA replication proteins MCM5 and MCM7 in plants. Plant Physiol. 2009 Jun;150(2):658-69. doi: 10.1104/pp.109.136614. Epub 2009 Apr 8. PMID: 19357199; PMCID: PMC2689970.

Steiner-Lange S., Unte U. S., Eckstein L., Yang C., Wilson Z. A., Schmelzer E., Dekker K., Saedler H. Disruption of Arabidopsis thaliana MYB26 results in male sterility due to non-dehiscent anthers. Plant J. 2003 May;34(4):519-28. doi: 10.1046/j.1365-313x.2003.01745.x. PMID: 12753590.

Storf S., Jansson S., Schmid V. H. Pigment binding, fluorescence properties, and oligomerization behavior of Lhca5, a novel light-harvesting protein. J Biol Chem. 2005 Feb 18;280(7):5163-8. doi: 10.1074/jbc.M411248200. Epub 2004 Nov 23. PMID: 15563470.

Sun L., Zhu B., Zhang X. *et al.* Transcriptome profiles of three Muscat table grape cultivars to dissect the mechanism of terpene biosynthesis. *Sci Data* **6**, 89 (2019). https://doi.org/10.1038/s41597-019-0101-y

Swanson R., Clark T. & Preuss D. Expression profiling of Arabidopsis stigma tissue identifies stigma-specific genes. Sex Plant Reprod 18, 163–171 (2005). https://doi.org/10.1007/s00497-005-0009-x

Tantikanjana T., Rizvi N., Nasrallah M. E., Nasrallah J. B. A Dual Role for the *S*-Locus Receptor Kinase in Self-Incompatibility and Pistil Development Revealed by an *Arabidopsis rdr6* Mutation, *The Plant Cell*, Volume 21, Issue 9, September 2009, Pages 2642–2654, https://doi.org/10.1105/tpc.109.067801

Terrier N., Glissant D., Grimplet J., Barrieu F., Abbal P., Couture C., ... & Hamdi S. (2005). Isogene specific oligo arrays reveal multifaceted changes in gene expression during grape berry (Vitis vinifera L.) development. *Planta*, *222*(5), 832-847.

Tjaden G., Edwards J. W., Coruzzi G. M. Cis elements and trans-acting factors affecting regulation of a nonphotosynthetic light-regulated gene for chloroplast glutamine synthetase. Plant Physiol. 1995 Jul;108(3):1109-17. doi: 10.1104/pp.108.3.1109. PMID: 7630938; PMCID: PMC157463.

The UniProt Consortium UniProt: the universal protein knowledgebase in 2021 Nucleic Acids Res. 49:D1 (2021).

Thomas M.R., Matsumoto S., Cain P. *et al.* Repetitive DNA of grapevine: classes present and sequences suitable for cultivar identification. *Theoret. Appl. Genetics* 86, 173–180 (1993). https://doi.org/10.1007/BF00222076

To A., Joubès J., Barthole G., Lécureuil A., Scagnelli A., Jasinski S., Lepiniec L., Baud S., WRINKLED Transcription Factors Orchestrate Tissue-Specific Regulation of Fatty Acid Biosynthesis in Arabidopsis , The Plant Cell, Volume 24, Issue 12, December 2012, Pages 5007–5023, https://doi.org/10.1105/tpc.112.106120

Tsuwamoto R., Fukuoka H., Takahata Y. GASSHO1 and GASSHO2 encoding a putative leucine-rich repeat transmembrane-type receptor kinase are essential for the normal development of the epidermal surface in Arabidopsis embryos. Plant J. 2008 Apr;54(1):30-42. doi: 10.1111/j.1365-313X.2007.03395.x. Epub 2007 Dec 15. PMID: 18088309.

Velasco R., Zharkikh A., Troggio M., Cartwright D. A., Cestaro A., Pruss D., ... & Viola R. (2007). A high-quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PloS one*, *2*(12), e1326.

Vellosillo T., Martínez M., López M. A., Vicente J., Cascón T., Dolan L., Hamberg M., Castresana C. Oxylipins produced by the 9-lipoxygenase pathway in Arabidopsis regulate lateral root development and defense responses through a specific signaling cascade. Plant Cell. 2007 Mar;19(3):831-46. doi: 10.1105/tpc.106.046052. Epub 2007 Mar 16. PMID: 17369372; PMCID: PMC1867370.

Velt A., Frommer B., Blanc S., Holtgräwe D., Duchêne È., Dumas, V., ... & Rustenholz C. (2022). An improved reference of the grapevine genome supports reasserting the origin of the PN40024 highly-homozygous genotype. bioRxiv.

Villain P., Mache R., Zhou D. X. The mechanism of GT element-mediated cell type-specific transcriptional control. J Biol Chem. 1996 Dec 20;271(51):32593-8. doi: 10.1074/jbc.271.51.32593. PMID: 8955086.

Vucich V. A. and Gasser C. S. Novel structure of a high molecular weight FK506 binding protein from Arabidopsis thaliana. Mol Gen Genet. 1996 Oct 16;252(5):510-7. doi: 10.1007/BF02172397. PMID: 8914512.

Wang B., Jin S. H., Hu H. Q., Sun Y. G., Wang Y. W., Han P., Hou B. K. UGT87A2, an Arabidopsis glycosyltransferase, regulates flowering time via FLOWERING LOCUS C. New Phytol. 2012 May;194(3):666-675. doi: 10.1111/j.1469-8137.2012.04107.x. Epub 2012 Mar 9. PMID: 22404750.

Wang T., Liang L., Xue Y. et al. A receptor heteromer mediates the male perception of female attractants in plants. Nature 531, 241–244 (2016). https://doi.org/10.1038/nature16975

Wang X., Xu X., Mo X. et al. Overexpression of TCP8 delays Arabidopsis flowering through a FLOWERING LOCUS C-dependent pathway. BMC Plant Biol 19, 534 (2019). https://doi.org/10.1186/s12870-019-2157-4

Waters D. L., Holton T. A., Ablett E. M., Lee L. S., & Henry R. J. (2005). cDNA microarray analysis of developing grape (Vitis vinifera cv. Shiraz) berry skin. *Functional & integrative genomics*, *5*(1), 40-58.

Wen Y. Q., Zhong G. Y., Gao Y., Lan Y. B., Duan C. Q., Pan Q. H. Using the combined analysis of transcripts and metabolites to propose key genes for differential terpene accumulation across two regions. BMC Plant Biol. 2015 Oct 6;15:240. doi: 10.1186/s12870-015-0631-1. PMID: 26444528; PMCID: PMC4595271.

Wenkel S., Turck F., Singer K., Gissot L., Le Gourrierec J., Samach A., Coupland G. CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis. Plant Cell. 2006 Nov;18(11):2971-84. doi: 10.1105/tpc.106.043299. Epub 2006 Nov 30. PMID: 17138697; PMCID: PMC1693937.

Wientjes E. and Croce R. The light-harvesting complexes of higher-plant Photosystem I: Lhca1/4 and Lhca2/3 form two red-emitting heterodimers. Biochem J. 2011 Feb 1;433(3):477-85. doi: 10.1042/BJ20101538. PMID: 21083539.

Wientjes E., van Stokkum I. H., van Amerongen H., Croce R. The role of the individual Lhcas in photosystem I excitation energy trapping. Biophys J. 2011 Aug 3;101(3):745-54. doi: 10.1016/j.bpj.2011.06.045. PMID: 21806943; PMCID: PMC3145314.

Winter D., Vinegar B., Nahal H., Ammar R., Wilson G. V., & Provart N. J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PloS one*, *2*(8), e718.

Wong D. C. J., Schlechter R., Vannozzi A., Höll J., Hmmam I., Bogs J., Tornielli G. B., Castellarin S. D., Matus J. T. A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family uncovers new insights into the regulation of stilbene accumulation, DNA Research, Volume 23, Issue 5, October 2016, Pages 451–466, https://doi.org/10.1093/dnares/dsw028

Xia C., Wang Y. J., Liang Y., Niu Q. K., Tan X. Y., Chu L. C., Chen L. Q., Zhang X. Q., Ye D. The ARID-HMG DNA-binding protein AtHMGB15 is required for pollen tube growth in Arabidopsis thaliana. Plant J. 2014 Sep;79(5):741-56. doi: 10.1111/tpj.12582. Epub 2014 Jul 23. PMID: 24923357.

Yanagisawa S. Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. Plant J. 2000 Feb;21(3):281-8. doi: 10.1046/j.1365-313x.2000.00685.x. PMID: 10758479.

Yang C., Xu Z., Song J., Conner K., Vizcay Barrena G., Wilson Z. A.. Arabidopsis MYB26/MALE STERILE35 regulates secondary thickening in the endothecium and is essential for anther dehiscence. Plant Cell. 2007 Feb;19(2):534-48. doi: 10.1105/tpc.106.046391. Epub 2007 Feb 28. PMID: 17329564; PMCID: PMC1867336.

Yang Y., Yu Y., Liang Y., Anderson C. T., & Cao J. (2018). A profusion of molecular scissors for pectins: classification, expression, and functions of plant polygalacturonases. Frontiers in plant science, 9, 1208.

Yu F., Shi J., Zhou J., Gu J., Chen Q., Li J., Cheng W., Mao D., Tian L., Buchanan B. B., Li L., Chen L., Li D., Luan S. ANK6, a mitochondrial ankyrin repeat protein, is required for male-female gamete

recognition in Arabidopsis thaliana. Proc Natl Acad Sci U S A. 2010 Dec 21;107(51):22332-7. doi: 10.1073/pnas.1015911107. Epub 2010 Dec 1. PMID: 21123745; PMCID: PMC3009778.

Zander M., Lewsey M. G., Clark N. M. et al. Integrated multi-omics framework of the plant response to jasmonic acid. Nat. Plants 6, 290–302 (2020). https://doi.org/10.1038/s41477-020-0605-7

Zhang D., Liu D., Lv X., Wang Y., Xun Z., Liu Z., Li F., Lu H. The cysteine protease CEP1, a key executor involved in tapetal programmed cell death, regulates pollen development in Arabidopsis. Plant Cell. 2014 Jul;26(7):2939-61. doi: 10.1105/tpc.114.127282. Epub 2014 Jul 17. PMID: 25035401; PMCID: PMC4145124.

Zhang Y., Li Z., Liu J. et al. Transposable elements orchestrate subgenome-convergent and -divergent transcription in common wheat. Nat Commun 13, 6940 (2022). https://doi.org/10.1038/s41467-022-34290-w

Zhang Z. L., Xie Z., Zou X., Casaretto J., Ho T. H., Shen Q. J. A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. Plant Physiol. 2004 Apr;134(4):1500-13. doi: 10.1104/pp.103.034967. Epub 2004 Mar 26. PMID: 15047897; PMCID: PMC419826.

Zhong R., Richardson E. A., Ye Z. H. The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. Plant Cell. 2007 Sep;19(9):2776-92. doi: 10.1105/tpc.107.053678. Epub 2007 Sep 21. PMID: 17890373; PMCID: PMC2048704.

Zhou D. X. (1999). Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. Trends in plant science, 4(6), 210-214.

Zhu L., Gazin C., Lawson N., Pagès H., Lin S., Lapointe D. et al. (2010) Chippeakanno: a bioconductor package to annotate chip-seq and chip-chip data. *BMC Bioinformatics*, **11**, 237. https://doi.org/10.1186/1471-2105-11-237.

Paper published in The Plant Journal (https://doi.org/10.1111/tpj.15686)

CHAPTER IV: Direct regulation of shikimate, early phenylpropanoid, and stilbenoid pathways by Subgroup 2 R2R3-MYBs in grapevine

Luis Orduna [1], Miaomiao Li[2], David Navarro-Payà[1], Chen Zhang[1], Antonio Santiago[1], Pablo Romero[1], Ziva Ramsak[3], Gabriele Magon[4], Janine Holl[5], Patrick Merz[5], Kristina Gruden[3], Alessandro Vannozzi[4], Dario Cantu[6], Jochen Bogs[5], Darren C. J. Wong[7], Shao-shan Carol Huang[2] and Josè Tomàs Matus[1]*

[1] Institute for Integrative Systems Biology (I2SysBio), Universitat de Valencia-CSIC, Paterna, 46908, Valencia, Spain

[2] Center for Genomics and Systems Biology, Department of Biology, New York University, USA

[3] Department of Biotechnology and Systems Biology, National Institute of Biology, Vecna pot 111, 1000, Ljubljana, Slovenia

[4] Department of Agronomy, Food, Natural resources, Animals, and Environment (DAFNAE), University of Padova, Legnaro 35020, Italy

[5] Dienstleistungszentrum Landlicher Raum Rheinpfalz, Viticulture and Enology Group, Neustadt/W, Germany

[6] Department of Viticulture and Enology, University of California Davis, Davis, California, USA

[7] Ecology and Evolution, Research School of Biology, The Australian National University, Acton, Australia

*Correspondence: tomas.matus@gmail.com, tomas.matus@uv.es

4.1 Abstract

The stilbenoid pathway is responsible for the production of resveratrol in grapevine (*Vitis vinifera* L.). A few transcription factors (TFs) have been identified as regulators of this pathway but the extent of this control has not been deeply studied. Here we show how DNA affinity purification sequencing (DAP-Seq) allows for the genome-wide TF-binding site interrogation in grape. We obtained 5190 and 4443 binding events assigned to 4041 and 3626 genes for MYB14 and MYB15, respectively (approximately 40% of peaks located within -10 kb of transcription start sites). DAP-Seq of MYB14/MYB15 was combined with aggregate gene co-expression networks (GCNs) built from more than 1400 transcriptomic datasets from leaves, fruits, and flowers to narrow down bound genes to a set of high confidence targets. The analysis of MYB14, MYB15, and MYB13, a third uncharacterized member of Subgroup 2 (S2), showed that in addition to the few previously known stilbene synthase (STS) targets, these regulators bind to 30 of 47 STS family genes. Moreover, all three MYBs bind to several PAL, C4H, and 4CL genes, in addition to shikimate pathway genes, the WRKY03 stilbenoid co-regulator and resveratrol-modifying gene

candidates among which ROMT2-3 were validated enzymatically. A high proportion of DAP-Seq bound genes were induced in the activated transcriptomes of transient MYB15-overexpressing grapevine leaves, validating our methodological approach for delimiting TF targets. Overall, Subgroup 2 R2R3-MYBs appear to play a key role in binding and directly regulating several primary and secondary metabolic steps leading to an increased flux towards stilbenoid production. The integration of DAP-Seq and reciprocal GCNs offers a rapid framework for gene function characterization using genome-wide approaches in the context of non-model plant species and stands up as a valid first approach for identifying gene regulatory networks of specialized metabolism.

Keywords: *secondary metabolism, regulatory networks, transcription factors, transcriptional regulation, DNA affinity purification sequencing*

4.2 Introduction

The evolved complexity of plant specialized metabolism can be traced back to the colonization of dry land by green algal-derived ancestors (Waters, 2003). Since their origin, land plants have been accompanied by a persistent range of abiotic and biotic stresses such as ultraviolet (UV) radiation, desiccation, or unfavorable microbial communities that have led to the selective emergence of novel protective metabolites derived from primary metabolism (Kenrick & Crane, 1997). A clear example of this is the phenylpropanoid pathway (PPP), a major source of aromatic secondary metabolites in plants, with many of their roles related to tolerance and adaptation to the abovementioned stresses (Davies et al., 2020). While some branches of the pathway are ubiquitous in the plant kingdom (such as those producing flavonoids), others such as the stilbene pathway are restricted to a small number of species across at least 10 unrelated families, including *Vitaceae* (e.g. *Vitis vinifera* L.) and *Moraceae* (e.g. *Morus alba*) (Dubrovina & Kiselev, 2017). Grapevine is not only an important crop species but an interesting model to study the complexity of the stilbene pathway given the remarkable expansion of the stilbene synthase (STS) family in its genome through segmental and tandem gene duplications, reaching up to, in total, 47 genes (Parage et al., 2012; Vannozzi et al., 2012), albeit 12 of these are considered pseudogenes. Stilbenes are phytoalexins, which are small, lipophilic compounds with key roles in plant defense, and they accumulate in response to a range of abiotic and biotic stresses. In recent years, the grapevine STS gene family has been studied to reveal a high degree of responsiveness and effectiveness against different biotic or abiotic stresses. For instance, ectopic expression of VqSTS36 from the Chinese wild species *Vitis quinquangularis*, in both *Arabidopsis* and tomato, enhanced resistance to powdery mildew and osmotic stress (Huang et al., 2018), while the expression of VqSTS29 in *Arabidopsis* also led to powdery mildew resistance (Xu et al., 2019). Moreover, the induction of STS gene expression in different *V. vinifera* tissues has been observed in response to fungal infection, UVC, or heat treatments (Lecourieux et al., 2017; Vannozzi et al., 2012; Yin et al., 2016). STS enzymes are direct competitors of chalcone synthases for pathway precursors. Both proteins are closely related type III polyketide synthases, generating a tetraketide intermediate from the condensation of p-coumaroylcoenzyme A with three molecules of malonyl-coenzyme A, which depending on STS/chalcone synthase activity, will generate resveratrol or naringenin chalcone, respectively, thus defining the entry point of the stilbene and flavonoid branches. The main and first stilbene produced in grape tissues is resveratrol, a well-known nutraceutical with many characterized properties ranging from antioxidant to antiviral activities (e.g. it has been recently shown to inhibit SARSCoV-2 in vitro replication in human lung cells; Pasquereau et al., 2021). Resveratrol is derivatized into a broad range of stilbenes such as pterostilbene, viniferins, piceid, and

piceatannol, which involve methoxylation, oligomerization, glucosylation, and hydroxylation, respectively. The enzymes catalysing these reactions are mostly uncharacterized in grapevine except for a resveratrol Omethyltransferase (ROMT1) responsible for the production of pterostilbene (Schmidlin et al., 2008) and a resveratrol glucosyl transferase in Vitis labrusca, leading to the production of piceid (Hall & De Luca, 2007). Hydroxylation of resveratrol into piceatannol could be carried out by cytochrome P450 oxidoreductases but no candidates have yet been identified in grape. The stilbene pathway in grapevine is mainly regulated at the transcriptional level through transcription factors (TFs). In particular, the R2R3-type MYB14 and MYB15 [members from Subgroup 2 (S2)] have been shown to activate a few STS promoters specifically (STS29 and STS41) in transient reporter assays (Holl et al., 2013), but modulation of other stilbenoid branch enzymes remains unexplored. Furthermore, gene co-expression networks (GCNs) and further correlation with stilbene accumulation have pointed out MYB13 as an additional putative regulator of stilbene accumulation (Wong et al., 2016) although it has not yet been validated in planta. We have initially explored TF regulatory networks interrogated by the use of GCNs, leading to the identification of members of the AP2/ERF, bZIP, and WRKY gene families as potential regulators of STS expression (Wong & Matus, 2017); however, experimental evidence of binding is necessary to prove regulatory causality. Nevertheless, systems biology approaches initially conducted in Wong et al. (2016), applied to the regulation of transcription in grapevine, have paved the way for the functional characterization of additional stilbene pathway regulators such as bZIP1, ERF114, MYB35A, and WRKY53 in recent years (Vannozzi et al., 2018; Wang et al., 2019; Wang & Wang, 2019), proving their efficacy in hypothesis-driven research for TF discovery. Interestingly, one of the newly identified stilbene regulators, WRKY53, binds to a subset of STS genes (STS32 and STS41) and it is thought to form a regulatory complex with MYB14 and MYB15, probably increasing its activity (Wang et al., 2020). Moreover, WRKY03 has also been shown to work in synergy with MYB14 in the upregulation of STS29 expression (Vannozzi et al., 2018). Among the identified STS regulators, MYB14 and MYB15 have been proposed as upstream TFs in the regulatory cascade governing stilbenoid accumulation mainly due to their rapid activation response. However, this high hierarchy has no experimental validation to date. MYB14 and MYB15 have been shown to be differentially expressed under biotic and abiotic stresses in different *Vitis* species and cultivars with varying stilbene accumulation levels. For instance, the higher stilbene content of *V. labrusca* cv. 'Concord', in response to UVC and Al3+, compared with *V. vinifera* cv. 'Cabernet Sauvignon' has been directly linked to a greater MYB14 promoter activity (Bai et al., 2019). The detailed characterization of these two potential regulators is hence of great importance. In addition, no other processes controlled by these TFs have been identified. This study combines genome-wide TF binding-site

interrogation using DNA affinity purification sequencing (DAP-Seq), and aggregate whole genome co-expression networks to lay out MYB14 and MYB15 cistrome landscapes and identify their complete repertoire of target genes. We have also analyzed the yet uncharacterized MYB13. Our results suggest that these three MYBs bind to regulatory elements in most members of the STS gene family, other shikimate and early phenylpropanoid genes as well as bind to WRKY regulators of stilbene synthesis, representing major regulators of this specialized metabolic pathway.

4.3 Results

4.3.1 MYB14 and MYB15 bind proximal upstream regions of a large set of STS genes

Our DAP-Seq analysis reported 5190 and 4443 TF-binding events, i.e. peaks (Fig. 1.4A and Dataset S1), which were assigned to 4041 and 3626 different genes for MYB14 and MYB15, respectively. An initial inspection of all binding events showed that 73% and 75% of the peaks are present between 10 kb upstream of transcription start sites (TSSs) and 2 kb downstream of annotated gene ends for MYB14 and MYB15, respectively. In total, 30% and 32% of MYB14 and MYB15 peaks, respectively, are found within 5 kb upstream of TSSs. MYB14 and MYB15 shared, in total, 2709 bound genes as well as an almost identical DNA-binding motif obtained from the enrichment analysis of the top 600 most significant peaks sequences (Fig. 1.4B). The cistromes of MYB14 and MYB15 *Arabidopsis* orthologues have not been previously studied. A closer look at the STS gene family revealed that 22 of the 47 family members have MYB14/MYB15 DAP-Seq peaks associated to them (including five of 13 pseudogenes). Examining the STS genomic regions revealed clear DNA-binding signals upstream of TSSs for both MYB14 and MYB15 TFs, which was not observed in the pIX-HALO non-specific DNA binding control (Fig. 2.4). A selection of housekeeping genes was used as a negative control with no specific DNA binding of MYB14 or MYB15 TFs around their TSSs. Interestingly, many DNA bound genes belong to the shikimate pathway or even to the first committed steps of the PPP. In addition to STS, other enzyme categories such as O-methyltransferases and glucosyltransferases are also represented among bound genes.

4.3.2 Inspection of MYB14 and MYB15 gene-centered co-expression networks

We generated condition-dependent leaf and fruit aggregate whole genome co-expression networks using public data uploaded in Sequence Read Archive (SRA), composed of >1400 runs belonging to >70 independent experiments (Dataset S2). Network performance was assessed through the area under the receiver operator characteristic curve (AUROC) measurement of the final aggregate network, obtaining AUROC values of approximately 0.75. The final fruit/flower network consisted of 31.723 genes as individual nodes and 13.323.660 co-expression connections as directed edges (weighted by frequency of coexpression across the selected SRA studies). The leaf network consisted of 31.759 genes and 13.338.780 edges. The two networks share, in total, 1.795.970 common interactions, showing a great number of tissue-specific co-expression connections. GCNs centered on individual genes were extracted from the final aggregate leaf and fruit/ flower networks in the form of the top 420 interactions for each

particular gene (corresponding to the top 1% of all grapevine genes according to the V.Cost annotation). All network data as well as a range of visualization tools are publicly available at VitViz (https://tomsbiolab.com/vitviz). As MYB14 and MYB15 have very low expression in flower tissues, any detected co-expression relationship with other genes in the whole genome network are most surely attributable to fruit samples, therefore, fruit/flower GCNs are referred from here on as fruit GCNs. Individual MYB14- and MYB15-gene centered fruit GCNs revealed that both TFs shared 146 genes of the 420 GCN members. In addition, MYB14 is present in the GCN of MYB15 and vice versa. We found 13 STS genes present in both fruit GCNs. In addition, 28 STS genes are exclusively present in the fruit GCN of MYB14 meaning that almost all known STS genes in the grapevine genome (41 of 47) are present in any one of the fruit GCNs. The six STS genes, which are not present in either TF-centered GCN, are STS4/ 11/33/34/44, all of which are putative pseudogenes. Interestingly, MYB15 does not have any exclusive STSs in its fruit GCN, suggesting a more direct relationship of MYB14 with this gene family. Other notable genes that are common to both GCNs are PALs, WRKYs, and other genes related to secondary metabolism. There are five PALs shared by both fruit GCNs, while the MYB14 and MYB15 fruit GCNs have two and one exclusive PAL gene, respectively. A comparison of MYB14 and MYB15 leaf GCNs revealed 180 common genes of 420, with MYB14 and MYB15 again being present in each other's GCNs. Regarding STS genes, 28 of 47 are present in both MYB GCNs, while four and eight are exclusive for MYB14 and MYB15, respectively. The presence of other notable genes of the secondary metabolism pathways in both GCNs, such as PALs and WRKYs, is also remarkable. Results in both leaf and fruit GCNs greatly overlap with DAP-Seq data, providing transcriptional regulation evidence for many co-expression relationships (Fig. 3.4). In total, the number of co-expressed genes in both fruit GCNs is of 694, of which 23% were identified as MYB-bound genes by DAP-Seq. A similar tendency is observed for the comparison between leaf GCNs and DAP-Seq data. Moreover, among these are genes of interest coding for TFs such as WRKYs, NACs, or enzymes within the shikimate, early phenylpropanoid or stilbenoid pathways. Most of the genes identified both by DAP-Seq and co-expression analyses had the MYB binding motifs identified within 5 kb upstream of TSSs. An overlap of these data with the grapevine reference gene catalogue v1.1 (Navarro-Paya et al., 2022) available at Integrape (http://www.integrape.eu/index.php/resources/genomes/) revealed a considerable representation of secondary metabolism genes (all the provided gene symbols in this study are in accordance with the reference gene catalogue, v1.1).

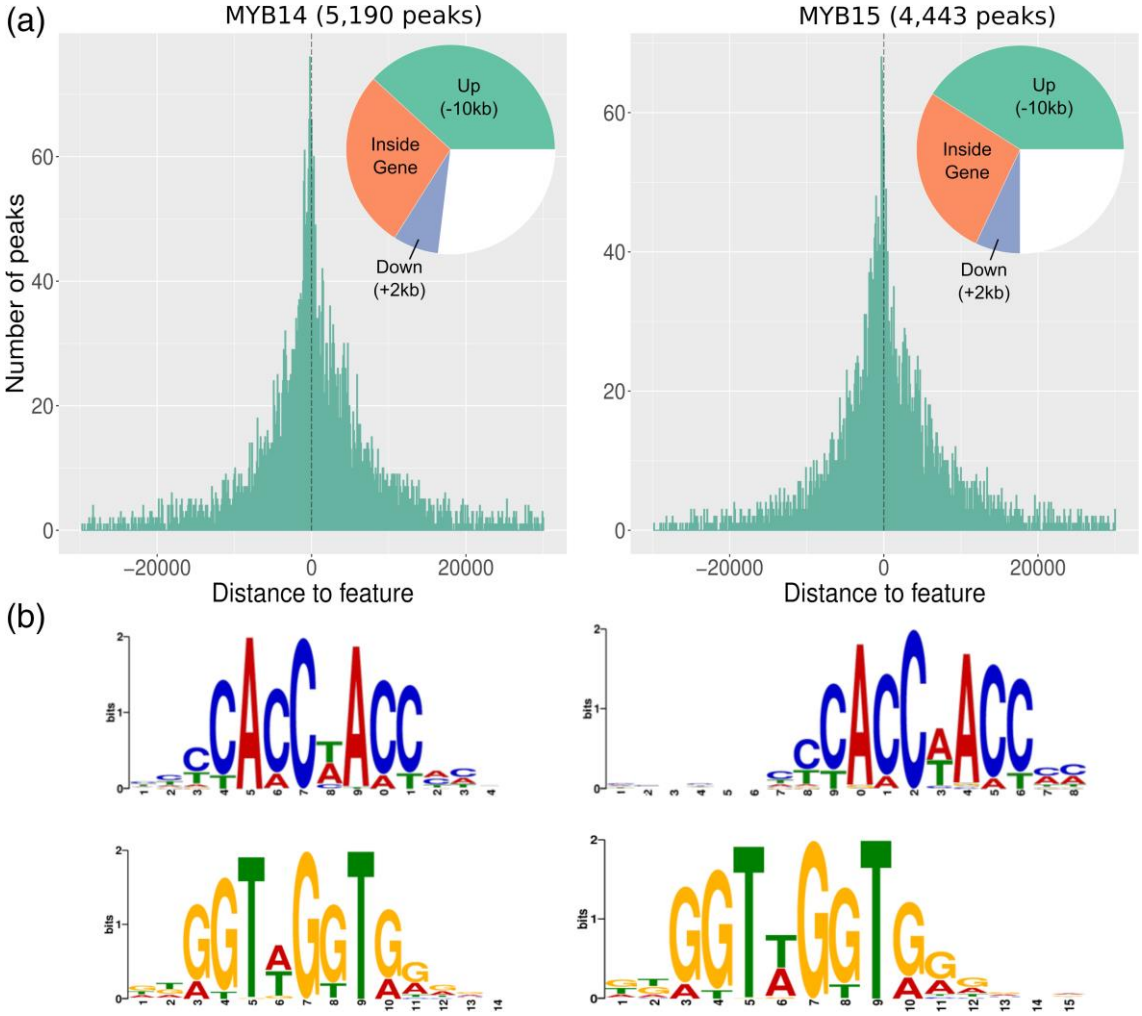


Figure 1.4. MYB14 and MYB15 DAP-Seq derived cistrome landscapes in *Vitis vinifera* cultivar 'PN40024'. (A) DNA-binding events with respect to all transcription start sites of assigned genes. The proportion of binding peaks 10 kb upstream of transcription start sites, inside genes or 2 kb downstream of genes are represented within the pie-charts in green, orange, and blue, respectively. (B) *De novo* binding motifs, forward and reverse, obtained from the top 600 scoring peaks of MYB14 and MYB15 using MEME suite.

4.3.3 Integrating GCNs and DAP-Seq data to identify high confidence targets

Bearing in mind that MYB14 and MYB15 may be major regulators and hence be co-expressed with a large number of genes, a fraction of biologically relevant co-expressed genes is expected to be absent in their GCNs, which have a 420-gene cut-off. Therefore, in addition to the MYB14 and MYB15 GCNs, the GCNs of MYB-bound genes were also inspected to interrogate the presence of either MYB14 or MYB15. These 'reciprocal' gene co-expression relationships allowed further prediction of a reliable list of targets. Gene-centered GCNs were extracted for each DAP-Seq identified gene from the whole genome co-expression network (4041 and 3626 for MYB14 and MYB15 respectively). This allowed for further integration of co-expression and

TF-binding results by overlapping DAP-Seq results with the reciprocal GCNs. For this overlap, only MYB-bound genes with peaks within 10 kb upstream of the TSS and 2 kb downstream of the end of their gene feature were considered. Bound genes with a co-expression relationship present in at least one of the two GCNs (i.e. MYB14 being present in a MYB14-bound gene GCN and/or vice versa) were considered high confidence targets (HCTs). This integration of gene-centered GCNs increased the number of MYB-bound genes supported by network data. For instance, in fruit HCTs an increase from 76 to 145 and from 56 to 127 was observed for MYB14 and MYB15, respectively. The different tissues used for the aggregate networks did not affect the total number of MYB HCTs, obtaining 146 and 145 MYB14 HCTs in leaf and fruit, respectively. Gene enrichment analysis conducted for each MYB HCT list, using either Gene Ontology (GO) or Kyoto Encyclopedia of Genes and Genomes (KEGG), show trihydroxystilbene synthase activity as the most significant term (Fig. S1A and Dataset S2). Moreover, there are a number of interesting terms such as shikimate 3-dehydrogenase activity, phenylalanine ammonia-lyase activity, and other shikimate/early phenylpropanoid related terms. Many of the 57 common fruit HCTs in fact correspond to PAL and STS genes. Although gene set enrichment analyses offer similar results for both tissue specific HCTs, the defense response term was exclusive to leaf. Common MYB14/MYB15 fruit HCTs show the presence of at least one known co-regulator of the stilbenoid pathway, i.e. WRKY03, as well as specific shikimate pathway enzymes, such as DAHPS3 and SDH4, and key early phenylpropanoid enzymes such as 4CL8 and C4H1–3. Common fruit HCTs also contain 12 STS genes while individual fruit HCTs contain seven exclusive STS genes in the case of MYB14 and one in the case of MYB15. Common MYB14/ 15 leaf HCTs show similar composition to fruit HCTs, with the presence of WRKY03, 4CL8, C4H2, and 17 different STS genes. Eight STS genes appear as HCTs for fruit and leaf in both MYB14 and MYB15. Only one STS bound gene, the pseudogene STS3, is not an HCT (Fig. S1B). MYB-bound genes are distributed among the three PPP branches (stilbenoid, lignin, and flavonoid); however, a lack of HCTs is observed among the lignin or flavonoid branches in both fruit and leaf HCTs (Fig. S2 and Dataset S2). To corroborate MYB targets further, we overexpressed MYB15 in grapevine leaves.

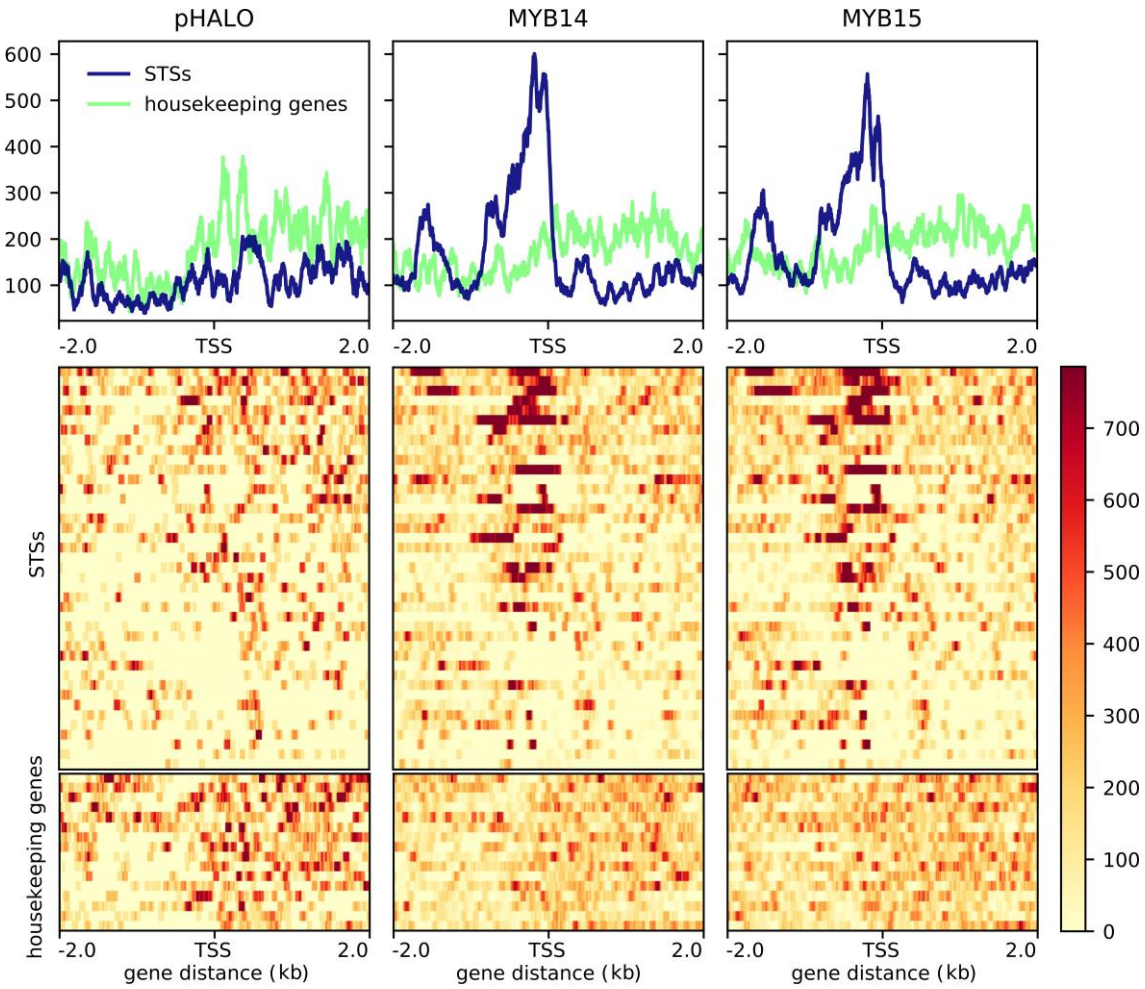


Figure 2.4. MYB14 and MYB15 DNA-binding events in the promoter regions of stilbene synthase (STS) genes. DAP-Seq binding signal for MYB14 and MYB15 at 2 kb and +2 kb from the transcription start site (TSS) of STS genes is high when compared with background housekeeping genes both in the density plots, which show the average binding signal, and the heatmaps showing individual gene profiles. STS and housekeeping genes used for this figure are listed in Dataset S1. The figures were generated using the Deeptools suite v.3.3.2, computing and normalizing the coverage for each BAM file with a BinSize = 10 and RPKM normalization. RPKM value for each bin is the average between the 10 positions that define each bin. BigWig files for the individual replicates for each transcription factor were merged using bigWigMerge v.2 and bedGraphToBigWig v.4.

4.3.4 Transient MYB15 overexpression in grapevine confirms many fruit and leaf predicted targets

We validated MYB15 HCTs by transiently overexpressing it in grapevine leaves from 10-week-old plants and monitoring gene expression changes through microarrays. Endogenous and total MYB15 expression showed a short-term increase at the initial time-points probably due to the agroinfiltration per se (i.e. wounding stress). Nonetheless, a greater and longer-lasting increase in expression was observed thereafter in the MYB15-agroinfiltrated samples, which can be fully attributed to the overexpression of the transgene (Fig. 4.4A and S3). WGCNA analysis resulted

in 18 modules (Fig. 4.4B, left panel). We found the MYB15 probe in module eigengene 5 (ME5), which is nearly identical to ME14, ME2, and ME4, all showing a higher expression in 35S:MYB15 leaves compared with controls at all time-points. These four modules also clustered closely with ME12 and ME16, which still showed similar Z-score patterns across the different samples. We determined that these four modules (ME5, ME2, ME4, and ME14) hold most of MYB15 targets (Fig. 4.4B). The average gene expression of these four modules is indeed higher in MYB15-overexpressed samples with respect to control time-points (Figure 4C). GO and KEGG enrichment analyses on each module (Fig. S4) revealed interesting enrichment terms such as the expected trihydroxystilbene synthase term (ME5). In addition, other terms were O-methyltransferase activity (ME5), other shikimate/early phenylpropanoid enzymatic terms, and defense response terms such as 'response to stress', 'response to fungus', and 'chitin metabolic process'. Probe-associated genes with positive fold changes cluster together and they mostly belong to the four modules of interest. Moreover, the probe representing ROMT1– 6 shows a strong fold induction across all time-points (Fig. 4.4D). A similar induction dynamic is observed for other phenylpropanoid-related enzymes (e.g. 4CL8) and TFs (e.g. WRKY03, WRKY08, or WRKY34).

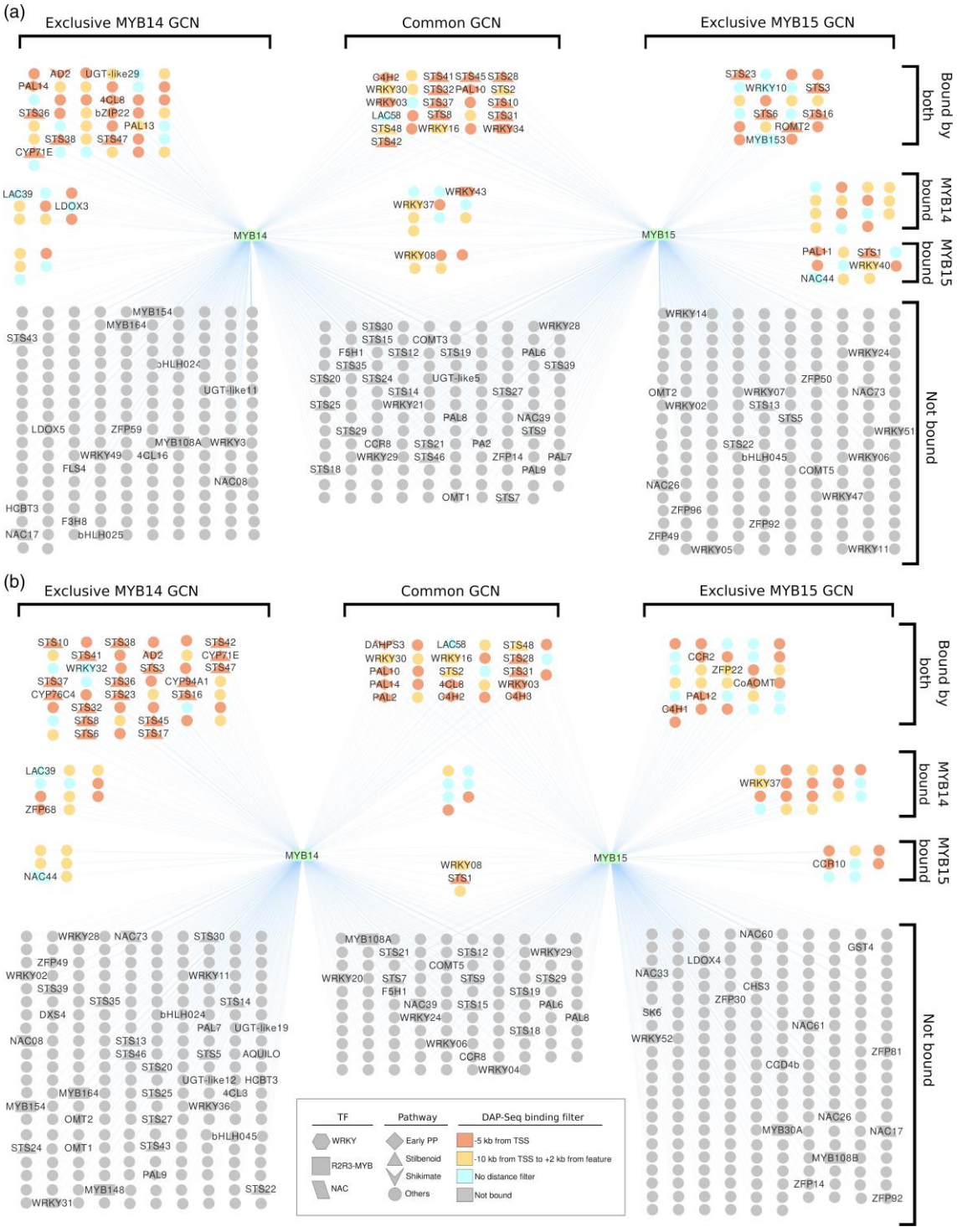


Figure 3.4. MYB14 and MYB15 leaf and fruit gene co-expression networks (GCNs) share many co-expressed genes, many of which are also detected with DNA affinity purification sequencing (DAP-Seq). MYB14 and MYB15 GCNs consist of their top 420 co-expressed genes. The color of each node (i.e. gene) depicts the different distance filters met by the closest DAP-Seq peaks. When genes are bound by both MYB14 and MYB15 the distance filters are only indicated if both transcription factors (TFs) have peaks within them. The shape of each node determines the metabolic pathway or TF family it belongs to. (A) Leaf GCNs. (B) Fruit GCNs.

4.3.5 VviMYB13 shares a high proportion of bound genes with its two S2 co-members

R2R3-MYB S2 is composed of MYB13/14/15 in grapevine and *Arabidopsis*. Alignments and phylogenetic analyses of this subfamily in these and other plant species show that this close evolutionary relationship is in part explained by the conservation of the R2/R3 repeats and their C-terminal FW1 and FW2 domains (Fig. S5). The closest grape MYBs to S2 have lost at least one of these domains (i.e. VviMYB135-138). Thus, we additionally checked the potential contribution of VviMYB13 to the regulation of secondary metabolism in grapevine. Our VviMYB13 DAP-Seq analysis reported 17.019 binding events assigned to 10.624 different genes (Fig. S6 and Dataset S1). Most VviMYB14/15 bound genes are contained within VviMYB13 bound genes (Figure S6B). Approximately 74% of peaks are present between 10 kb upstream of TSSs and 2 kb downstream of the annotated gene ends. VviMYB13 binds to STS promoter regions in a similar manner to VviMYB14/15 (Fig. 5.4B and S7). The DNA-binding motif obtained for VviMYB13 is practically identical to that of VviMYB14/15 (Fig. 1.4A) and even to the binding motif identified for AtMYB13 by (O'Malley et al., 2016) (Fig. 5.4A). Interestingly, VviMYB13 was observed to bind at 2.3 kb from the TSSs of both VviMYB13 itself and VviMYB14. Gene set enrichment analyses for AtMYB13 and VviMYB13/14/15 bound genes illustrate the functional relatedness of grape S2 MYBs with an Arabidopsis homologue (Fig. 5.4C and Dataset S1). Surprisingly, AtMYB13 also presents shikimate- and early phenylpropanoid-related terms as its *V. vinifera* homologues. As *Arabidopsis thaliana* does not produce stilbenes, the trihydroxystilbene synthase activity term is only significantly enriched for VviMYB13/14/15. An interesting term that is only present for AtMYB13 is lignin biosynthesis suggesting a potential regulatory diversification between *Arabidopsis* and grape homologues. Despite the similarities of VviMYB13/14/15 through DAP-Seq analysis, an overall lack of expression of VviMYB13 across the SRA experiments used in the aggregate whole genome co-expression networks meant that no valuable co-expression data could be extracted for VviMYB13 (Fig. 5.4D). Transient overexpression of VviMYB13 in grapevine, or further inspection of datasets where MYB13 is induced could be used to narrow down its bound genes to a list of HCTs. As shown previously, VviMYB13/14/15 bound genes include many members of the shikimate pathway, early PPP, and stilbenoid pathways (Fig. S8). Surprisingly, we also found bound genes among the lignin and flavonoid branches, many of which are known to be oppositely expressed compared with STS and MYB14/15 genes; as an example, chalcone synthase genes have been shown to be strongly repressed under UVC or fungal stress conditions, which in turn largely activate MYB14/15 and STS gene expression (Blanco-Ulate et al., 2015; Vannozzi et al., 2012). This observation and the fact that these

additional bound genes are not HCTs (Fig. S1) may suggest that they are negatively regulated by the interaction of other proteins with R2R3-MYBs from S2.

4.3.6 Characterization of new stilbenoid-pathway genes identified as S2 MYB targets

Several MYB14-15 HCTs and MYB13-bound genes encode different types of functionally uncharacterized enzymes. Among those potentially producing stilbene derivatives, we found several O-methyltransferases, laccases, hydroxylases (cytochrome P450s), and glycosyltransferases, all representing interesting cases for further functional validation, which would position DAP-Seq as a tool for novel enzyme identification. As a first approach to study these potentially new enzymes of the stilbenoid pathway, we conducted a 7-day time-course experiment of grapevine (cv. 'Gamay Freaux') cell cultures, elicited with methyl-jasmonate and cyclodextrins (MeJA + CD) that are known to activate, largely and specifically, the expression of MYB/14/15 (Almagro et al., 2014). As cell suspensions from this teinturier (i.e. red fleshed) cultivar accumulate anthocyanins ectopically in the presence of light, we grew and subcultured the cells for several passes in dark conditions to avoid the premature commitment of the PPP for flavonoid production of flavonoids in detriment of stilbenoids. Anthocyanin-devoid (i.e. white) cells were elicited for the quantification of secondary metabolites at 4 and 7 days. An accumulation of resveratrol was observed only inside MeJA-treated cells at both time-points while anthocyanin content was greatly reduced compared with the control at day 7 (Fig. 6.4 and S9). Anthocyanins show an increasing tendency only in the control due to the effect of sugar replenishment (conducted at the beginning of the time-course) while in the elicited cells jasmonate seems to direct preferentially the flux of the PPP for stilbene accumulation rather than flavonoids despite the effect of the sugars. Moreover, there was an increase of piceatannol and viniferin (particularly on day 7), while piceid accumulation did not present significant changes across treatment and time-points. We compared the accumulation of metabolites with the expression of the shikimate, early phenylpropanoid, and stilbene pathway genes (including the potentially new pathway genes) by reanalysing the previously published microarray study of Almagro et al., (2014) of 24 h MeJA + CD-treated cv. 'Gamay Freaux' cells (Fig. 6.4). We first manually curated and improved the current MapMan grapevine ontology with the complete list of genes within the newly created stilbene pathway terms, among which we created 'Trans-resveratrol di-Omethyltransferase activity' and 'Resveratrol Glycosyltransferase activity' (Dataset S3). The combined DAP-Seq results of MYB13/14/15 were included in the context of these pathways, highlighting those genes that are bound by the S2 MYBs. The metabolite profile observed in our experiment correlates with the activation of MYB15 and almost all its bound

genes (no probe available for MYB14/ MYB13), including resveratrol-modifying candidate genes. Production of viniferin and piceatannol also matched the upregulation of their putative related enzymes. In particular, LAC59 and several cytochrome P450 CYP76C4 genes, identified as MYB14/15 HCTs, could be directly responsible for the enzymatic reactions producing these compounds in our elicited grape cells. The microarray data also support the lack of pterostilbene accumulation, as no upregulation of ROMT1 or ROMT-like genes was observed. We inspected additional public RNA-Seq datasets to see if MYB15/14 targets with potential resveratrol-modifying activity were co-expressed with their corresponding stilbenoid metabolites. For instance, by reanalysing the transcriptome responses upon Botrytis infection stage 2 in cv. 'Semillon' (Blanco-Ulate et al., 2015) we see a large group of potential resveratrol glycosyl-transferase (RsGTs), CYP450 and laccases, many of which were identified as HCT, being highly induced (Fig. S10). In this case, the expression of laccases and RsGTs matched the differential accumulation of viniferins and piceid in response to infection. As both ROMT transcripts and pterostilbene were undetected in the elicited-grape cells or in the Botrytis-infected samples, we further inspected their expression in other transcriptomics datasets to look for their condition-specific induction. The cv. 'Corvina' atlas (Fasoli et al., 2012) represents a suitable dataset with >50 samples corresponding to all types of vegetative and reproductive tissues and green-to-mature stages. The stilbenoid pathway and the regulators WRKY03 and MYB14/15 show high expression in leaf and berry development, in particular at senescent or mature stages for both types of organs. For the case of ROMT1 and ROMT-like genes, we see an almost specific upregulation in senescent leaves (Fig. S11). Thus, we extracted RNA from these senescent grapevine organs and amplified two ROMT-like genes, here named ROMT2 and ROMT3. Together with the previously characterized ROMT1 gene used as a positive control, we transiently overexpressed them in *Nicotiana benthamiana* leaves in combination with STS48 that is also an MYB14/MYB15 HCT. All these three ROMT genes, which show several MYB-binding events upstream of their TSSs, were able to promote pterostilbene accumulation in tobacco leaves as seen by liquid chromatography (LC)–mass spectrometry (MS) analysis (Fig. 7.4). On the contrary, the sole overexpression of STS48 only produced piceid accumulation while the agroinfiltrated empty vector did not produce any stilbenes. Overall, the DAP-Seq data presented here allowed us to select candidate pathways genes for their enzymatic characterization *in planta*.

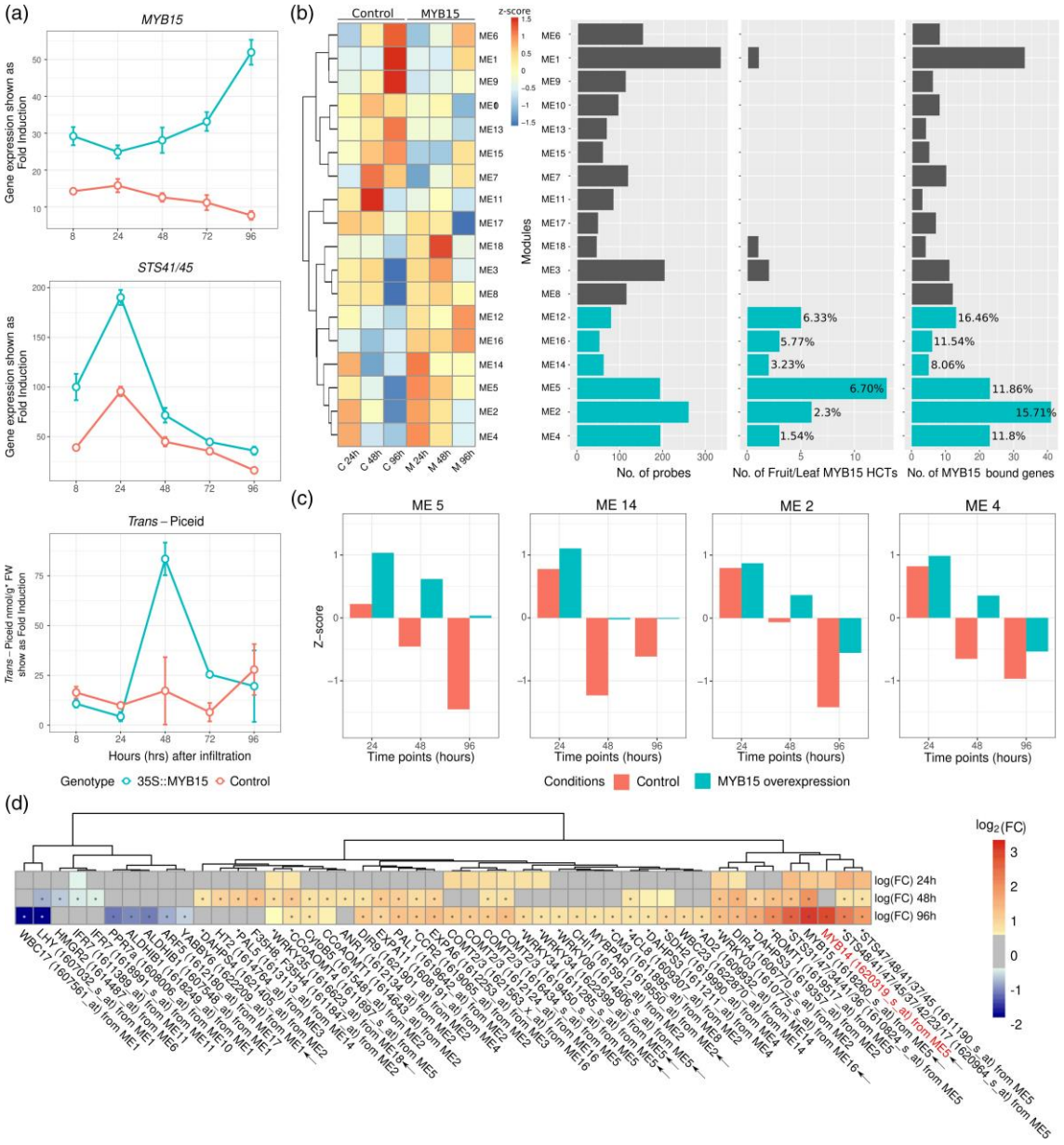


Figure 4.4. Secondary metabolism genes are induced upon overexpression of MYB15 in grapevine leaves. (A) Increase in MYB15 and STS expression as well as in trans-piceid content expressed as fold induction compared with non-infiltrated leaves for both 35S::VviMYB15 and empty vector control transformations (blue and red, respectively). (B) Clustered mean Z-scores across each time-point for modules obtained by WGCNA from log2(RMA + 1) expression data. Bar plots for every module show: (i) number of probes per module, (ii) number of probes with at least one MYB15 leaf or fruit high confidence target (HCT) assigned, and (iii) number of probes with at least one MYB15 DNA affinity purification sequencing bound gene assigned. Percentages are provided for the four cluster modules of interest (marked in blue). ME, module eigengene. (C) Mean Z-scores across different leaf samples grouped by agroinfiltration treatment and time-point for the four modules of interest. (D) Temporal gene expression changes in response to MYB15 overexpression for MYB15-bound genes (as well as MYB14) present in the grapevine reference gene catalogue (v1.1) with a Log2 FC greater than 0,53 or smaller than -0,53 in at least one time-point. Log2 FC values between -0,53 and 0,53 are greyed out. FC, fold change.

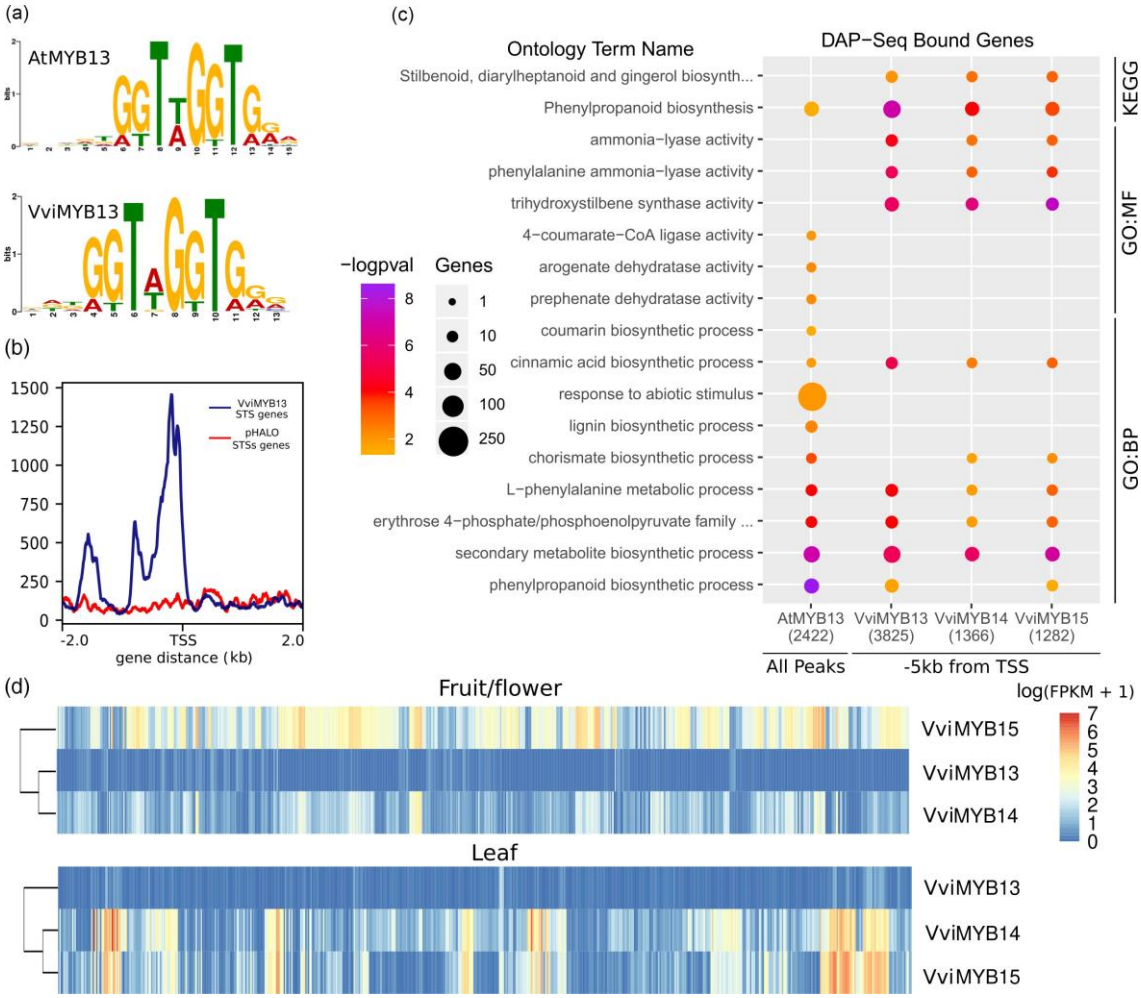


Figure 5.4. VviMYB13-binding features are similar to both VviMYB14 and VviMYB15 despite its different tissue-specific gene expression patterns. (A) AtMYB13 (O'Malley et al., 2016) and VviMYB13 DNA-binding motifs identified using MEME suite. (B) VviMYB13 DNA affinity purification sequencing (DAP-Seq) binding peaks events with respect to 2 kb upstream and downstream of stilbene synthase (STS) transcription start sites. The figure was generated using the Deeptools suite v.3.3.2, computing and normalizing the coverage for each BAM file with a BinSize = 10 and RPKM normalization. RPKM value for each bin is the average between the 10 positions that define each bin. BigWig files for the individual replicates for the transcription factors were merged using bigWigMerge v.2 and bedGraphToBigWig v.4. TSS, transcription start site. (C) Selection of significantly enriched terms from Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) ontologies is shown for AtMYB13 and VviMYB13/14/15 bound genes. A -log10pval scale is provided where a higher value represents a greater statistical significance on a continuous color scale from orange to purple. Number of bound genes intersecting with each ontology term is represented by point size. GO levels are indicated as GO: MF (molecular function) and GO:BP (biological process). (d) A continuous color scale from blue to red represents the ln(x) of FPKM + 1 for VviMYB13/14/15 across different SRA runs for every experiment used to build the fruit/flower and leaf aggregate networks.

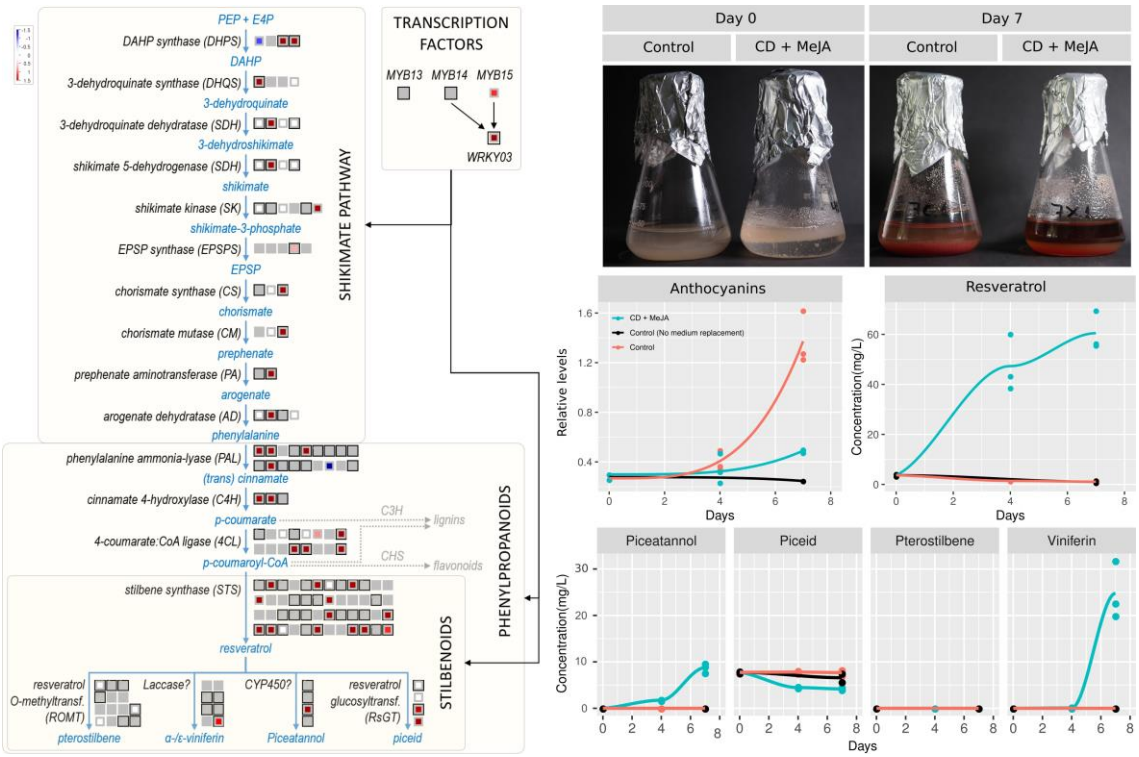


Figure 6.4. MYB13/14/15 DNA binding of shikimate, early phenylpropanoid and stilbenoid pathway genes shown together with microarray gene expression of methyljasmonate and cyclodextrins (MeJA + CD) elicited cells (Almagro et al., 2014). Left panel: genes are ordered from left to right and top to bottom. Genes that are surrounded by a black box correspond to MYB-bound genes (within -10 kb of the transcription start site, the gene body itself, or +2 kb from the end of the gene). Significant (0.05 P-value threshold) positive log2 fold-changes are shown in red and negative changes in blue. Genes with no associated microarray probe are greyed out while white boxes mark genes with no significant differential gene expression. WRKY03 is included as it has been shown to cooperate with R2R3-MYBs for inducing STS gene expression (Vannozzi et al., 2018). Right panel, top: grape cell suspensions at the beginning of the elicitation experiment (Day 0) and 7 days after (Day 7). Bottom: anthocyanin content, measured by spectrophotometry, and stilbenoid quantification by liquid chromatography–mass spectrometry (i.e. resveratrol, piceatannol, pterostilbene, piceid and viniferin in the grape cells at Days 0, 4, and 7. An additional sample, corresponding to non-elicited grape cells subcultured in the old same growth media (i.e. with no sugar replenishment) was taken at Day 7.

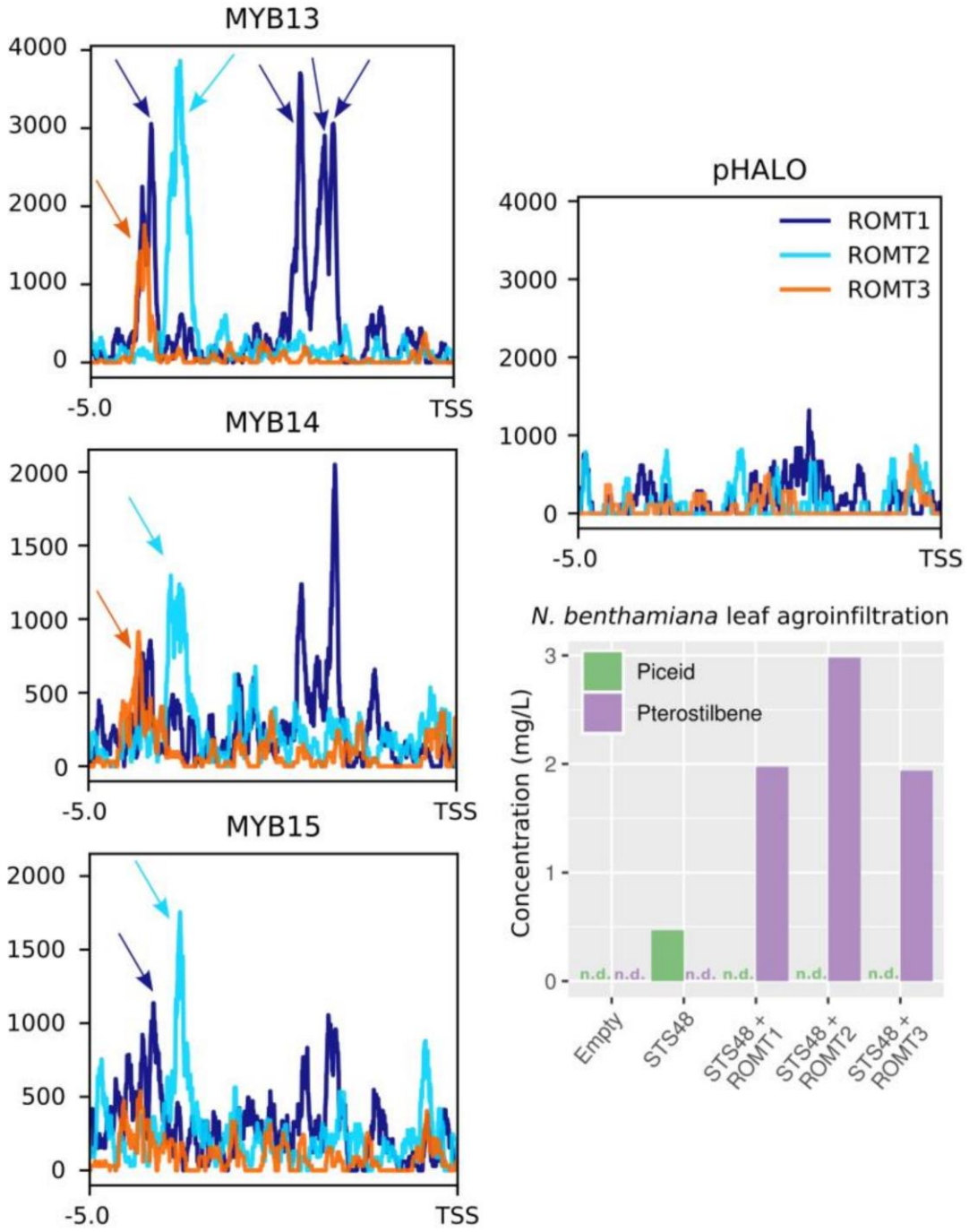


Figure 7.4. MYB13/14/15 DNA-binding events in close proximity to ROMT genes. Peaks are shown in comparison with pHALO up to 5 kb upstream of ROMT1–3 transcription start sites (TSSs). Detected peaks, marked by arrows, are found at about 4.5 and 2 kb upstream of TSSs. Agroinfiltration of both STS48 and ROMT1–3 leads to the accumulation of pterostilbene with respect to the empty vector control. Chromatograms and mass spectrometry-detected transitions for piceid and pterostilbene are found in Figure S12. n.d., non-determined metabolites due to very low concentrations.

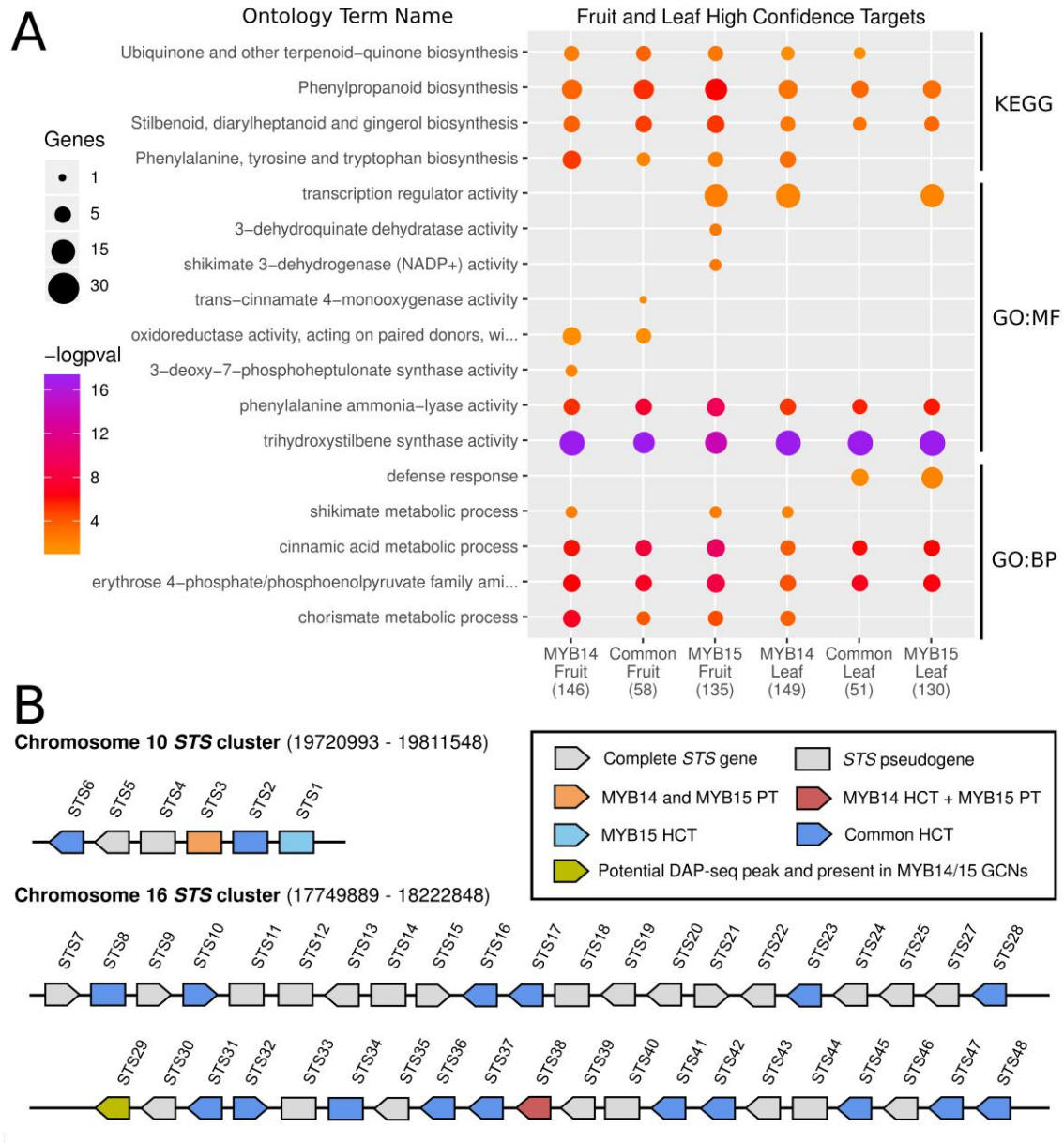


Figure S1. (A) Gene set enrichment analysis of high confidence targets (HCTs) and (B) STS gene synteny

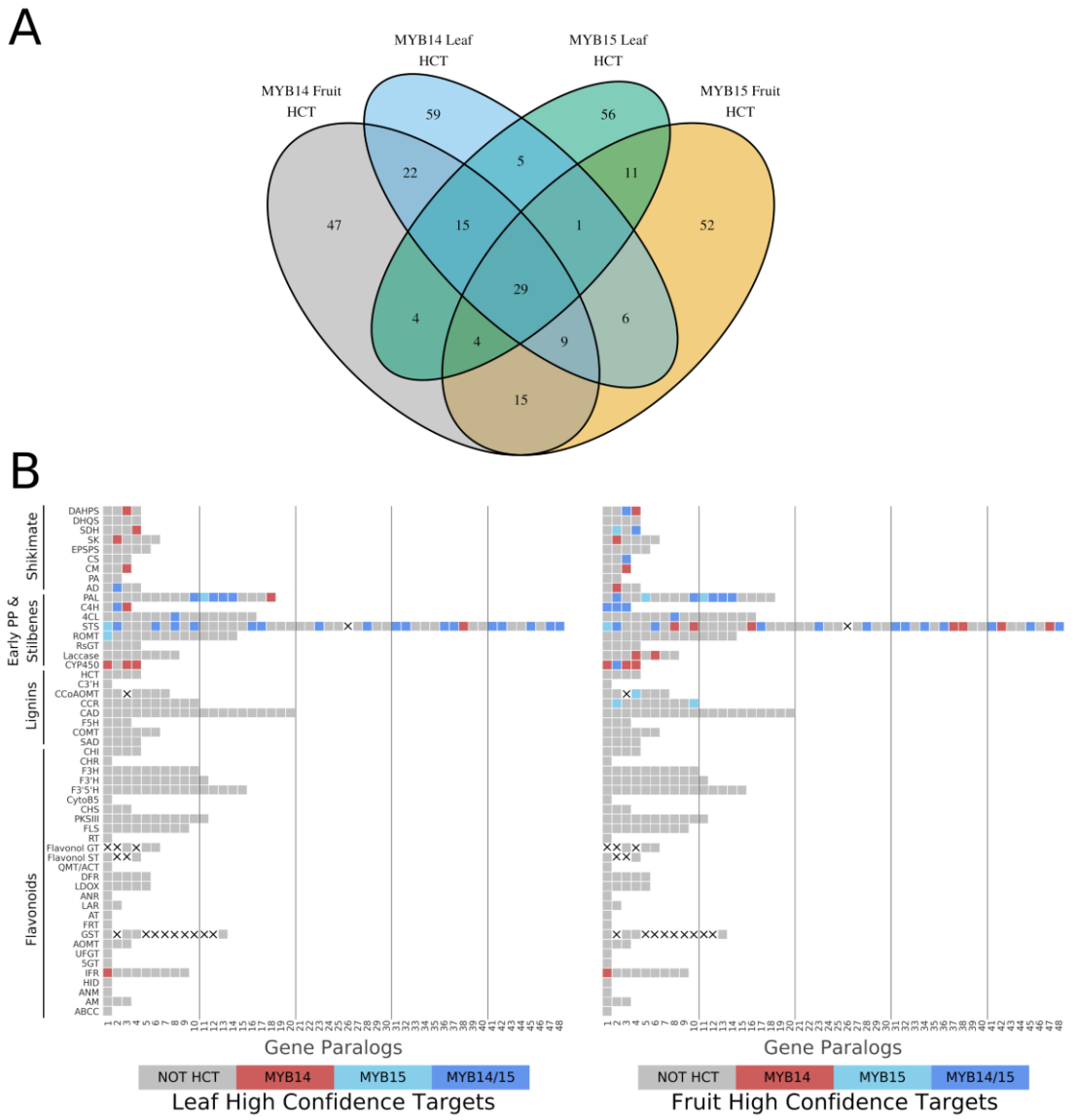


Figure S2. MYB14 and MYB15 high confidence targets (HCTs) among shikimate, phenylpropanoid, and stilbenoid pathway genes.

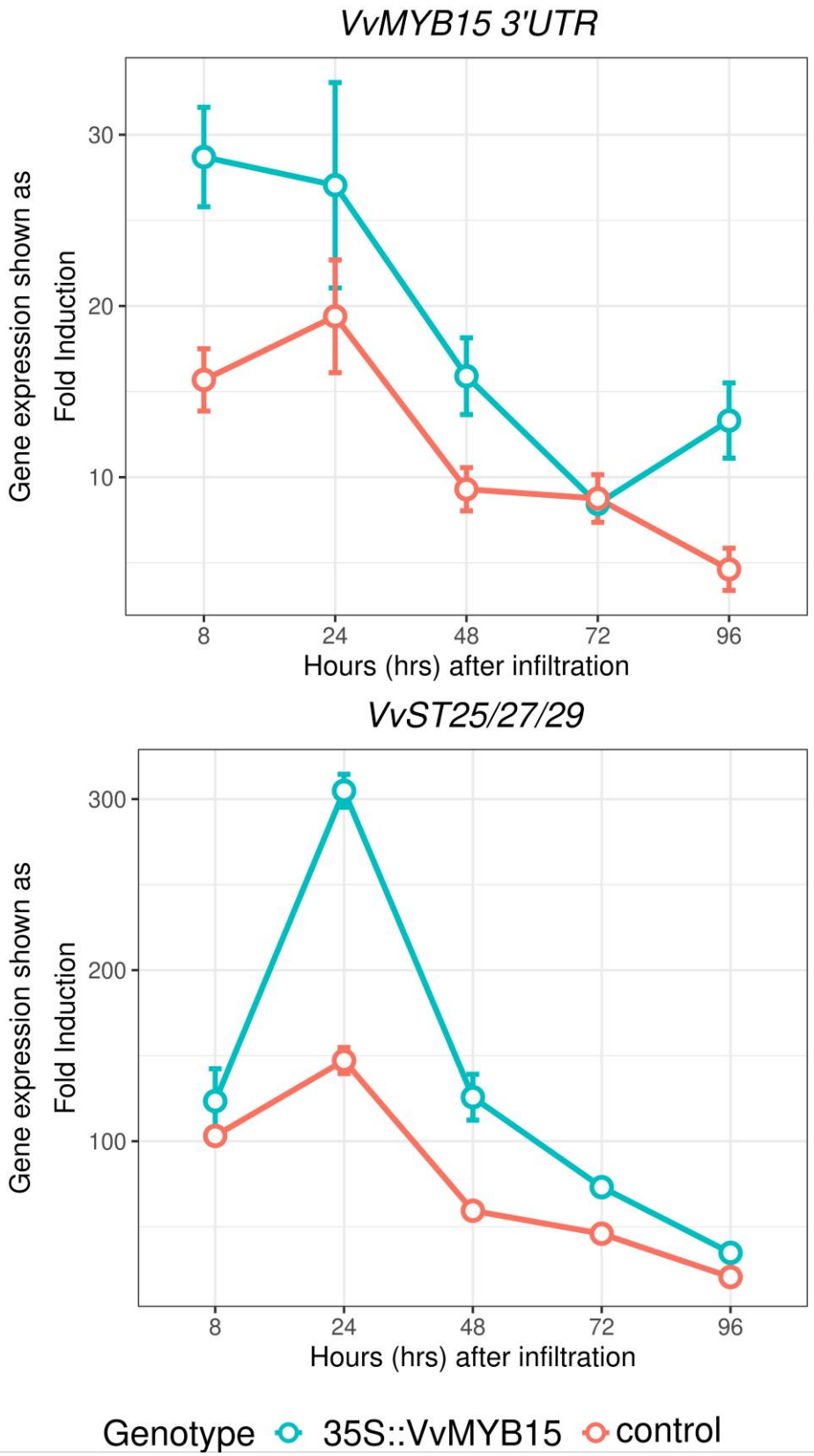


Figure S3. Endogenous MYB15 and STS25/2729 expression in grapevine leaves overexpressing MYB15.

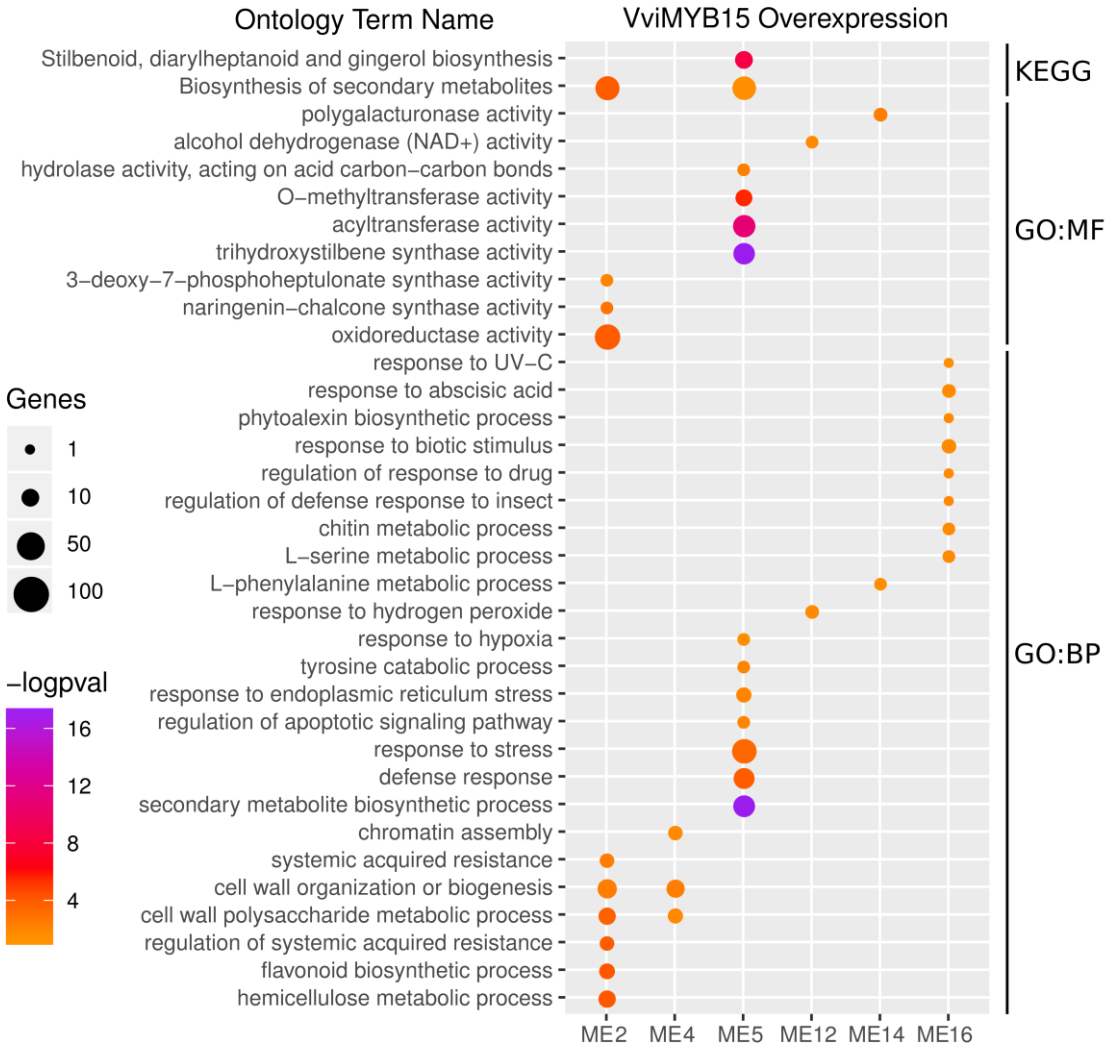


Figure S4. GO and KEGG gene set enrichment analysis for selected WGCNA cluster modules in grapevine leaves overexpressing MYB15.

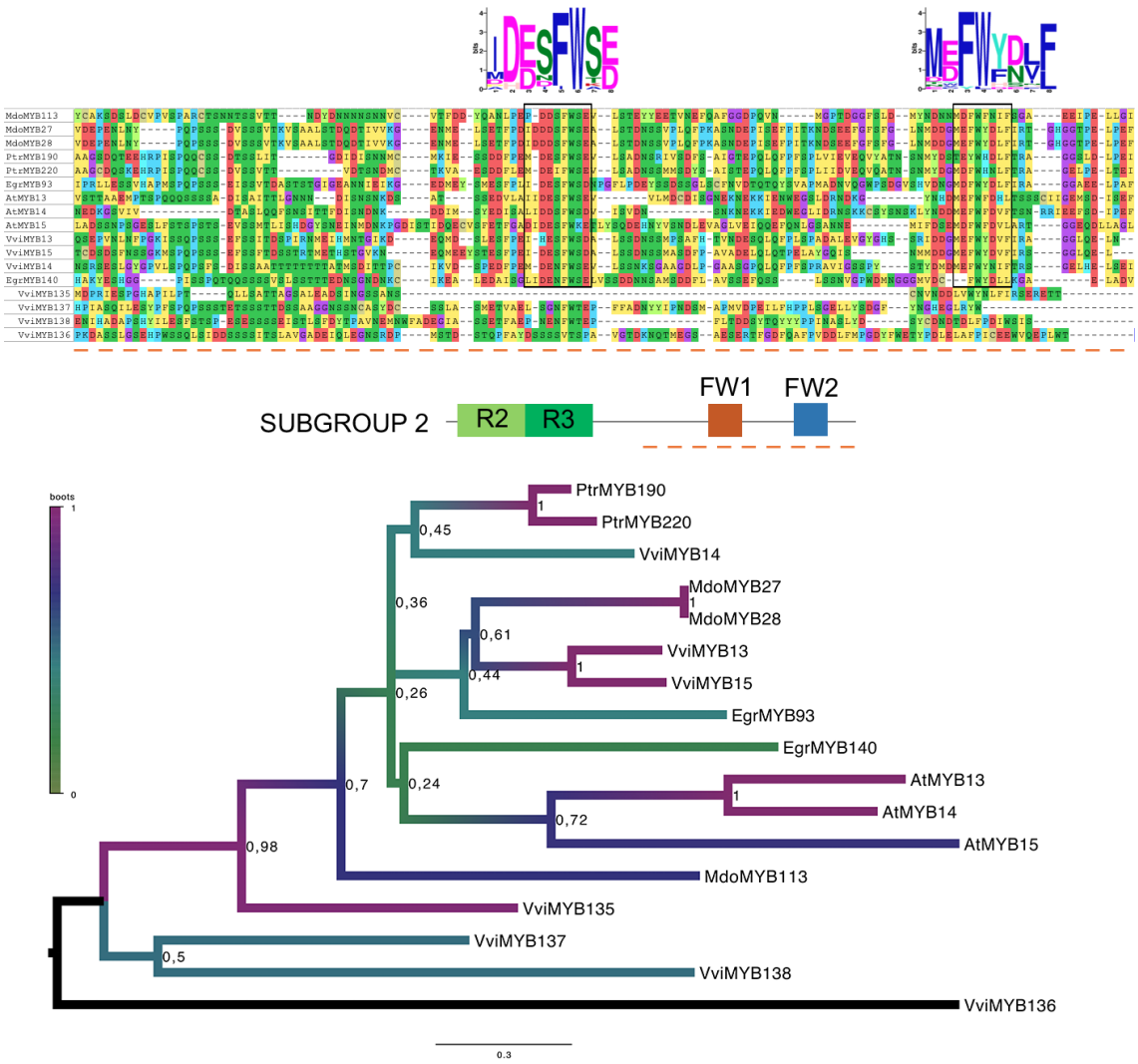


Figure S5. Phylogenetic and protein motif analyses of Subgroup 2 R2R3-MYB proteins and their closest relatives.

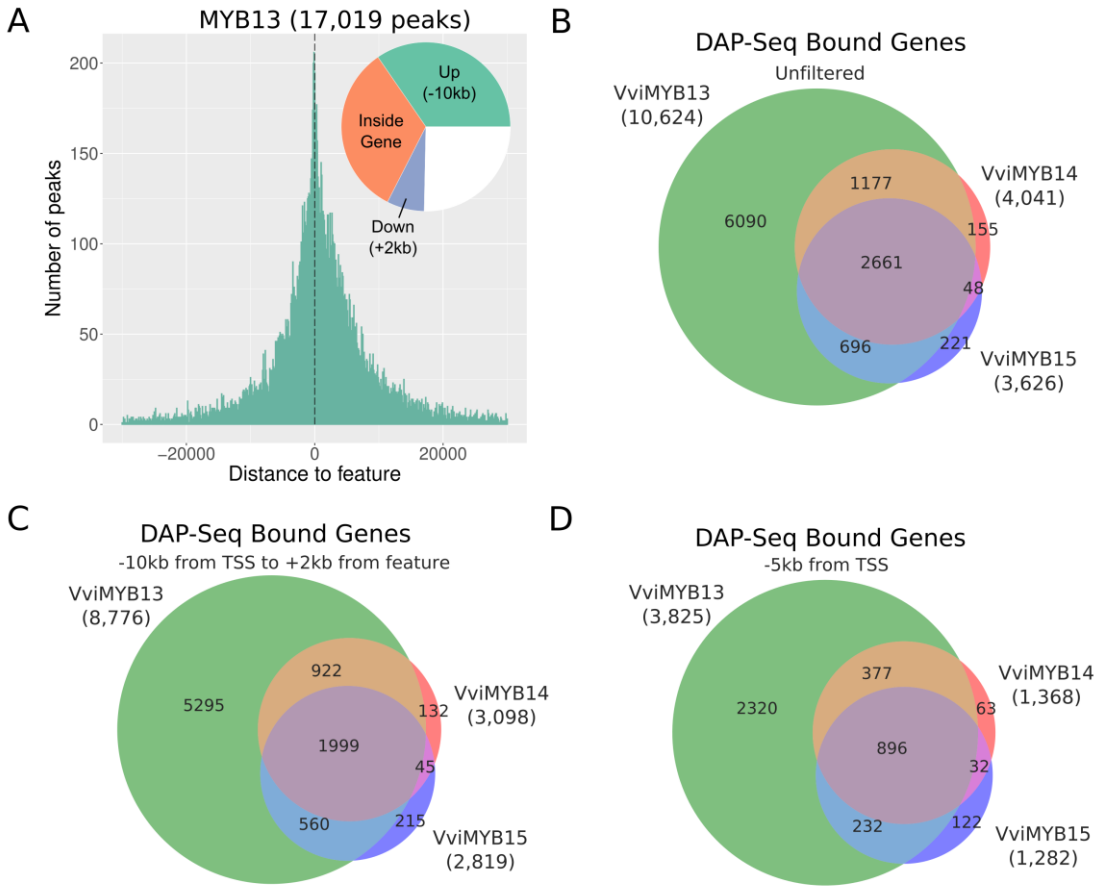


Figure S6. MYB13 DAP-Seq binding peak distribution and overlap of MYB13/14/15 bound genes.

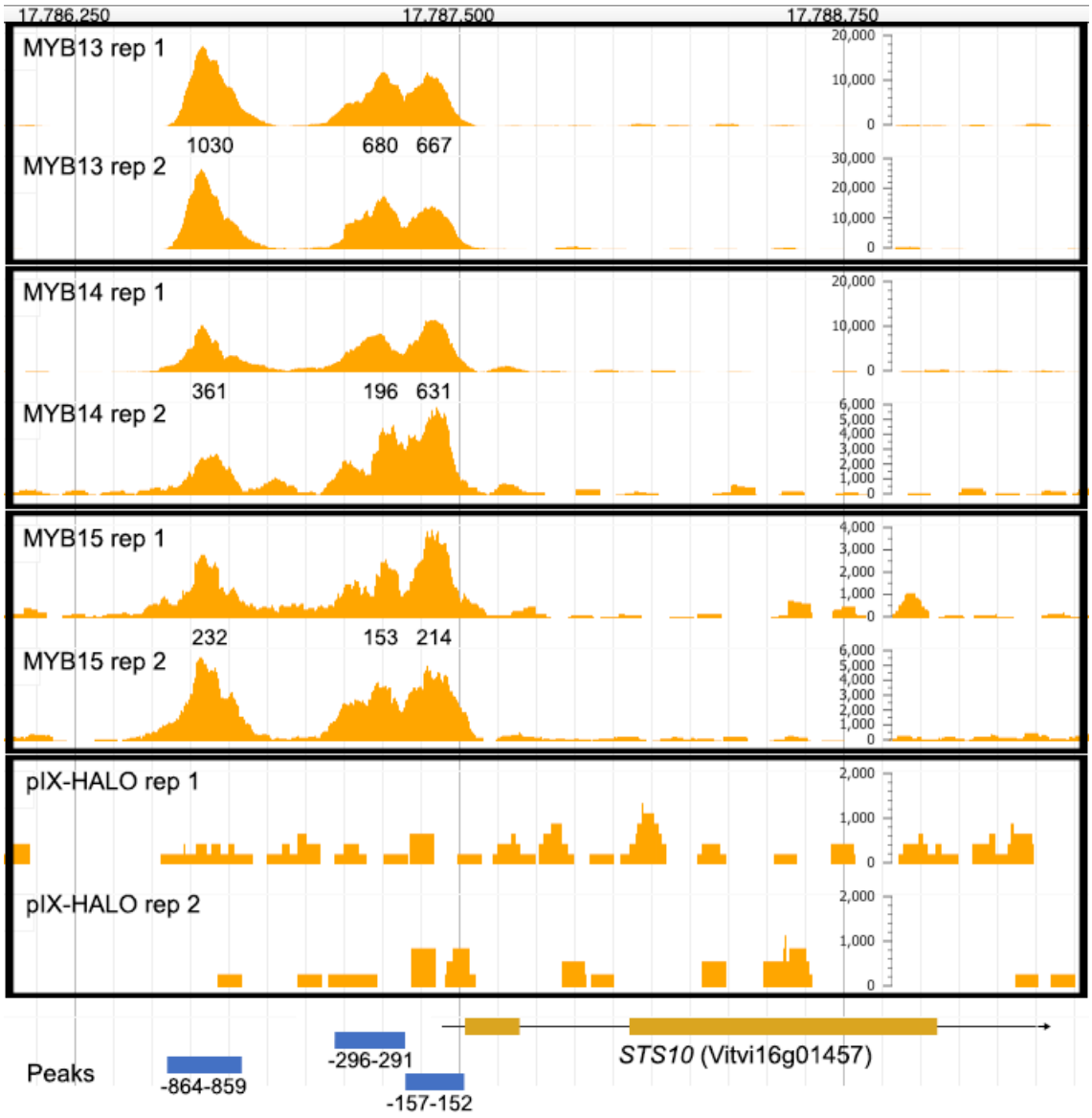


Figure S7. Genome Browser screenshot of MYB13/14/15 DAP-Seq detected peaks for STS10.

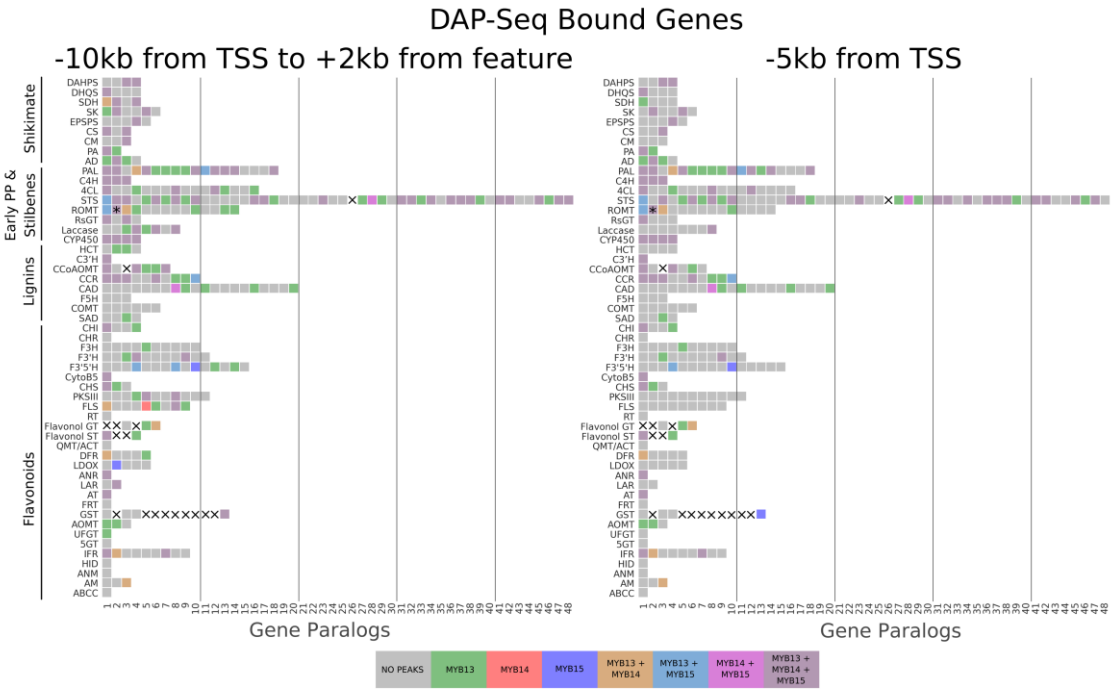


Figure S8. MYB13, MYB14, and MYB15 DNA binding among the shikimate, stilbenoid, lignin, and flavonoid pathways.

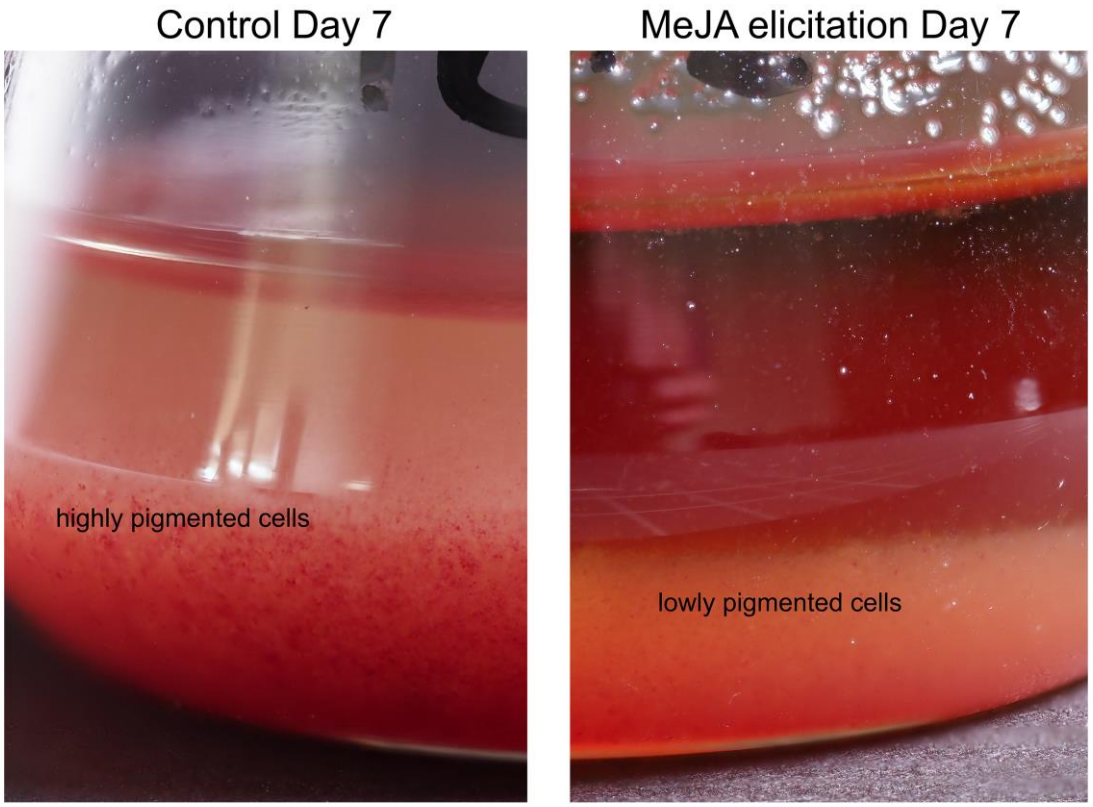


Figure S9. Lack of anthocyanin pigmentation in jasmonate (MeJA) + cyclodextrins (CD)-elicited grapevine cell cultures, compared with control cells at day 7.

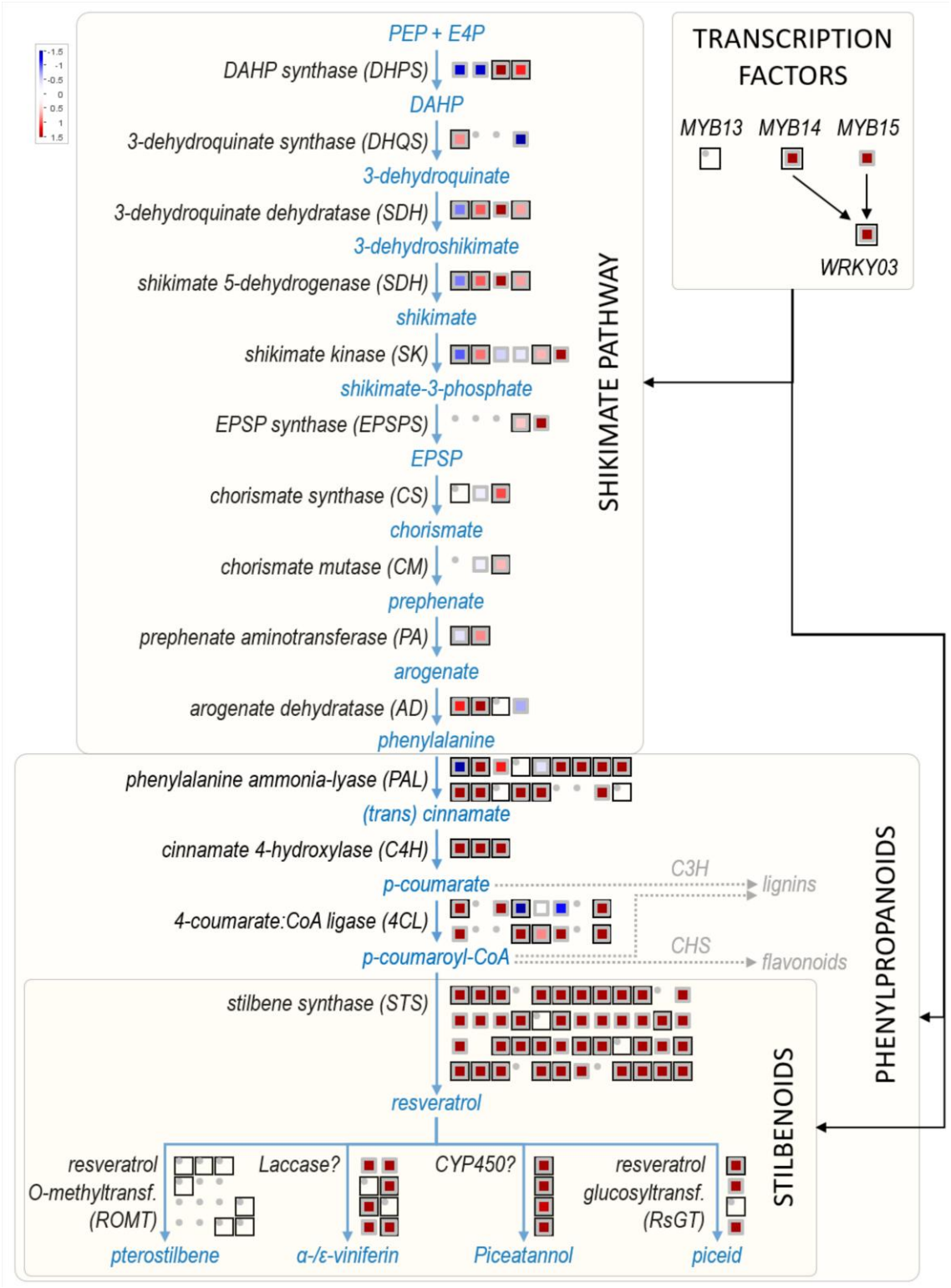


Figure S10. MYB13, MYB14, and MYB15 DNA binding with respect to differential gene expression of shikimate, early phenylpropanoid, and stilbenoid pathway genes during stage 2 noble rot infection of berries, reanalysed from Blanco-Ulate et al. (2015).

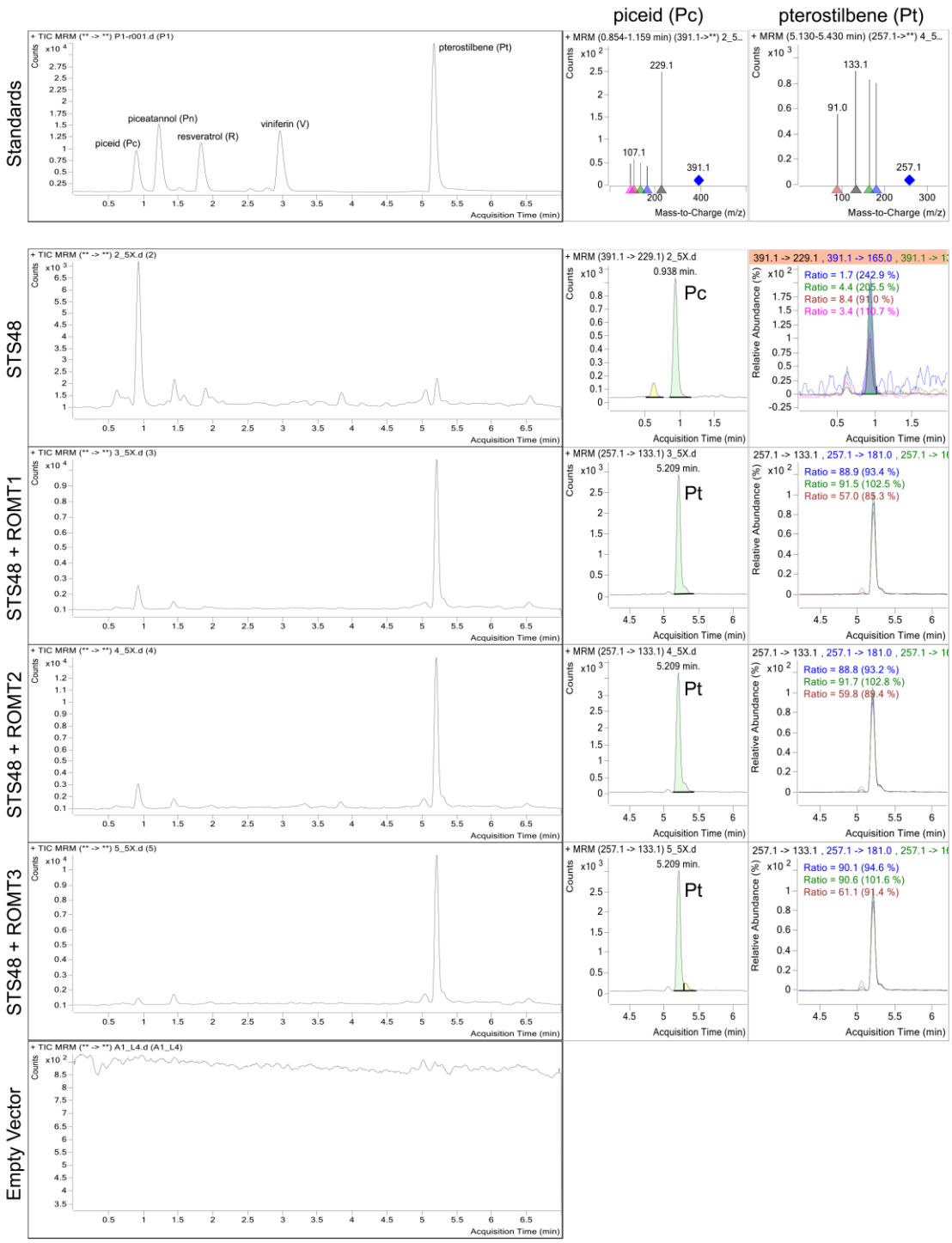


Figure S12. Piceid (Pc) and pterostilbene (Pt) accumulation in tobacco leaves upon overexpression of STS48 and ROMT1–3 genes.

Figure S13. JBrowse adaptation for the visualization of DAP-Seq data in grapevine (DAP-Browse App, available at https:// tomsbiolab.com/vitviz).

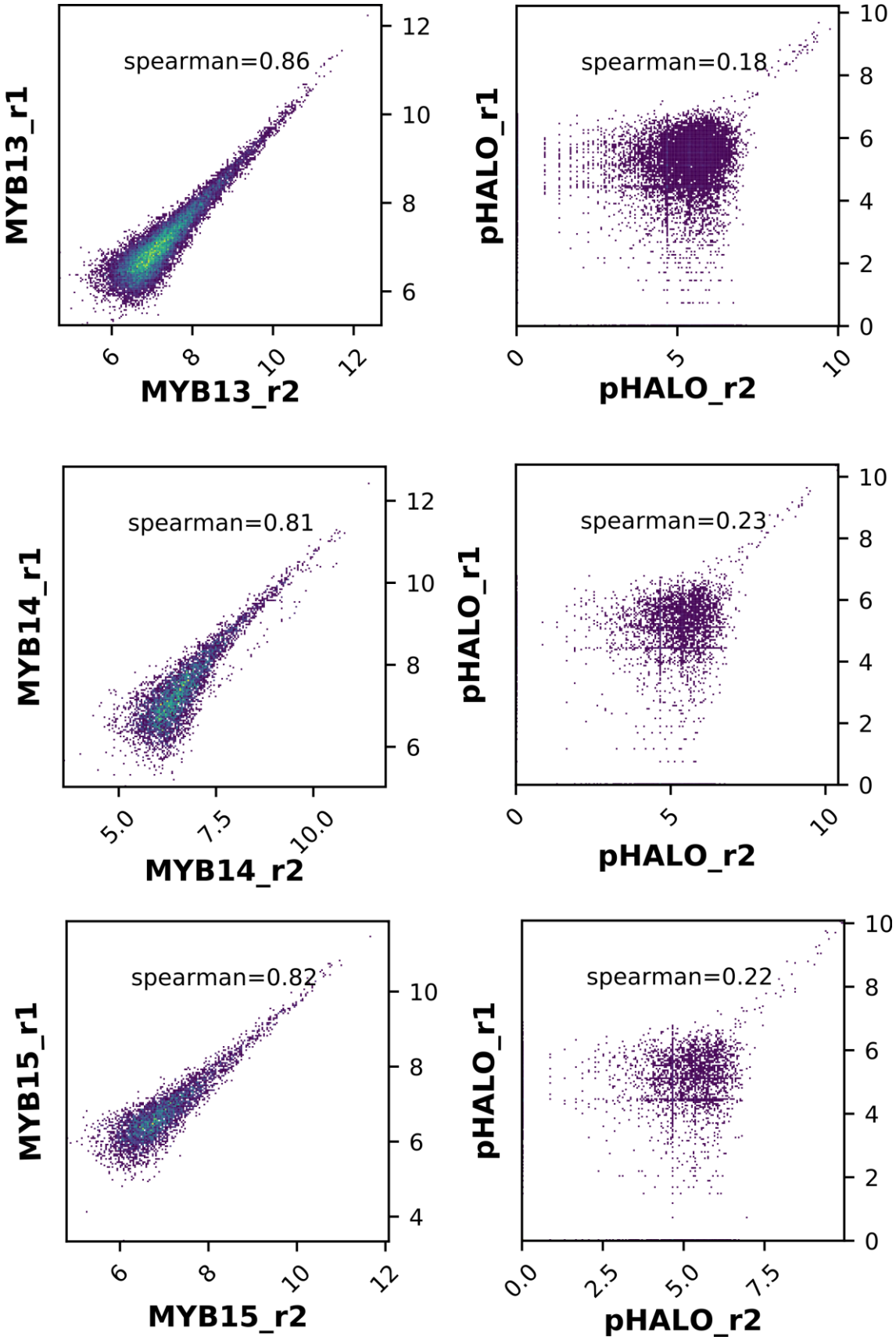


Figure S14. Reproducibility and similarity between DAP-Seq replicates.

4. 4 Discussion

4.4.1 DNA binding and gene co-expression as a proxy for gene regulatory networks

GCNs are based on the 'guilt-by-association' principle whereby correlation in expression implies biological association (Wolfe et al., 2005). This results in a promising tool for predicting gene regulatory networks, in particular to identify gene targets and their regulators. Network aggregation has been shown to improve performance with the use of microarray grapevine datasets (Wong, 2020), and here it has been applied to all the publicly available leaf and fruit RNA-Seq data for grapevine. The resulting network performance, estimated through the AUROC measurement (using MapMan BIN categories) was in our case approximately 0,75, which is an improvement over previously published RNA-Seq aggregate networks (Wong, 2020). Although network performance was high, a co-expression relationship does not necessarily indicate a biological connection (Gillis & Pavlidis, 2012) and the overlap with other datasets of experimental origin is required. This study represents a successful application of genome-wide DNA binding site interrogation in combination with reciprocal GCNs used to draw the gene regulatory networks of TFs controlling specialized metabolism. The three MYBs studied here belong to S2 of the R2R3- MYB subfamily. Some members of this clade have been functionally characterized in other species. For instance, in *A. thaliana* AtMYB13 is involved in response to UVB (Qian et al., 2020), AtMYB14 is related to cold tolerance (Chen et al., 2013) and AtMYB15 is involved in the lignin biosynthesis and immune responses (Kim et al., 2020). In this study, we show that the grape members of this subgroup are all tightly connected to the shikimate pathway and PPP, particularly the early steps of the latter and most of the stilbenoid branch within this pathway. We show that not only STS genes are S2 MYB targets but also resveratrol-modifying enzymes such as glycosyltransferases and O-methyltransferases, characterized here. We also suggest laccases and CYP450 hydroxylases as potential viniferin- and piceatannol-producing enzymes based on DAP-Seq, co-expression networks, and the inspection of large transcriptomic datasets where these genes are expressed in correlation with stilbenoid accumulation. In this work, we show considerable overlap between MYB-related GCNs and their cistrome data, validating the use of co-expression relationships as a predictor of regulation. In addition, many of the HCTs identified by the overlap of these datasets was also confirmed by the transcriptomic analysis of transient overexpressing plants. Most of the statistically significant induced genes upon MYB15 overexpression are indeed HCTs (Fig. 5.4). This highlights the power of combining both DAP-Seq and reciprocal GCN data as a similar biological meaning is drawn from this approach when compared with the overexpression analysis. However, in the

case of MYB13, due to its low expression in both leaf and fruit SRA experiments, we could not generate it, as it would not have been very informative. Beyond target identification, GCNs can also point to MYB-interacting partners such as WRKY03 or WRKY43, which interact with MYB14 to increase STS expression (Vannozzi et al., 2018). Our GCNs also show a high co-expression of S2 MYBs with MYB154 from S14, which has been recently shown to activate a few STS gene promoters in *V. quinquangularis* (Jiang et al., 2021).

4.4.2 Shared and exclusive regulatory features of S2 MYB TFs

MYB13, MYB14, and MYB15 share a high proportion of bound genes as well as identical DNA-binding motifs suggesting a partial overlap of their regulatory roles. This conserved motif has already been described as a MYB *cis*-regulatory element; in particular, the AC-element motif (CACC[T/A]ACC) previously identified for AtMYB15 in *A. thaliana* (Romero et al., 1998). As described in Kelemen et al. (2015), *Arabidopsis* S2 R2R3-MYBs, as well as other subgroups such as S1, S3, S13, and S24, show a high degree of specificity towards the AC-element motif. Within the stilbenoid pathway, we found that MYB-binding patterns, i.e. in terms of their sequence and positions, across individual STS genes were remarkably similar for these three R2R3-MYBs, providing strong evidence that the recent tandem duplications observed for the STS gene family involved upstream regions to a certain length extension. MYB13/14/15 TFs bind in two distinct STS promoter elements; the first at 400 bp from the TSS that is present in most STSs, and a second at 2000 bp, which is only present for a subgroup of them. This observation suggests that STSs may possess gene-specific differences in their regulation. The presence of two binding sites in some of them also offers the possibility of synergistic or dual regulation. MYB TFs regulating other specialized metabolic pathways have been known to interact with other proteins from different TF families such as bHLHs or WDRs as well as other MYBs. According to Holl et al. (2013) MYB15 and MYB14 lack a bHLH-interacting motif within their R3 repeat, which is present in other bHLH-interacting MYBs (e.g. those from S5 and S6) but it is unclear if they have dimerization domains. Further studies could address if MYB14 and MYB15 homo- or heterodimerize. If this is the case, the idea of simultaneous binding at different regions of the promoter gains strength. Our GCNs may help to point out in vivo regulatory differences between TFs with otherwise identical in vitro DNA-binding profiles. This appears to be the case for MYB14/15 as the exclusive presence of 28 STS genes in the MYB14 fruit network points to a very interesting MYB14-specific relationship for a subset of STS genes. This observed contrast in MYB14 and MYB15 co-expressed genes could be explained by promoter differences in MYB14/15 leading to separate regulation by upstream TFs. On the other hand, structural changes between the TFs, leading to distinct interactions with other co-regulators, could also explain this observation. Most STSs are also

MYB14 and MYB15 fruit HCTs. The two previously reported STS targets of MYB14/MYB15 (Holl et al., 2013) were either HCTs (STS41) or present in the MYB14/15 GCNs (STS29). STS29 was not defined as a fruit or leaf HCT as no binding events were automatically called by the peak detection software. However, a potential peak is observed for the three MYBs when examining the mapped sequencing reads in the promoter region of the gene. The enrichment analyses of leaf and fruit HCTs, resulting from the overlap of GCNs and DAP-Seq data, corroborate the newly found connections of MYB13/14/15 with shikimate and early phenylpropanoid genes. Several PAL genes appear as HCTs in addition to STSs, supporting the idea that MYB14 and MYB15 may have the ability to increase and favor the activity of the PPP for the formation of stilbenes, in detriment of lignins or flavonoids. In fact, we may suggest that MYB14/15 might be involved in the repression, direct or not, of chalcone synthases, which are direct competitors of STSs activity. We also suggest that these MYB TFs may rewire the whole metabolic grid of the cell and favor the increase of carbon flux into the PPP by activating the shikimate pathway. Direct DNA-binding observed for at least one isoenzyme of every shikimate and early PPP step highlights the central character of the studied R2R3-MYB regulators in activating stilbenoid metabolism. Despite not being able to perform leaf or fruit GCNs for MYB13, we suggest that these three proteins are highly redundant in their regulatory potential. This redundancy is still partial as GCNs also suggest differences between these two TFs. TFs act in a hierarchical manner and based on our data, MYB14 and MYB15 appear to be near the top of this hierarchy, not only because they are able to control a large portion of the shikimate/early phenylpropanoid and stilbene pathways but also because approximately 21% of their targets are TFs, respectively. In fact, at least nine WRKY TF genes are present among the common MYB14/15 HCTs. WRKY TFs have been suggested as MYB co-regulators in the control of STS expression (Wang et al., 2020; Xi et al., 2014; Xu et al., 2019) and in many defense responses to pathogens. In particular, WRKY03 is bound by both TFs and present in MYB14 and MYB15 leaf and fruit GCNs. This observation adds one more layer of complexity to the hierarchical regulation of MYB14, which requires WRKY03 to increase STS gene expression (Vannozzi et al., 2018). The data presented here suggest that grapevine S2 members act mainly as positive regulators of transcription; however, based on the reduction of gene expression of some bound genes upon MYB15 induction, S2 MYBs could also be involved in negative regulation possibly requiring other co-repressors. An alternative hypothesis could be that MYB15 may be competing for promoter space with other stronger transcriptional activators that recognize the same target motifs, as suggested by Tamagnone et al. (1998). An interesting example of this repression is observed among circadian rhythm-related genes. Given the known circadian behavior of STS gene expression (Carbonell-Bejerano et al., 2014), we further explored the potential relationship of

MYB15 with the core molecular components of the circadian clock. MYB15 leaf overexpression led to a strong downregulation of LHY (Dataset S3), one of the two components of the main oscillator within the system. Moreover, different components of the clock such as LHY itself, PPR7a/b, TOC-like, GI, and ELF4 are bound by either of MYB13/14/15 (Dataset S1), which further suggest a direct connection with the circadian rhythm. RVE8-type MYBs have been previously shown to have roles regarding the control of the circadian rhythm in *A. thaliana* (Shalit-Kaneh et al., 2018); however, the involvement of R2R3-MYBs has not been previously described.

4.4.3 Integration of DAP-Seq and expression data as a tool for the identification of putative secondary metabolic pathway enzymes

By overlapping the different datasets generated in this work, we were able to identify putative stilbenoid genes among MYB targets, such as laccases (i.e. multicopper oxidases potentially involved in resveratrol oligomerization), resveratrol hydroxylases, resveratrol O-methyltransferases (ROMTs) and resveratrol glycosyltransferases (RsGTs). Within O-methyltransferases, ROMT1 has been the only gene described to date as involved in the synthesis of pterostilbene (Schmidlin et al., 2008), and here we show that MYB13 and MYB15 bind within its 5-kb upstream region. In addition, we show that the three S2 MYBs bind in a similar region with respect to ROMT2 and ROMT10 promoters. Despite none of these ROMTs appearing in MYB14 or MYB15 fruit GCNs, some of them do appear in leaf GCNs. This is in line with the lack of expression of ROMT genes in fruit or flower tissues (Figure S11), highlighting the importance of tissue-specificity when dealing with coexpression networks. In fact, the transient overexpression of MYB15 led to an upregulation of all ROMT1-6 enzymes recognized by the microarray probes. Clustering of microarray data (Figure 4) also showed that the expression profiles of the ROMT enzymes was very similar to MYB15 and hence the gene set enrichment analysis of cluster module ME5, where MYB15 is found, showed enrichment of the O-methyltransferase activity ontology term. We also identified cytochrome P450 and laccase (multicopper oxidases) among MYB targets. Both of these enzyme classes are suggested to take part in the hydroxylation and oligomerization of resveratrol. This has not been shown experimentally except for the production of piceatannol by a transiently expressed human cytochrome P450 in *Nicotiana bethamiana* (Martınez-Marquez et al., 2016). We identify particular genes from these two families, which are highly induced in stilbenoid accumulating conditions such as *Botrytis* infection (e.g. at Stage 2 of the infection; Figure S10) as well as being present in S2 MYB GCNs or HCT lists. Nevertheless, further studies are required for their functional demonstration. The approach of integrating DAP-Seq data and reciprocal GCNs is particularly relevant in non-model species; however, it needs sufficient public transcriptomic

data for a given species. Online tools for the inspection of DAP-Seq data would certainly help in the identification of novel TF targets for their selection and further characterization. Therefore, we developed a public grapevine DAP-Seq visualization application named DAP-Browse, present in the Vitis Visualization Platform (Figure S13; VitViz; available at https://tomsbiolab.com/vitviz).

4.5 Experimental Procedures

4.5.1 DNA affinity purification sequencing

Genomic DNA (gDNA) was purified from young grapevine leaves of cv. 'Pinot Noir' (clone PN-94) according to Chin et al. (2016). gDNA library construction, DAP-Seq, peak calling, and motif analysis were performed as described in Bartlett et al. (2017) with the following modifications. The gDNA samples (800 ng/library) were sonicated into 200-bp fragments on a focused ultrasonicator (Covaris, Woburn, MA, USA). The gDNA library then underwent end-repair, A-tailing, and adapter ligation for subsequent Illumina-compatible sequencing. Successful library construction was verified by gel electrophoresis of sonicated gDNA and qPCR of adapter ligated fragments. MYB14, MYB15, and MYB13 were amplified from cv. 'Pinot Noir' and cloned into the pIX-HALO expression plasmid (TAIR Vector: 6530264275) with the HALO-tag at the N-terminal. Clones were verified by XhoI digestion. The MYB-expression vectors were used in a coupled transcription/translation system (Promega, Madison, WI, USA) to produce MYB Halo tagged proteins. Protein production was confirmed by western blot using an anti-HaloTag antibody (Promega). HaloTag-ligand conjugated magnetic beads (Promega) were used to pull down the HaloTag-fused TFs. Pulled-down TFs were exposed to DAP-Seq gDNA libraries for MYB-DNA binding. Four hundred nanograms of gDNA library was used per DAP-Seq reaction. The bound DNA was then eluted and sequencing libraries were generated by PCR amplification. Sequencing was carried out using Illumina NextSeq 500 set up to obtain 30 million 1 x 75 bp single end reads. An empty pIX-HALO expression vector was used as a negative control to account for non-specific DNA binding as well as copy number variation at specific genomic loci. Two replicates (individual and independent gDNA-TF interaction samples) were used for all experiments including the control. The latest published genome assembly for *V. vinifera* cv. 'PN40024' is the 12X.v2, which is associated to V1 and VCost.v3 annotations, containing 29.971 and 42.414 gene models, respectively (Canaguier et al., 2017; Jaillon et al., 2007). Although the VCost.v3 annotation has been manually curated for the STS family, it also contains non-protein coding genes (such as lncRNAs and miRNAs) most of which have only been automatically annotated. As the main focus of this work was to study the regulation of protein-coding genes in secondary metabolism (stilbenoid pathway in particular), a new annotation file combining all V1 gene models with the updated STS VCost.v3 gene models was created. The merged gff3 annotation file is available at http:// tomsbiolab.com/scriptsandfiles. DAP-Seq reads were mapped to the 'PN40024' 12X.v2 reference genome using bowtie2 (Langmead & Salzberg, 2012), version 2.0- beta7, with default parameters and post-processing to remove reads that have MAPQ scores lower than 30. Peak detection was performed using GEM peak caller (Guo et al., 2012) version 3.4 with the 12X.v2

genome assembly using the following parameters: '-q 1 -t 1 -k_min 6 -kmax 20 -k_seqs 600 -k_neg_dinu_shuffle', limited to nuclear chromosomes. The replicates were analyzed as multi-replicates with the GEM replicate mode. Peak summits called by GEM were associated with the closest gene model in the custom annotation file using the BioConductor package ChIPpeakAnno (Zhu et al., 2010) with default parameters (i.e. NearestLocation). De novo motif discovery was performed by retrieving 200- bp sequences centered at GEM-identified binding events for the 600 most enriched peaks and running the meme-chip tool (Machanick & Bailey, 2011) in MEME suite 4.10.1 with default parameters. Gene set enrichment analysis were conducted for the three analyzed MYB-bound gene lists (Dataset S1). DAP-Seq replicate reproducibility was assessed using Deeptools suite v.3.3.2 (Fig. S14). Briefly, for each TF, genomic regions where peaks had been detected were extracted and divided into bins. The DAP-Seq binding 'signal' (read counts normalized by library size) was extracted in all the bins and averaged to create an overall binding signal for each peak. This was done independently for replicate 1 (y-axis) and replicate 2 (x-axis) for all peaks, computing the Spearman correlation coefficient for each TF. This analysis allowed us to demonstrate the high reproducibility between TF replicates, with correlations over 0,8. We computed the same plots for the pIXHALO replicates across the peaks detected for each TF, resulting in low correlation, as expected, because the reads sequenced in these experiments are mostly random noise. We thus corroborate the robustness of the results derived from using two replicates, as in shown in O'Malley et al. (2016) (Figure S1B), improved compared with ChIP-seq, where replicate correlation ranges between 0.6 and 0.7).

4.5.2 Generation of condition-dependent aggregate whole genome co-expression networks and extraction of gene-centered co-expression networks

Transcriptomic RNA-Seq SRA studies (Illumina sequencing) from fruit/flower and leaf tissue samples were downloaded. SRA studies were manually inspected and filtered to keep those that were correctly annotated and contained four or more data sets (i.e. runs) while also excluding SRA studies concerning microRNAs, small RNAs and non-coding RNAs. In total, 35 and 42 SRA studies were obtained, encompassing 807 and 670 runs from fruit/flower and leaf samples, respectively (Dataset S2). Reads were trimmed with fastp (Chen et al., 2018), version 0.20.0 with the following parameters: '-detect_adapter_for_pe -n_ base_limit 5 cut_front_window_size 1 cut_front_mean_quality 30 - cut_front cut_tail_window_size 1 cut_tail_mean_quality 30 - cut_tail -l 20'. After trimming, the runs are aligned with STAR (Dobin et al., 2012), version 2.7.3a, using default parameters. Raw counts were computed using FeatureCounts (Liao et al., 2013), version 2.0.0, and the VCost.v3_27.gff3 gene models (available at http:// www.integrape.eu/). Each SRA study was analyzed individually to build a highest reciprocal rank (HRR) matrix (Mutwil

et al., 2011). Each raw counts matrix was normalized to FPKMs, and genes with less than 0.5 FPKMs in every run of the SRA study were removed. The Pearson's correlation coefficients of each gene against the remaining genes was then calculated for each SRA study (across all run matrices) and ranked in descending order. Ranked Pearson's correlation coefficient values were used to compute HRRs among the top 420 ranked genes (420 roughly equals 1% of all VCost.v3 gene models), using the following formula: HRR(A,B) = max(rank (A,B), rank(B,A)), generating a HRR matrix for each SRA study. To construct the aggregate whole genome co-expression network, the frequency of co-expression interaction(s) across individual HRR matrices was used as edge weights, and after ranking in descending order, the top 420 frequency values for each gene were chosen to build the final aggregate networks. The top 420 most highly co-expressed genes for any gene of interest, was used to generate individual GCNs. Network functional connectivity (i.e. performance) across all given annotations and genes was assessed as in Wong (2020) by neighbor-voting, a machine learning algorithm based on the guilt-by-association principle, which states that genes sharing common functions are often co-ordinately regulated across multiple experiments (Verleyen et al., 2014). The evaluation was performed using the EGAD R package (Ballouz et al., 2016) with default settings. The network was scored by the AUROC across MapMan V4 BIN functional categories associated to the VCost.v3 annotation (Dataset S2) using threefold cross-validation. MapMan BIN ontology annotations were limited to groups containing 20– 1000 genes to ensure robustness and stable performance when using the neighbor-voting algorithm.

4.5.3 Prediction of HCTs in grape reproductive and vegetative organs

To identify MYB14 and MYB15 HCTs, TF-bound genes identified with DAP-Seq were overlapped with data extracted from the aggregate whole genome co-expression network derived from fruit/flower and leaf transcriptomic data. Individual GCNs were extracted from the whole genome network for each MYB14/15 bound gene and for both MYBs to check if a particular MYB-to-target gene pair had at least one co-expression relationship. This relationship was considered positive if either the DNA bound gene was present in the MYB GCN or the respective MYB gene was present in the GCN of the candidate gene. Thus, bound genes for each MYB TF with a positive co-expression relationship with their respective MYB TF were considered as fruit/flower or leaf HCTs. GO and KEGG enrichment analyses for HCTs were performed using the gprofiler2 R package (Kolberg et al., 2020) with default settings. A significance threshold of 0,05 was chosen for P-values adjusted with the Benjamini–Hochberg correction procedure (Benjamini & Hochberg, 1995).

4.5.4 Transient MYB15 overexpression, microarray time-course, and transpiceid quantification in leaves

Vitis vinifera cv. 'Shiraz' leaves from 10-week-old plants were transformed via *Agrobacterium*-mediated infiltration as in Merz et al. (2015) with MYB15 (35S::VviMYB15) or an empty vector control. A green fluorescent protein vector (35S, GFP-GUS) was used in each agroinfiltration to estimate transformation efficiency. Samples were subsequently collected at five different time-points (8, 24, 48, 72, and 96 h) in separate replicate pools (L1 and L2) consisting of three leaves from independent plants per time-point (n = 2 biological replicates). Microarray analysis, using the Affymetrix V. vinifera Grape Array (A-AFFY-78), was carried out on extracted leaf RNA from three of the time-points described for both control and MYB15 infiltrated samples. Microarray data are available at Dataset S3 (Microarray dataset sheet). Affymetrix probe-to-gene associations were reassigned as here explained. Affymetrix probe sequences (originally associated to V1) were realigned using bowtie with default parameters against the 'PN40024' transcriptome derived from the VCost.v3 annotation to improve V1 probe-to-gene assignments (Dataset S3). As each Affymetrix probe has a different number of sequences associated to it (ranging from 8 to 20), a probe-to-gene association was accepted when more than 40% of the probe's sequences matched the same gene with no mismatches. The percentage was increased to 60% for hits with one mismatch, 70% with two mismatches and 80% with three mismatches. From, in total, 16 602 probes, 59.8% were assigned to a VCost.v3 gene model. The LIMMA R package (Ritchie et al., 2015) was used for RMA normalization of fluorescence values after which they were log2- transformed. Normalized probes were analyzed with the NOISeq R package (Tarazona et al., 2011). Briefly, batch effect was corrected using the ARSyNseq function with default parameters, and then differentially expressed genes were obtained comparing the two MYB15 overexpression lines against the two control lines for each time-point using the NOISeqBIO function. Genes were considered as differentially expressed when posterior probability >0,95 and ¦log2FC¦>0.53 in at least one time-point, obtaining 2297 probes corresponding to differentially expressed genes. RMA values of the filtered probes were clustered applying the weighted GCN analysis (WGCNA) R package (Langfelder & Horvath, 2008) generating 18 modules for the signed network. Briefly, WGCNA R package blockwiseModules function was used with a soft-threshold power value of 18 (as no free-scale topology was reached), a deepSplit of 4 and a mergeCutHeight of 0,1, obtaining, in total, 18 modules for the signed network (i.e. considering the sign of correlation). Microarray analysis results are available at Dataset S3 (MA analysis sheet). The pheatmap R package was used to represent the modules using the calculated Z-scores across samples. The probe-assigned genes in each module were used for GO and KEGG enrichment analyses as described for HCTs. Quantitative expression

analysis of MYB15 or STS genes and trans-piceid quantification were conducted using L2 leaf pools across the different samples. Trans-piceid was quantified on a reverse phase high-performance LC (HPLC) instrument (Kronton, Reno, NV, USA). Quantitative PCR was run in triplicate per time-point and treatment using primers against the endogenous (aligning to the UTR) and transgenic MYB15, STS25/27/29, and STS41/45. Transcript levels were corrected using the housekeeping gene Ubiquitin1. All used primers are shown in Dataset S3. The fold of induction reported for both microRNAs and metabolites were calculated relative to background levels in non-infiltrated leaves.

4.4.5 Grapevine cell culture elicitation time-course

MeJA-CD elicitation of cell suspensions was carried out as described in Bru et al. (2006) with the following modifications. Liquid cell suspensions were initiated by transferring solid calli into 250-ml flasks containing liquid culture media (with Gamborg basal salts and Morel vitamins) described in Bru et al. (2006) and grown in an orbital shaker at 120 rpm. Liquid cell cultures were maintained in dark conditions for 2 months, with subcultures being performed every 2 weeks by moving half of the extracellular volume to a new flask and restoring the original volume with fresh media (half medium replenishment). For the elicitation time-course, four flasks of cell suspensions were mixed together, and three-quarters of the suspension was filtered using a sterile glass filter (90– 150 lm pore size) and a vacuum filtration system to separate the cells from the media. In total, 15 100 ml flasks were prepared with three biological replicates for each time-point and treatment. Day 0 cells were immediately collected after filtering and frozen in liquid nitrogen. For day 4 and day 7 samples, 8 g of filtered cells and 32 g of new liquid media were added in each flask (complete medium replenishment). Treated (elicited) cells were grown in the liquid media with 50 mM methyl-beta-cyclodextrin (CAS no. 128446-36-6) and 84 ll of 47.5 mM MeJA dissolved in 100% MeOH while untreated (control) cell suspensions were grown in liquid media with 84 ll of 100% MeOH. The rest of the original suspension (one-quarter) was subcultured as usual and kept growing at dark for being used as an additional control for half medium replenishment at day 7. All the biological replicates were grown by orbital shaking at 120 rpm in the conditions described above. Cell cultures were sampled at their corresponding time-point, filtered using a vacuum-aided system, and then washed with cold sterile water before freezing liquid nitrogen. Cells were lyophilized at -52°C and 0,63 mbar for 48 h and then transferred to -80°C.

4.4.6 STS and ROMT overexpression in tobacco leaves

ROMT2-3 were amplified using TOPO-D (Invitrogen, Waltham, MA, USA) compatible primers from cDNA derived from grapevine senescent leaves. RNA extraction and cDNA synthesis were performed using the SpectrumTM Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO, USA) and NZY First-Strand cDNA Synthesis Kit, respectively. The final ROMT2/3-containing pENTR/D-TOPO plasmids were verified by Sanger sequencing. ROMT2 and ROMT3 were transferred into the binary destination vector pB2GW7 including Cauliflower mosaic virus (CaMV) 35S promoter through Gateway LR clonase recombination (Invitrogen). Together with ROMT1 and STS48 constructs (Santos-Rosa et al., 2008; Schmidlin et al., 2008) all these were chemically transformed in *Agrobacterium* tumefaciens strain C58 for subsequent agroinfiltration of *N. benthamiana* leaves. The bacterial suspension was kept at an OD600 of 0,2 for each construct before infiltration of leaves on their abaxial side (three leaves of one plant for each experimental combination) using a 1 ml syringe without a needle. Agroinfiltration was conducted with the individual empty vector-, STS-, or combined STS- and ROMT-containing bacterial cultures. After 72 h, leaf-tissue samples were collected by dissection to avoid leaf nerves. Collected samples were frozen at -80°C.

4.6.7 Stilbenoid and anthocyanin quantification of grape cell cultures and *Nicotiana* agroinfiltrated leaves

Stilbenoid and anthocyanin quantification of grape cell cultures and *Nicotiana* agroinfiltrated leaves was performed as in Martınez-Marquez et al. (2016) and Nakata and Ohme-Takagi (2014), respectively, with modifications. Stilbenoid metabolites were extracted using 10 mg of cells in 1 ml of 80% methanol with 220 rpm overnight shaking at 4°C. The extracted solution was centrifuged at 14 000 g for 10 min and collected in another tube for HPLC–MS detection and quantification, as described in Hurtado-Gaitan et al. (2017). For the case of tobacco samples, leaves were weighed and grounded in liquid nitrogen. Extraction using 100% MeOH was carried out keeping a ratio of 8 ml of MeOH per 1 g of fresh tissue with 220 rpm overnight shaking at 4°C as described in Martınez-Marquez et al. (2016). After centrifuging at 14 000 g for 10 min the supernatant was used for HPLC–MS determination and quantification, as described in Hurtado-Gaitan et al. (2017). Anthocyanin quantification was performed as described in Nakata and Ohme-Takagi (2014). Briefly, lyophilized cells (0,2 ml) were treated with 1 ml of extraction buffer (45% MeOH and 5% HCL). After a 10 min vortex and centrifugation at 12 000 g for 5 min, the supernatant was collected into a new tube and centrifuged again under the same conditions. Supernatant absorbance was measured at 530 nm and 637 nm and the anthocyanin relative levels were calculated using the formula: (Abs530 – (0.25 9 Abs637)) 9 5/(mg of cells).

4.6.8 Re-annotation of the grapevine MapMan ontology and representation of previously published microarray and RNA-Seq data

MapMan v4 (Schwacke et al., 2019) annotation was generated based on EGAD, resulting in high-quality annotations for 13.318 of 41.413 VCost.v3 genes (32%) (Dataset S3). A pathway image for shikimate, phenylpropanoid, and stilbenoid pathway was created and annotated using the MapMan tool (Thimm et al., 2004) for both Vcost.v3 and V1 of the grapevine genome annotations. This image includes the DAP-Seq results whether a gene bound by MYB14, MYB15, and MYB13 is shown in black squares. All files are available for download from the GoMapMan exports page (Ramsak et al., 2014) (gomapman.nib.si/export). Significant log2 fold-change values from a microarray study involving jasmonate elicitation of cv. 'Gamay Freaux' cell cultures (Almagro et al., 2014) and from an RNA-Seq study of *Botrytis*-infected berries of cv. 'Semillon' (Blanco-Ulate et al., 2015) were represented within the newly generated MapMan pathway image.

ACKNOWLEDGEMENTS

This work was supported by Grant PGC2018-099449-A-I00 and by the Ramon y Cajal program (grant RYC-2017-23 645), both awarded to JTM, and to the FPI scholarship (PRE2019-088044) granted to LO from the Ministerio de Ciencia, Innovacion y Universidades (MCIU, Spain), Agencia Estatal de Investigacion (AEI, Spain), and Fondo Europeo de Desarrollo Regional (FEDER, European Union). CZ is supported by China Scholarship Council (CSC; no. 201906300087). KG and ZR were supported by the Slovenian Research Agency (grants P4-0165 and Z7-1888). SCH is partially supported by the National Science Foundation (grant PGRP IOS-1916804). This article is based upon work from COST Action CA 17111 INTEGRAPE, supported by COST (European Cooperation in Science and Technology). Data have been treated and uploaded in public repositories according to the FAIR principles, in accordance with the guidelines found at INTEGRAPE. The genomic data presented here will be re-analyzed and associated to the PN40024.v4 assembly and its structural annotation upon its release. Grapevine cell cultures in solid media were kindly provided by Roque Bru (Universidad de Alicante). Special thanks to Pere Mestre and Philippe Hugueney (INRAE Colmar) for providing the 35S:STS48 construct and the 35S:ROMT1 positive control used for agroinfiltration, to Anne-Francoise Adam-Blondon and Nicolas Francillonne (URGI INRAE Versailles) for guidance in the adaptation of JBrowse during an INTEGRAPE Short-Term Scientific Mission of LO, and to Susana Selles-Marchart, for her help with stilbenoid quantification.

AUTHOR CONTRIBUTIONS

LO, DN-P, and JTM designed the research; ML, JH, PM, GM, PR, and CZ collected plant material and performed the experiments; LO, AS, DN-P, S-sCH, ZR, and DW conducted all bioinformatic analyses; LO, DN-P, and JTM wrote the paper. JH, PM, JB, AV, and DC contributed with reagents. All authors contributed to interpretations and revisions of the manuscripts.

CONFLICT OF INTERESTS

The authors declare that they have no competing interests.

4.7 Supplementary Information

Additional supporting information too big to be displayed here, please consult at the following link: https://doi.org/10.1111/tpj.15686

Figure S11. Microarray expression of shikimate, early phenylpropanoid, and stilbenoid pathway genes across the *Vitis vinifera* cv. 'Corvina' expression atlas (Fasoli et al., 2012). Too big to be displayed here, please consult this figure at the indicated link.

Dataset S1. VviMYB13/14/15 DAP-Seq peaks, Gene Set Enrichment Analysis (GSEA) tables, STS genes and housekeeping genes.

Dataset S2. RNA-Seq SRA runs for network construction, MYB14 and MYB15 GCNs, HCTs, GSEA tables, and VCost.v3 to MapMan Bin associations.

Dataset S3. Primers, MYB15 overexpression cluster modules, affymetrix microarray dataset, microarray analysis (affymetrix probe-to-gene assignments and log2 fold-changes) and updated MapMan ontology.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally shareable data necessary to reproduce the reported results. The data is available at https://www.ncbi.nlm.nih. gov/geo/query/acc.cgi?acc=GSE180450

DATA AVAILABILITY STATEMENT

DAP-Seq and microarray data can be found in the Gene Expression Omnibus (GEO) database of the NCBI under the accessions GSE180450 and GSE195646, respectively. All relevant data can be found within the manuscript and its supporting materials.

BIBLIOGRAPHIC REFERENCES

Almagro L., Carbonell-Bejerano P., Belchi-Navarro S., Bru R., Martinez-Zapater J. M., Lijavetzky D. et al. (2014) Dissecting the transcriptional response to elicitors in Vitis vinifera cells. PLoS ONE, 9, e109777. https:// doi.org/10.1371/journal.pone.0109777.

Bai R., Luo Y., Wang L., Li J., Wu K., Zhao G. et al. (2019) A specific allele of MYB14 in grapevine correlates with high stilbene inducibility triggered by Al3+ and UV-C radiation. Plant Cell Reports, 38, 37–49. https://doi.org/10.1007/s00299-018-2347-9.

Ballouz S., Weber M., Pavlidis P. & Gillis J. (2016) EGAD: ultra-fast functional analysis of gene networks. Bioinformatics, 33, 612–614. https://doi. org/10.1093/bioinformatics/btw695.

Bartlett A., O'Malley R., Huang S.-S., Galli M., Nery J., Gallavotti A. et al. (2017) Mapping genome-wide transcription-factor binding sites using dap-seq. Nature Protocols, 12, 1659–1672. https://doi.org/10.1038/ nprot.2017.055.

Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B: Methodological, 57, 289–300. https://doi.org/ 10.1111/j.2517-6161.1995.tb02031.x https://rss.onlinelibrary.wiley.com/ doi/abs/10.1111/j.2517-6161.1995.tb02031.x.

Blanco-Ulate B., Amrine K. C., Collins T. S., Rivero R. M., Vicente A. R., Morales-Cruz A. et al. (2015) Developmental and metabolic plasticity of white-skinned grape berries in response to botrytis cinerea during noble rot. Plant Physiology, 169, 2422–2443. https://doi.org/10.1104/pp.15. 00852.

Bru R., Selles S., Casado-Vela J., Belchi-Navarro S. & Pedreno M. A. (2006) Modified cyclodextrins are chemically defined glucan inducers of defense responses in grapevine cell cultures. Journal of Agricultural and Food Chemistry, 54, 66–71. https://doi.org/10.1021/jf051485j.

Carbonell-Bejerano P., Rodrıguez V., Royo C., Hernaiz S., Moro-Gonzalez L.C., Torres-Vinals M. et al. (2014) Circadian oscillatory transcriptional programs in grapevine ripening fruits. BMC Plant Biology, 14, 1–15. https://doi.org/10.1186/1471-2229-14-78.

Chen S., Zhou Y., Chen Y. & Gu J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34, i884–i890. https://doi.org/10. 1093/bioinformatics/bty560.

Chen Y., Zhangliang C., Kang J., Kang D., Gu H. & Qin G. (2013) Atmyb14 regulates cold tolerance in arabidopsis. Plant Molecular Biology Reporter, 31, 87–97. https://doi.org/10.1007/s11105-012-0481-z.

Chin C. S., Peluso P., Sedlazeck F. J., Nattestad M., Concepcion G. T., Clum A. et al. (2016) Phased diploid genome assembly with single-molecule real-time sequencing. Nature Methods, 13, 1050–1054. https:// doi.org/10.1038/nmeth.4035.

Davies K. M., Jibran R., Zhou Y., Albert N. W., Brumell D. A., Jordan B. R. et al. (2020) The evolution of flavonoid biosynthesis: a bryophyte perspective. Frontiers in Plant Science, 11, 7. https://doi.org/10.3389/fpls. 2020.00007.

Dobin A., Davis C. A., Schlesinger F., Drenkow J., Zaleski C., Jha S. et al. (2012) STAR: ultrafast universal RNA-Seq aligner. Bioinformatics, 29, 15– 21. https://doi.org/10.1093/bioinformatics/bts635.

Dubrovina A. S. & Kiselev K. V. (2017) Regulation of stilbene biosynthesis in plants. Planta, 246, 597–623. https://doi.org/10.1007/s00425-017- 2730-8.

Fasoli M., Dal Santo S., Zenoni S., Tornielli G. B., Farina L., Zamboni A. et al. (2012) The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell, 24, 3489–3505. https://doi.org/10.1105/tpc.112.100230.

Gillis J. & Pavlidis P. (2012) "Guilt by association" is the exception rather than the rule in gene networks. PLoS Computational Biology, 8, 1–13. https://doi.org/10.1371/journal.pcbi.1002444.

Guo Y., Mahony S. & Gifford D. K. (2012) High resolution genome wide binding event finding and motif discovery reveals transcription factor spatial binding constraints. PLoS Computational Biology, 8, 1–14. https:// doi.org/10.1371/journal.pcbi.1002638.

Hall D. & De Luca V. (2007) Mesocarp localization of a bi-functional resveratrol/hydroxycinnamic acid glucosyltransferase of Concord grape (*Vitis labrusca*). Plant Journal, 49, 579–591. https://doi.org/10.1111/j.1365-313X. 2006.02987.x.

Holl J., Vannozzi A., Czemmel S., D'Onofrio C., Walker A., Rausch T. et al. (2013) The r2r3-myb transcription factors myb14 and myb15 regulate stilbene biosynthesis in *Vitis vinifera*. The Plant Cell, 25, 4135–4149. https://doi.org/10.1105/tpc.113.117127.

Huang L., Yin X., Sun X., Yang J., Rahman M., Chen Z. et al. (2018) Expression of a grape vqsts36-increased resistance to powdery mildew and osmotic stress in *Arabidopsis* but enhanced susceptibility to botrytis cinerea in *Arabidopsis* and tomato. International Journal of Molecular Sciences, 19, 2985. https://doi.org/10.3390/ijms19102985.

Hurtado-Gaitan E., Selles-Marchart S., Martinez-Marquez A., Samper-Herrero A. & Bru-Martinez R. (2017) A focused multiple reaction monitoring (MRM) quantitative method for bioactive grapevine stilbenes by ultra-high-performance liquid chromatography coupled to triple-quadrupole mass spectrometry (UHPLC-QqQ). Molecules, 22, 418. https://doi.org/10.3390/molecules22030418.

Jaillon O., Aury J. -M., Noel B., Policriti A., Clepet C., Casagrande A. et al. (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature, 449, 463–467. https://doi.org/10.1038/nature06148.

Jiang C., Wang D., Zhang J., Xu Y., Zhang C., Zhang J. et al. (2021) VqMYB154 promotes polygene expression and enhances resistance to pathogens in Chinese wild grapevine. Horticulture Research, 8, 1–17. https://doi.org/10.1038/s41438-021-00585-0.

Kelemen Z., Sebastian A., Xu W., Grain D., Salsac F., Avon A. et al. (2015) Analysis of the DNA-binding activities of the *Arabidopsis* r2r3-myb transcription factor family by one-hybrid experiments in yeast. PLoS ONE, 10, 1–22. https://doi.org/10.1371/journal.pone.0141044.

Kenrick P. & Crane P. (1997) The origin and early evolution of plants on land. Nature, 389, 33–39. https://doi.org/10.1038/37918.

Kim S. H., Lam P. Y., Lee M. -H., Jeon H. S., Tobimatsu, Y. & Park O. K. (2020) The *Arabidopsis* r2r3 myb transcription factor myb15 is a key regulator of lignin biosynthesis in effector-triggered immunity. Frontiers in Plant Science, 11, 1456. https://doi.org/10.3389/fpls.2020.583153 https:// www.frontiersin.org/article/10.3389/fpls.2020.583153.

Kolberg L., Raudvere U., Kuzmin I., Vilo J. & Peterson H. (2020) gprofiler2 – an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. F1000Research, 9, 709. https:// doi.org/10.12688/f1000research.24956.2.

Langfelder P. & Horvath S. (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics, 9, 1–13.

Langmead B. & Salzberg S. (2012) Langmead b, salzberg sl.. fast gapped-read alignment with bowtie 2. Nature Methods, 9, 357–359. https://doi. org/10.1038/nmeth.1923.

Lecourieux F., Kappel C., Pieri P., Charon J., Pillet J., Hilbert G. et al. (2017) Dissecting the biochemical and transcriptomic effects of a locally applied heat treatment on developing Cabernet Sauvignon grape berries. Frontiers in Plant Science, 8, 53. https://doi.org/10.3389/fpls.2017.00053.

Liao Y., Smyth G.K. & Shi W. (2013) featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features. Bioinformatics, 30, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

Machanick P. & Bailey T. (2011) Meme-chip: motif analysis of large dna datasets. Bioinformatics (Oxford, England), 27, 1696–1697. https://doi. org/10.1093/bioinformatics/btr189.

Martinez-Marquez A., Morante-Carriel J. A., Ramirez-Estrada K., Cusido, R. M., Palazon, J. & Bru-Martinez R. (2016) Production of highly bioactive resveratrol analogues pterostilbene and piceatannol in metabolically engineered grapevine cell cultures. Plant Biotechnology Journal, 14, 1813–1825. https://doi.org/10.1111/pbi.12539.

Merz P. R., Moser T., Holl J., Kortekamp A., Buchholz G., Zyprian E. et al. (2015) The transcription factor VvWRKY33 is involved in the regulation of grapevine (*Vitis vinifera*) defense against the oomycete pathogen *Plasmopara viticola*. Physiologia Plantarum, 153, 365–380. https://doi.org/10. 1111/ppl.12251.

Mutwil M., Klie S., Tohge T., Giorgi F. M., Wilkins O., Campbell M. M. et al. (2011) PlaNet: combined sequence and expression comparisons across plant networks derived from seven species. The Plant Cell, 23, 895–910. https://doi.org/10.1105/tpc.111.083667.

Nakata M. & Ohme-Takagi M. (2014) Quantification of anthocyanin content. Bio-Protocol, 4, e1098. https://doi.org/10.21769/BioProtoc.1098.

Navarro-Payà D., Santiago A., Orduna L., Zhang C., Amato A., Fattorini C. et al. (2022) The grape gene reference catalogue as a standard resource for gene selection and genetic improvement. Frontiers in Plant Science, 12, 803977. https://doi.org/10.3389/fpls.2021.803977.

O'Malley R., Huang S. S. C., Song L., Lewsey M., Bartlett A., Nery J. et al. (2016) Cistrome and epicistrome features shape the regulatory DNA landscape. Cell, 165, 1280–1292. https://doi.org/10.1016/j.cell.2016.04.038
http://www.sciencedirect.com/science/article/pii/S0092867416304810.

Parage C., Tavares R., Rety S., Baltenweck-Guyot R., Poutaraud A., Renault L. et al. (2012) Structural, functional, and evolutionary analysis of the unusually large stilbene synthase gene family in grapevine. Plant Physiology, 160, 1407–1419. https://doi.org/10.1104/pp.112.202705 http:// www.plantphysiol.org/content/160/3/1407.

Pasquereau S., Nehme Z., Haidar Ahmad S., Daouad F., Van Assche J., Wallet C. et al. (2021) Resveratrol inhibits HCoV-229E and SARS- CoV-2 coronavirus replication in vitro. Viruses, 13, 1–11. https://doi.org/10.3390/ v13020354 http://www.ncbi.nlm.nih.gov/pubmed/33672333.

Qian C., Chen Z., Liu Q., Mao W., Chen Y., Tian W. et al. (2020) Coordinated transcriptional regulation by the UV-b photoreceptor and multiple transcription factors for plant UV-b responses. Molecular Plant, 13, 777– 792. https://doi.org/10.1016/j.molp.2020.02.015 https://www.sciencedirect. com/science/article/pii/S1674205220300629.

Ramsak Z., Baebler S., Rotter A., Korbar M., Mozetic I., Usadel B. et al. (2014) GoMapMan: integration, consolidation and visualization of plant gene annotations within the MapMan ontology. Nucleic Acids Research, 42, 1167–1175. https://doi.org/10.1093/nar/gkt1056.

Ritchie M. E., Phipson B., Wu D., Hu Y., Law C. W., Shi W. et al. (2015) limma powers differential expression analyses for RNA-Sequencing and microarray studies. Nucleic Acids Research, 43, e47. https://doi.org/10. 1093/nar/gkv007.

Romero I., Fuertes A., Benito M. J., Malpica J. M., Leyva A. & Paz-Ares J. (1998) More than 80R2R3- MYB regulatory genes in the genome of *Arabidopsis thaliana*. Plant Journal, 14, 273–284. https://doi.org/10.1046/j. 1365-313X.1998.00113.x.

Santos-Rosa M., Poutaraud A., Merdinoglu D. & Mestre P. (2008) Development of a transient expression system in grapevine via agroinfiltration. Plant Cell Reports, 27, 1053–1063. https://doi.org/10.1007/ s00299-008-0531-z.

Schmidlin L., Poutaraud A., Claudel P., Mestre P., Prado E., Santos-Rosa M. et al. (2008) A stress-inducible resveratrol O- methyltransferase involved in the biosynthesis of pterostilbene in grapevine. Plant Physiology, 148, 1630–1639. https://doi.org/10.1104/pp.108.126003.

Schwacke R., Ponce-Soto G. Y., Krause K., Bolger A. M., Arsova B., Hallab A. et al. (2019) MapMan4: a refined protein classification and annotation framework applicable to multi-omics data analysis. Molecular Plant, 12, 879–892. https://doi.org/10.1016/j.molp.2019.01.003.

Shalit-Kaneh A., Kumimoto R. W., Filkov V. & Harmer S. L. (2018) Multiple feedback loops of the Arabidopsis circadian clock provide rhythmic robustness across environmental conditions. Proceedings of the National Academy of Sciences of the United States of America, 115, 7147–7152. https://doi.org/10.1073/pnas.1805524115.

Tamagnone L., Merida A., Parr A., Mackay S., Culianez-Macia F. A., Roberts K. et al. (1998) The AmMYB308 and AmMYB330 transcription factors from antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. The Plant Cell, 10, 135–154. https://doi.org/10. 1105/tpc.10.2.135.

Tarazona S., Garcia-Alcalde F., Dopazo J., Ferrer A. & Conesa A. (2011) Differential expression in RNA-seq: a matter of depth. Nucleic Acids Research, 43, e140. https://doi.org/10.1101/gr.124321.111.

Thimm O., Blasing O., Gibon Y., Nagel A., Meyer S., Kruger P. et al. (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant Journal, 37, 914–939. https://doi.org/10.1111/j.1365-313X.2004.02016.x.

Vannozzi A., Dry I., Fasoli M., Zenoni S. & Lucchin M. (2012) Genome-wide analysis of the grapevine stilbene synthase multigenic family: genomic organization and expression profiles upon biotic and abiotic stresses. BMC Plant Biology, 12, 8. https://doi.org/10.1186/1471-2229-12-130.

Vannozzi A., Wong D. C. J., Holl J., Hmmam I., Matus J. T., Bogs J. et al. (2018) Combinatorial Regulation of Stilbene Synthase Genes by WRKY and MYB Transcription Factors in Grapevine (*Vitis vinifera* L.). Plant and Cell Physiology, 59, 1043–1059. https://doi.org/10.1093/pcp/pcy045.

Verleyen W., Ballouz S. & Gillis J. (2014) Measuring the wisdom of the crowds in network- based gene function inference. Bioinformatics (Oxford, England), 31, 10. https://doi.org/10.1093/bioinformatics/btu715.

Wang D., Jiang C., Li R. & Wang Y. (2019) VqbZIP1 isolated from Chinese wild *Vitis quinquangularis* is involved in the ABA signaling pathway and regulates stilbene synthesis. Plant Science, 287, 110202. https://doi.org/ 10.1016/j.plantsci.2019.110202.

Wang D., Jiang C., Liu W., Wang Y. & Hancock R. (2020) The WRKY53 transcription factor enhances stilbene synthesis and disease resistance by interacting with MYB14 and MYB15 in Chinese wild grape. Journal of Experimental Botany, 71, 3211–3226. https://doi.org/10.1093/jxb/eraa097.

Wang L. & Wang Y. (2019) Transcription factor VqERF114 regulates stilbene synthesis in Chinese wild *Vitis quinquangularis* by interacting with VqMYB35. Plant Cell Reports, 38, 1347–1360. https://doi.org/10.1007/ s00299-019-02456-4.

Waters E. R. (2003) Molecular adaptation and the origin of land plants. Molecular Phylogenetics and Evolution, 29, 456–463. https://doi.org/10. 1016/j.ympev.2003.07.018 http://www.sciencedirect.com/science/article/ pii/S1055790303003130nn.

Wolfe C. J., Kohane I. S. & Butte A. J. (2005) Systematic survey reveals general applicability of 'guilt- by-association' within gene co-expression networks. BMC Bioinformatics, 6, 1–10. https://doi.org/10.1186/1471-2105-6- 227. Wong, D. (2020) Network aggregation improves gene function prediction of grapevine gene co-expression networks. Plant Molecular Biology, 103, 425–441. https://doi.org/10.1007/s11103-020-01001-2.

Wong D. C. J. & Matus J. T. (2017) Constructing integrated networks for identifying new secondary metabolic pathway regulators in grapevine: recent applications and future opportunities. Frontiers in Plant Science, 8, 1–8. https://doi.org/10.3389/fpls.2017.00505.

Wong D. C. J., Schlechter R., Vannozzi A., Holl J., Hmmam I., Bogs J. et al. (2016) A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family uncovers new insights into the regulation of stilbene accumulation. DNA Research, 23, 451–466. https://doi.org/10. 1093/dnares/dsw028.

Xi H., Ma L., Liu G., Wang N., Wang J., Wang L. et al. (2014) Transcriptomic analysis of grape (Vitis vinifera L.) leaves after exposure to ultraviolet C irradiation. PLoS ONE, 9, 1–24. https://doi.org/10.1371/journal. pone.0113772.

Xu W., Fuli M., Li R., Zhou Q., Yao W., Jiao Y. et al. (2019) Vpsts29/sts2 enhances fungal tolerance in grapevine through a positive feedback loop. Plant, Cell & Environment, 42, 2979–2998. https://doi.org/10.1111/ pce.13600.

Yin X., Singer S. D., Qiao H., Liu Y., Jiao C., Wang H. et al. (2016) Insights into the mechanisms underlying ultraviolet-c induced resveratrol metabolism in grapevine (*V. amurensis rupr*.) cv. "Tonghua-3". Frontiers in Plant Science, 7, 1–16. https://doi.org/10.3389/fpls.2016.00503.

Zhu L., Gazin C., Lawson N., Pages H., Lin S., Lapointe D. et al. (2010) Chippeakanno: a bioconductor package to annotate chip-seq and chip-chip data. BMC Bioinformatics, 11, 237. https://doi.org/10.1186/1471- 2105-11-237.

CHAPTER V: Functional characterization of *VviMYB108A* and *VviMYB108B*

5.1 Introduction

Taking advantage from the results obtained from the grapevine floral expression atlas (Chapter II) and from the DAP-Seq analysis of whorls-specific TFs (Chapter III), we focused on a specific floral whorl: the anther. We decided to concentrate on this particular tissue for several reasons: (i) among all the reproductive organs, the anther, together with the ovule, is one of the two major players in the process of fertilization *stricto sensu*. In this regard, a malfunctioning of the biogenetic process which specifically characterizes the anther would lead to a lack or reduced production of pollen, or in any case to a drastic decrease in plant fertility, with consequent absence or sub-efficiency in fecundation, from which would derive a loss in the final production, i.e. the berry. Knowing the genetic-molecular mechanisms that underlie microgametogenesis and the close relationships between pollen and anther, means being able to directly control the fertilization process and therefore having the possibility of qualitatively and quantitatively implementing the final production, acting in a targeted manner on the individual actors of the entire process, knowing exactly the metabolic pathways on which to interfere. The second main reason why our attention was focused on that organ is because the main regeneration procedures and *in vitro* tissue culture protocols for grapevine start from anthers and filament tissues explants (Gambino et al., 2007; Kikkert et al., 2005; Perrin et al., 2004; Stamp and Meredith, 1988). For these purposes, it was decided to consider the TF *VviMYB108A*, which – basing on our results – was observed to be specific of anther whorl (Chapter II) and the achievement of the high confidence regulative network gave back results well consistent with the putative biological function, also considering the *De novo* motif discovery analysis as further corroboration (Chapter III). Better to understand the role of this TF, it was decided to consider the related ortholog in *Arabidopsis thaliana*, namely *AtMYB108*. Its expression was observed to be restricted to anthers (Fig. 1.5B, C and D; Mandaokar and Browse, 2009), observation perfectly coherent with the strict association between this whorl and *VviMYB108A* that we described occurring in grapevine (Chapter II). Moreover, it seems to mediate stamens and pollen maturation in the transcriptional cascade taking place in response to jasmonate signaling together with MYB24 (Fig. 1.5A), indeed three different *myb108* mutant lines showed reduced male fertility (Fig. 2.5) associated with delayed anther dehiscence, reduced pollen viability

(Fig. 3.5) and a global decreased fecundity compared to the wild type (Mandaokar and Browse, 2009). In addition, the first time when *AtMYB108* was isolated, it was named *BOTRYTIS SUSCEPTIBLE 1 (BOS1)* due to its involvement in *Botrytis cinerea* response, observing the mutants (generated with T-DNA insertion by McElver et al., 2001) exhibiting enhanced disease symptoms. In this case, it was also proposed for the first time its putative involvement in jasmonate signaling pathway (Mengiste et a., 2003). Subsequent studies on the same *bos1* mutants – initially considered of "loss of function" type - have highlighted an uncontrolled PCD (programmed cell death) starting from the margins of a wound artificially inflicted on the *Arabidopsis* leaf surface, which propagates throughout the leaf up to reach the healthy tissues and to cause the death of the entire plant after a few weeks, with an independent effect from light and humidity (Fig. 4.5). After 5 days from the infliction of injury, *bos1* mutants showed a high accumulation of ROS in the tissues immediately surrounding the wounds. Furthermore, the exogenous application of SA, JA and ethylene did not have any influence on the progression of cell death (Cui et al., 2013). Incredibly, what subsequent studies show is totally the opposite of what was initially believed. Indeed, it was shown that *BOS1* promotes, rather than inhibits, the propagation of cell death. The previous studies used a single *Arabidopsis* mutant (*bos1-1*), which contains a T-DNA insertion in the 5'-untranslated region. This T-DNA insertion did not disrupt *BOS1* as expected, but instead, it altered the activation of *BOS1* upon wounding and pathogen attack due to sequences in the T-DNA that regulate gene activation (Cui et al., 2022). Summarizing all the observations expressed so far, there are strong elements that empower the hypothesis that MYB108 is therefore an intrinsic factor of susceptibility to *Botrytis*. At this point, better to understand the applicability to *Vitis vinifera* of the role of MYB108 observed in *A. thaliana*, it was decided to have a look at the phylogenetic relationships between *Arabidopsis* and grapevine R2R3-MYB families and from that, it was realized that what in the first case is a single ortholog gene, in *Vitis* it gave rise to two paralogue genes, namely *VviMYB108A* and *VviMYB108B* (Fig. 5.5; Wong et al., 2016). For this reason and better to understand the effect of the gene duplication and differentiation event in terms of functions repartition, mutual interaction and difference among the two paralogues, we proceeded with the functional characterization of both. In this regard, it was decided to express these two grapevine TFs in several heterologous systems. To this aim, the subcellular localization was investigated by means of a transient overexpression of the two paralogues in *N. benthamiana* leaves. Parallelly, stable overexpressing lines of tomato (*Lycopersicon esculentum* var. "Micro Tom") and *A. thaliana* were produced. Finally, the complementation of the mutants used in Mandaokar and Browse (2009) and in Cui et al. (2013) (SALK lines: SALK_056061

and SALK_024059) was performed. Due to the long timings required for obtaining transformed plants using the previously described techniques, even though the transformed material has been already collected, analyses has not been conducted by the end of the PhD activities. For this reason, except for the subcellular localization, only activities carried out to gather the biological material that will be analyzed in a very near future are reported . All the work described in this chapter was performed during an abroad research period at TOMSBiolab, based in the Institute for Integrative Systems Biology (I2SysBio), CSIC – Universitat de Valencia, in Spain.

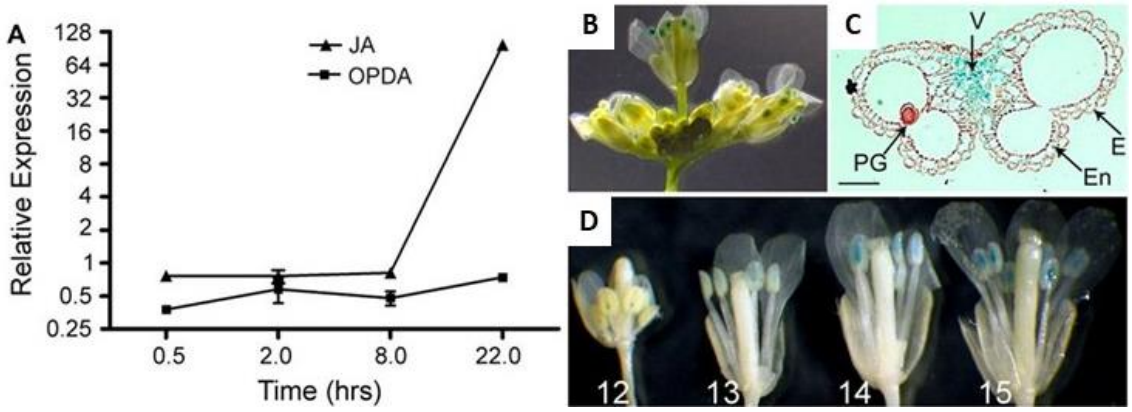


Figure 1.5. (A) At*MYB108* transcript levels in the jasmonate-synthesis mutant *opr3* stamens following treatment with jasmonate (JA) or OPDA (12-oxo-phytodienoic acid, a precursor of JA). (From B to D) GUS expression in the MYB108 gene-trap line. (B) Overview of the apical bud cluster. (C) Cross section of an anther from stage 13. Abbreviations: E, epidermis; En, endodermis; PG, pollen grain; V, vascular bundle. Bar = 50 μm. (D) Flowers from stages 12 to 15 (figure taken from Mandaokar and Browse, 2009).

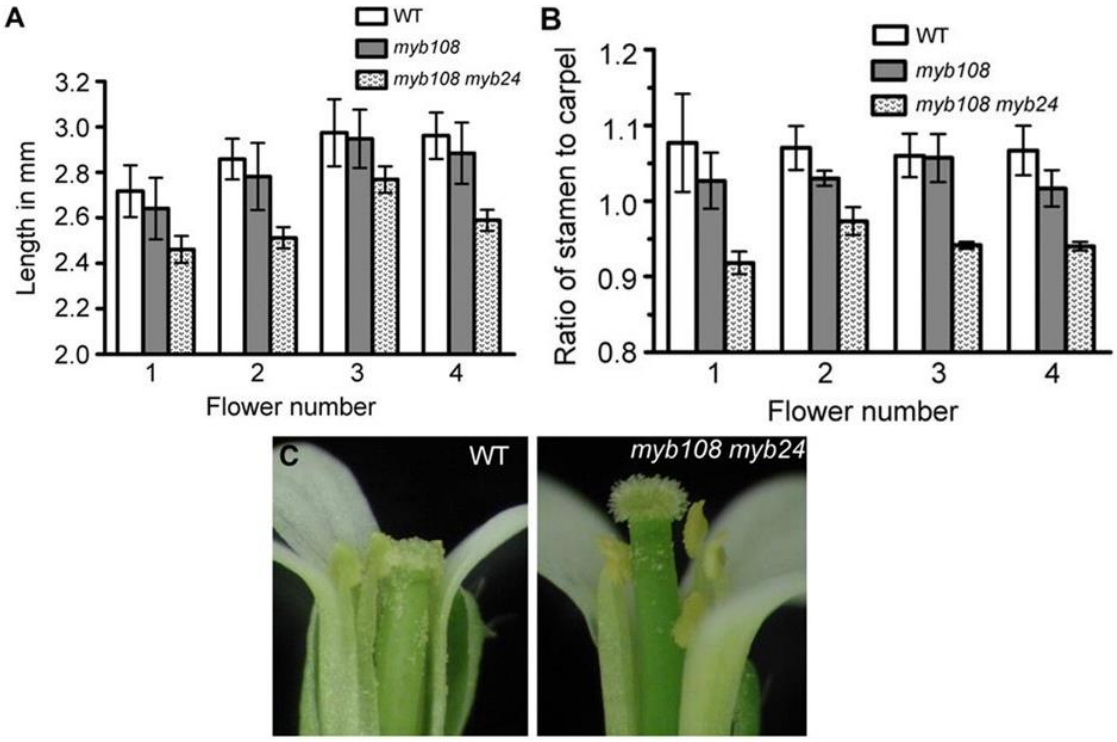


Figure 2.5. Stamen height is reduced in the *myb108 myb24* mutant. (A) Average length of filament plus anther measured for 12 long stamens from wild type (WT), *myb108*, and *myb108 myb24* plants at stages 1 to 4. (B) The ratio of stamen length to carpel length in WT, *myb108*, and *myb108 myb24* flowers. (C) Deposition of pollen on the carpels of WT and *myb108 myb24* flowers (figure taken from Mandaokar and Browse, 2009).

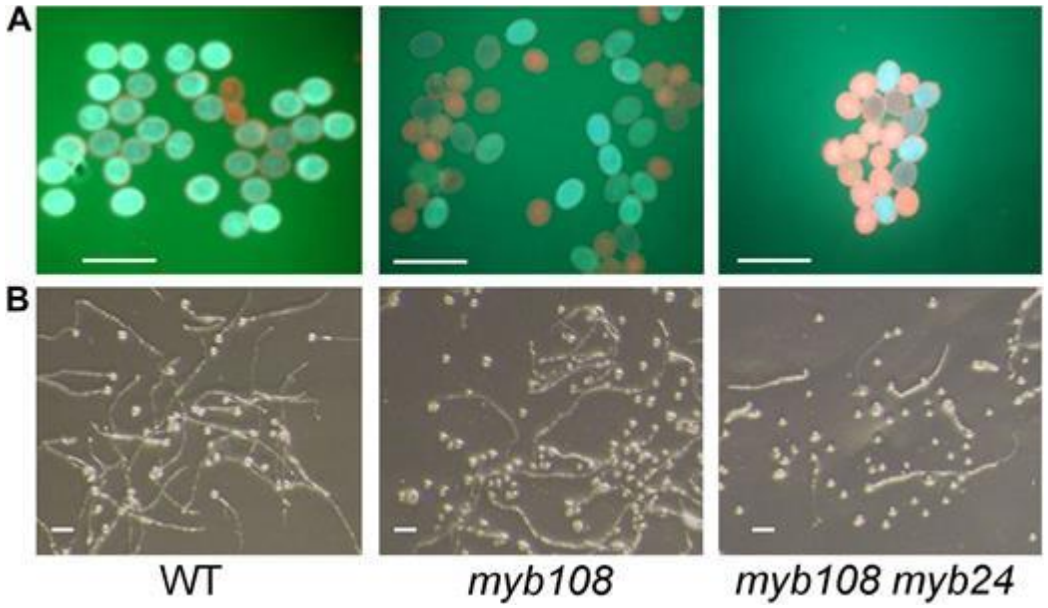


Figure 3.5. Reduced pollen viability and germination in *myb108* and *myb108 myb24* mutants. (A) Pollen from wild-type (WT) and mutant lines was stained with fluorescein diacetate and propidium iodide. This protocol stains viable pollen blue-green and inviable pollen red-brown. (B) Pollen harvested from mature open flowers was incubated on germination medium for 10 h (figure taken from Mandaokar and Browse, 2009).

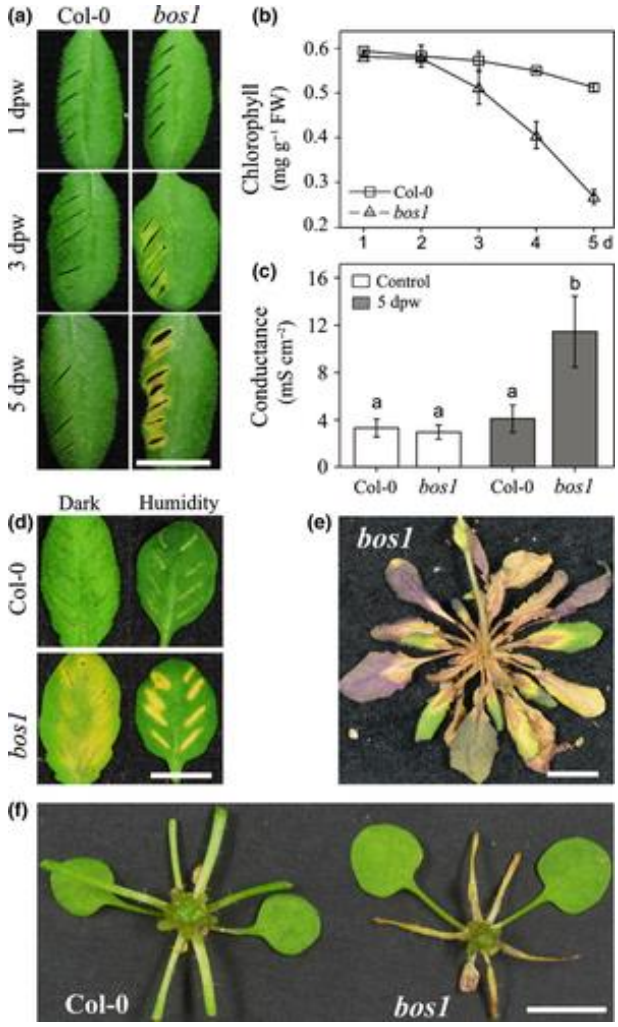


Figure 4.5. Cell death spread from wound sites in the *Arabidopsis thaliana* mutant *bos1*. Cell death spreading from a scalpel blade inflicted wound, measured at the indicated time in days post wounding (dpw) measured by: (A) photographs taken at 1, 3 and 5 dpw. (B) Chlorophyll retention assay. (C) Cell death quantified as ion leakage. Letters above error bars (± SE of means, $n = 6$) represent statistical significance (Student's *t*-test, $P < 0.01$). (D) Cell death spread from forceps pinching inflicted wound in *bos1* leaves also occurred when leaves were placed in the dark or under high humidity conditions immediately after wounding and remained there continuously until photographed. (E) Cell death spread from the wounding sites from the excision of four leaves advances through rosette core into unwounded leaves. This process takes 3–5 wk post-wounding (wpw), photograph taken at 4 wpw. (F) Photograph depicting cell death spread into petioles 7 days after leaf excision: Col-0 (wild type) petioles remain healthy (left; with six leaves removed) while

bos1 exhibited cell death from leaf excisions wounds had spread into the petioles (right; with five leaves removed). Bars, 1 cm. (Figure taken from Cui et al., 2013).

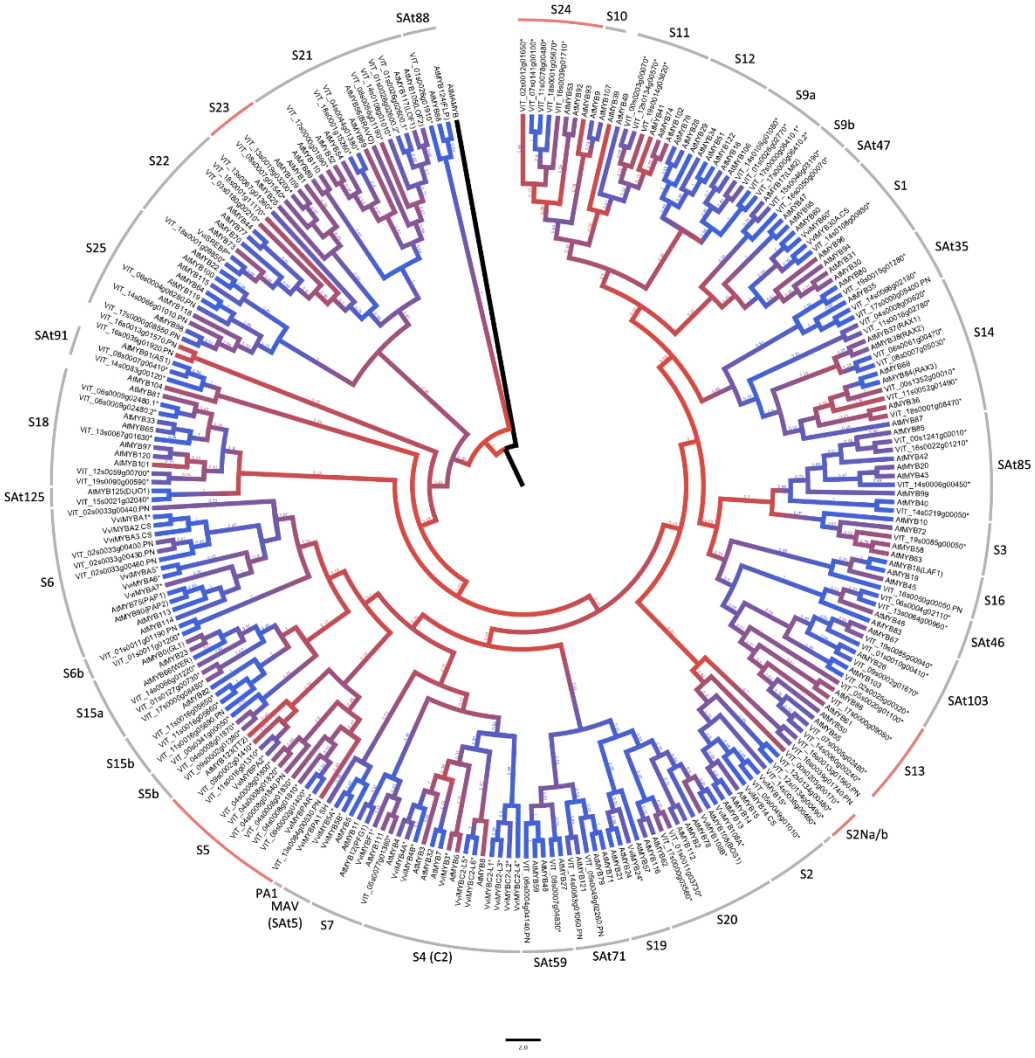


Figure 5.5. Phylogenetic relationships between grapevine and *Arabidopsis* R2R3-MYB families (figure taken from Wong et al., 2016).

5.2 Material and methods

5.2.1 Plasmid preparation

The ORFs of *VviMYB108A* and *VviMYB108B* were retrieved as described in Chap. III, Section "3.2 Material and methods", paragraph "3.2.1 Allele frequencies screening", synthesized directly in pDONR221 vector by ordering to GeneArt service (Thermo Fisher) and recombined in the overexpression vector pMpGWB406 (Fig. 6.5; Ishizaki et al., 2015) by using Gateway LR Clonase II Enzyme Mix (Thermo Fisher) setting up the following protocol as provided by manufacturer:

I. 50 µl reaction setup:
   - 1-7 µl of entry clone (total amount of 150 ng of pDONR221_ORF)
   - 1 µl of empty pMpGWB406 (150 ng)
   - Up to 8 µl of TE buffer
   - 2 µl of LR Clonase II enzyme mix

II. Incubate at 25°C for 1 hour

III. Add 1 µl of Proteinase K and incubate at 37°C for 10 minutes to inactivate the reaction

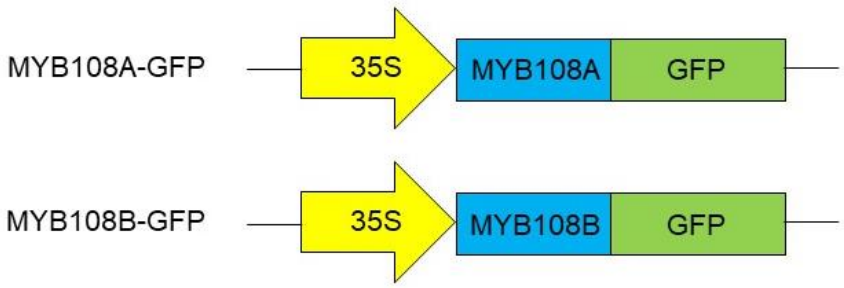


Figure 6.5. Highlight on the structure of pMpGWB406-ORFs plasmids: the yellow arrows represent the strong promoter CaMV 35S, the light blue rectangles are *VviMYB108A* and *B* ORFs and the green rectangles stand for the green fluorescent protein fused immediately after TF ORFs.

5.2.2 *Escherichia coli* culture transformation and plasmid collection

PROTOCOL

I. Add 5 µl of LR reaction obtained as described in previous paragraph to 40 µl of chemically competent (prepared as described in Chapter III, Section "3.2 Material and methods", paragraph "3.2.3 Preparation of chemically competent *Escherichia coli* strain TOP10 cells") *Escherichia coli* cells and mix gently not by pipetting up and down
II. Incubate on ice for 30 minutes
III. Heat-shock in 42°C bath for 30 seconds and back in in ice for 2 minutes
IV. Add 200 µl of LB and incubate at 37°C for 1 hour with constant shaking at 200 rpm
V. Plate on a 1.5% agarose LB Petri + Spectinomycin (100 mg/ml)
VI. Incubate overnight at 37°C
VII. Assess the presence of the plasmid with a colony-PCR. In our case, specific primers were used to confirm the presence of the predicted size amplicon. In particular, the forward primer was designed to match the CaMV 35S promoter of the pMpGWB406 plasmid (for this reason is the same for both MYB108A and MYB108B) and the reverse is gene-specific for the end of the ORF (Forward: 5'- CTTCGCAAGACCCTTCC -3'; MYB108A reverse: 5'- TCAGATGTTGGTGAACTGCGTCTGTACG -3'; MYB108B: 5'- CTAGAAGTCGTCAAAGAACTGTT -3')
VIII. Pick up a single transformed colony, put in 5 ml of LB + Kanamycin (50 mg/ml) and incubate overnight at 37°C with constant shaking at 200 rpm
IX. Plasmid purification miniprep using PureLink Quick Plasmid Miniprep Kit (Thermo Fisher) and sequencing to assess the correctness of the vector

5.2.3 Preparation of chemically competent *Agrobacterium tumefaciens* strain C58 and C58C1 (modified from Xu and Qingshun, 2008)

REAGENTS

Luria-Bertani broth (LB, pH7)	
Reagent	**Concentration (g/L)**
Triptone	10

NaCl	10
Yeast Extract	5
milliQ H_2O	up to volume

- 100 mM $CaCl_2$ + 15% glycerol

PROTOCOL

I. Grow a colony of *Agrobacterium tumefaciens* strains C58 and C58C1 (suitable for transient and stable overexpression, respectively) in 2 ml LB liquid media + Rifampicin (100 mg/ml; intrinsic resistance of C58 and C58C1 strains) by incubating at 28°C overnight with constant shaking (250 rpm)

II. Transfer the cells to 50 ml LB liquid media and incubate for 3 - 4 hours with constant shaking until it is reached an OD_{600} = 0.5 - 1.0

III. Chill the culture on ice for 5 minutes

IV. Centrifuge at 4000 rpm for 5 minutes at 4°C and discard the supernatant

V. Rinse the cell pellet with 10 ml of 100 mM ice-cold $CaCl_2$ and spin briefly again

VI. Add 1 ml of 100 mM ice-cold $CaCl_2$ + 15% glycerol to resuspend the cells gently on ice

VII. Aliquot 40 µl into 1.5 ml Eppendorf tubes and store at -80°C

5.2.4 Agrobacterium tumefaciens transformation

REAGENTS

Liquid MS medium with sucrose	
Reagent	**Concentration**
Murashige and Skoog basal medium	4.3 g/L

Sucrose	30 g/L
Acetosyringone (ACS)*	200 µM
milliQ H_2O	up to volume

*Add just before use

PROTOCOL

I. Let thaw on ice a 40 µl aliquot of competent *Agrobacterium tumefaciens*
II. Put 1 µl of plasmid into the thawed aliquot and mix gently not by pipetting up and down
III. Incubate on ice for 30 minutes
IV. Put the tube in liquid nitrogen for 1 minute and then transfer it into an incubator for 5 minutes at 37 °C
V. Add 400 µl of liquid LB and incubate at 28 °C for 3-4 hours with constant shaking at 200 rpm
VI. Plate on a 1.5% agarose LB Petri + Rifampicin and Spectinomycin (100 mg/ml)
VII. Incubate at 28 °C for 2-3 days
VIII. Assess the presence of the plasmid with a colony-PCR (primer reported at point VII of paragraph "5.2.2 *Escherichia coli* culture transformation and plasmid collection")
IX. Pick up a single transformed colony, put in 3 ml of LB + Rifampicin and Spectinomycin (100 mg/ml) and incubate overnight at 37°C with constant shaking at 200 rpm
X. Inoculate 20 µl of the overnight grown culture in 20 ml of LB + Rifampicin and Spectinomycin (100 mg/ml) and incubate again overnight at 37°C with constant shaking at 200 rpm
XI. Centrifuge at 6000 rpm for 10 minutes and discard supernatant
XII. Resuspend cellular pellet in 15 ml of Liquid MS medium with sucrose and incubate at 28 °C for 3-4 hours with constant shaking at 200 rpm
XIII. Measure the OD_{600} and dilute the culture with Liquid MS medium with sucrose to obtain 0.1-0.2 values
XIV. *Agrobacterium* culture is now ready to be used for plant transformation

5.2.5 Wild type tomato preparation

REAGENTS

Solid MS medium (pH 5.8-6.0)	
Reagent	**Concentration (g/L)**
Murashige and Skoog basal medium	4.3
Sucrose	20
Agarose	7
milliQ H_2O	up to volume

PROTOCOL

I. Work under a biological flow chamber in order to maintain the absolute sterility
II. Use medical gauze to prepare some packets of approximately 50 viable seeds each
III. Prepare a jar with a solution 50% bleach and add a few drops of polysorbate 20 (Tween-20)
IV. Prepare 3 jars with sterile water
V. Put the seed bags into the mixture prepared in step II and shake it gently for 30 minutes
VI. Transfer the bags into the first jar of sterile water and shake it gently for 5 minutes
VII. Transfer the bags into the second jar of sterile water and shake it gently for 10 minutes
VIII. Transfer the bags into the third jar of sterile water and shake it gently for 15 minutes
IX. Remove the packets from the jar and transfer the seeds into a Petri plate with sterile bibulous paper moistened with sterile water. Close the plate with parafilm in order to maintain sterility and incubate at 28 °C for 2-3 days with darkness
X. Transfer the seedlings on Solid MS medium and let them grow for 2-3 weeks with long photoperiod (16 hours of light and 8 hours of dark)

5.2.6 Cotyledons cut and infection

REAGENTS

Solid Pre- and Co-cultivation medium (pH 5.8-6.0)

Reagent	Concentration
Murashige and Skoog basal medium	4.3 g/L
Sucrose	20 g/L
Indole-3-acetic acid (IAA)*	0.2 mg/L
Zeatin (ZT)*	2 mg/L
Acetosyringone (ACS)*	200 µM
Agarose	7 g/L
milliQ H_2O	up to volume

*Add just before use

PROTOCOL

I. Work under a biological flow chamber in order to maintain the absolute sterility
II. When seedlings show well outspread cotyledons, it is the correct time to cut them
III. By picking the seedlings one by one, collect the cotyledons by cutting with a scalpel and remove the ends of the leaf surface with two cuts perpendicular to the longitudinal axis of the leaf (the point of insertion of the petiole and the exact opposite, where the leaf ends in a tip)
IV. Put the explants in plates with Solid Pre- and Co-cultivation medium, positioning the superior surface facing down in contact with the medium
V. Incubate at 25 °C in darkness for 3 days
VI. Dip the cotyledons into a C58C1 recombinant *Agrobacterium* suspension obtained as reported in step XIV of the paragraph "5.2.4 *Agrobacterium tumefaciens* transformation" and mix gently for 10 minutes
VII. Take out the cotyledons and put on sterilized filter paper to absorb the suspension, having care not to over dry

VIII. Put the cotyledons in plates with Solid Pre- and Co-cultivation medium having previously positioned a circle of sterilized filter paper on the medium surface, positioning the superior surface of the cotyledons facing down in contact with the medium (Fig. 7.5A)

IX. Incubate at 25 °C in darkness for 48 hours

5.2.7 Washing *Agrobacterium* from infected cotyledons

REAGENTS

Liquid MS medium	
Reagent	**Concentration**
Murashige and Skoog basal medium	4.3 g/L
Timentin*	300 mg/L
milliQ H_2O	up to volume

Solid Differential medium (pH 5.8-6.0)	
Reagent	**Concentration**
Murashige and Skoog basal medium	4.3 g/L
Sucrose	20 g/L
Indole-3-acetic acid (IAA)*	0.2 mg/L
Zeatin (ZT)*	2 mg/L
Timentin*	300 mg/L
Agarose	7 g/L
milliQ H_2O	up to volume

Solid Selective medium (pH 5.8-6.0)	
Reagent	**Concentration**
Murashige and Skoog basal medium	4.3 g/L
Sucrose	20 g/L
Indole-3-acetic acid (IAA)*	0.2 mg/L
Zeatin (ZT)*	2 mg/L
Timentin*	300 mg/L
Kanamycin	80 mg/L
Agarose	7 g/L
milliQ H_2O	up to volume

*Add just before use

PROTOCOL

I. Work under a biological flow chamber in order to maintain the absolute sterility
II. Dip the cotyledons in 30 ml of Liquid MS medium in a 50 ml Falcon tube and shake it softly for 4 minutes
III. Remove the Liquid MS medium, add 40 ml of sterile water, shake it softly for 2 minutes and remove the water
IV. Repeat steps I and II for two times
V. Take out the cotyledons and put on sterilized filter paper to absorb the suspension, having care not to over dry
VI. Put the cotyledons in plates with Solid differential medium positioning the superior surface of the facing down in contact with the medium (Fig. 7.5B)
VII. Leave the plates in long photoperiod conditions for 10 days
VIII. Transfer the cotyledons on Solid selective medium and reposition the plates in long photoperiod conditions

IX. Every 20 days, transfer the cotyledons into new Solid selective medium

5.2.8 Transformant whole plant regeneration

REAGENTS

Solid Rooting medium (pH 5.8-6.0)	
Reagent	**Concentration**
Murashige and Skoog basal medium	4.3 g/L
Sucrose	20 g/L
Indole-3-acetic acid (IAA)*	0.5 mg/L
Timentin*	200 mg/L
Agarose	7 g/L
milliQ H_2O	up to volume

*Add just before use

PROTOCOL

I. Work under a biological flow chamber in order to maintain the absolute sterility
II. When a resistant bud appears (Fig. 7.5C), cut it and place on Solid rooting medium keeping long photoperiodic conditions (Fig. 7.5D)
III. When the bud develops into a complete plantlet with a well expanded root system (Fig. 7.5E), cut the stem and put it into a new Solid rooting medium in order to conserve a copy of the plantlet, also by having care to leave at least one meristem in both the original plantlet and the copy, which allows them to regrow
IV. Put the plantlets in long photoperiod conditions
V. When a plantlet is completely regenerated again, take it off from the medium and carefully clean the roots under a water flow, having care of removing all the medium residues which – if not wiped away - would provoke the growth of fungal infestation due to their sucrose

content, and thus put it in a balanced growth substrate placed in a pot. Due to the *in vitro* growth conditions, the stomatal openings of the plant are wide opened, for this reason it is necessary to proceed with a gradual acclimatation of the plantlet, firstly by positioning a transparent cup covering all the epigeal apparatus in order to maintain high humidity levels (Fig. 8.5A). Let the plantlet with long photoperiod conditions and water abundantly

VI. After a week, replace the transparent cup with a suitably perforated one to further encourage the closure of stomata (Fig. 8.5B)

VII. After another week, remove completely the transparent cup and let pass 1 hour, then reposition the cup above the plant for another hour and take off it another time, proceeding in this way until the plant can remain with no coverage without assuming a patent bearing due to excessive evapotranspiration. After a week passed without evident stress, it is possible to collect a leaf sample for the plant, extract gDNA and test the presence of the insert with a PCR by using the primers reported in step VII of paragraph "5.2.2 *Escherichia coli* culture transformation and plasmid collection"

VIII. Let the plant grow in long photoperiod condition till the end of its cycle, hence gather the berry and from that the seeds

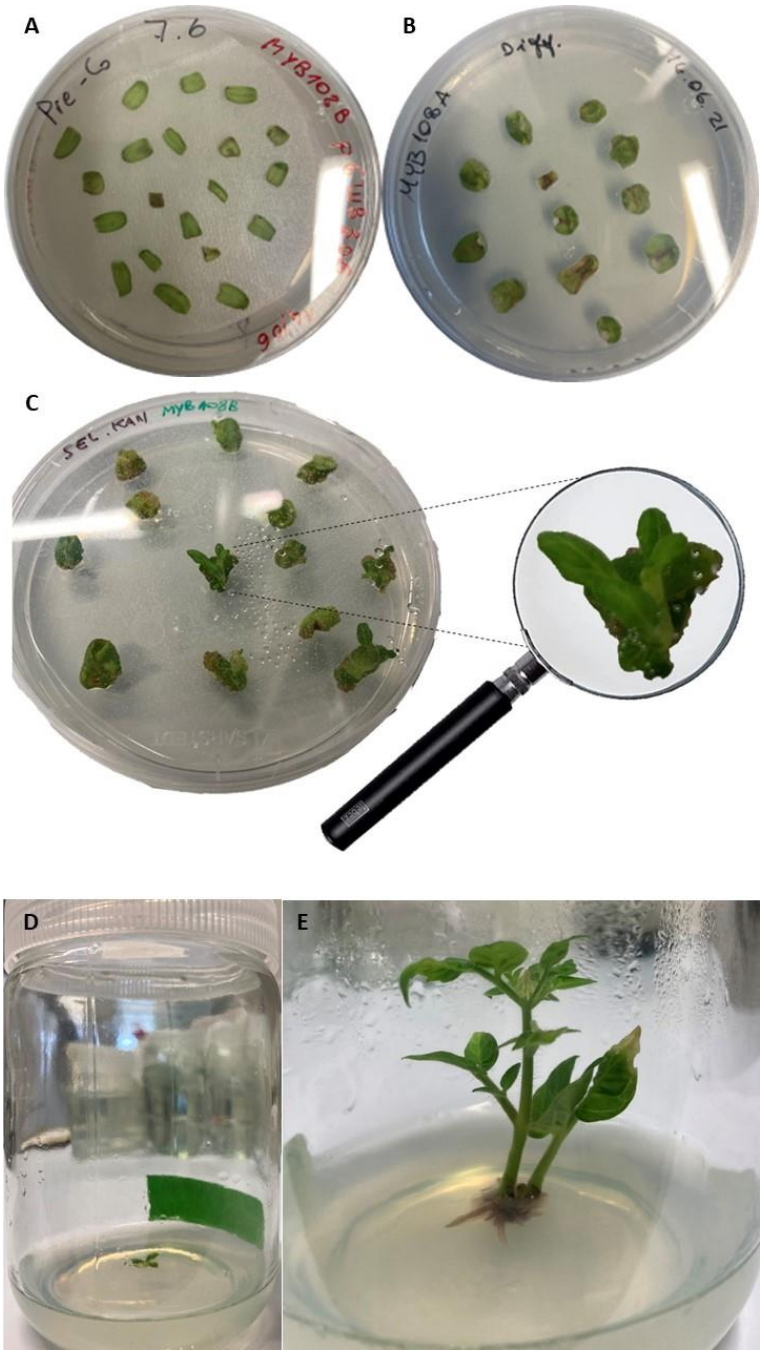


Figure 7.5. Tomato *in vitro* regeneration procedure. (A) Infected cotyledons in Pre- and Co-cultivation media with filter paper (as described in step VIII of the paragraph "5.2.6 Cotyledons cut and infection"). (B) Disinfected transformed cotyledons in Solid Differential media (as described in step VI of the paragraph "5.2.7 Washing *Agrobacterium* from infected cotyledons"). As can be seen, cotyledons have a well-developed callus cells cluster if compared to panel A. (C) Cotyledons in Solid Selective medium with some buds in the right stage to be cut and transplanted in Rooting medium. As can be seen from the zoom of a cotyledon, a bud can be transplanted when it shows visibly differentiated anatomical structures which make it appearing like a miniatured plant. (D) A transformant bud while regenerating in Solid Rooting medium. Pivotal is the size of the container

that allows the correct growth of the epigeal apparatus and the root system of the plant. (E) A well regenerated plantlet ready to be transplanted to a pot with containing some balanced substrate for tomato.

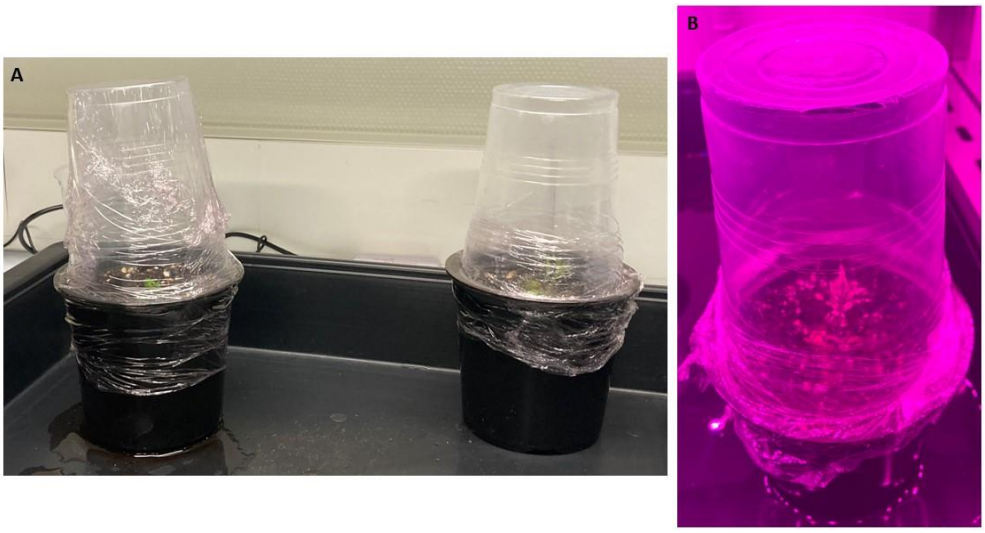


Figure 8.5. (A)Two plantlets just positioned in pots and covered with a transparent jar ready to start the acclimatation process. (B) A plantlet in a phytotron with holed coverage in the middle of the acclimatation procedure

5.2.9 *Arabidopsis thaliana* floral dip (modified from Zhang et al., 2006)

REAGENTS

Infection buffer	
Reagent	**Concentration**
$MgCl_2$	10 mM
Sucrose	5 %
Silwet L-77*	0.02 %
milliQ H_2O	up to volume

*Add just before use

PROTOCOL

I. Sow *Arabidopsis thaliana* wild type in a pot with soil surface covered by a cheesecloth to retain the substrate but with sufficiently large meshes to let the seedlings emerge and develop without constraints (Fig. 9.5). Let grow with long photoperiod conditions
II. When the primary inflorescence is completely issued (Fig. 10.5), cut it at the basis to stimulate the emission of the secondary ones which are more susceptible to the transformation than the principal
III. Once the secondary inflorescences are well developed and the flowers are still closed, the plants are ready to be transformed
IV. On the same day of the transformation, remove with a tweezer the precociously opened flowers eventually present
V. With the same procedure and proportions reported in the paragraph "5.2.4 *Agrobacterium tumefaciens* transformation", grow a recombinant *Agrobacterium* C58C1 suspension of 500 ml of final volume. Always check the culture with a colony PCR before to use
VI. Centrifuge at 4000 rpm for 10 minutes and discard supernatant
VII. Resuspend the pellet in 200 ml of Infection buffer
VIII. Measure the OD_{600} and dilute the culture with infection buffer to obtain 0.8-0.9 values
IX. Pour the suspension in a wide mouth container
X. Dip the inflorescence of *Arabidopsis* in the suspension for 20-30 seconds and let it drain off from excess of adhered suspension
XI. Wrap the plant with transparent film in order to maintain high humidity levels and position laying on the side for 24 in darkness
XII. Uncover the plant and let it grow with long photoperiod conditions till the siliques are totally dry and it is possible to easy collect seeds
XIII. Sterilize transformant seeds as described in paragraph "5.2.5 Wild type tomato preparation" and sow them in Solid MS medium + Kanamycin (80 mg/L) as selection factor (Fig. 11.5)

Figure 9.5. Pot with soil surface covered by a cheesecloth to retain the substrate but with sufficiently large meshes to let the seedlings emerge and correctly develop without constraints

Figure 10.5 *Arabidopsis thaliana* plant with principal inflorescence axis in the correct stage to be cut in order to stimulate the emission of secondary floral stems

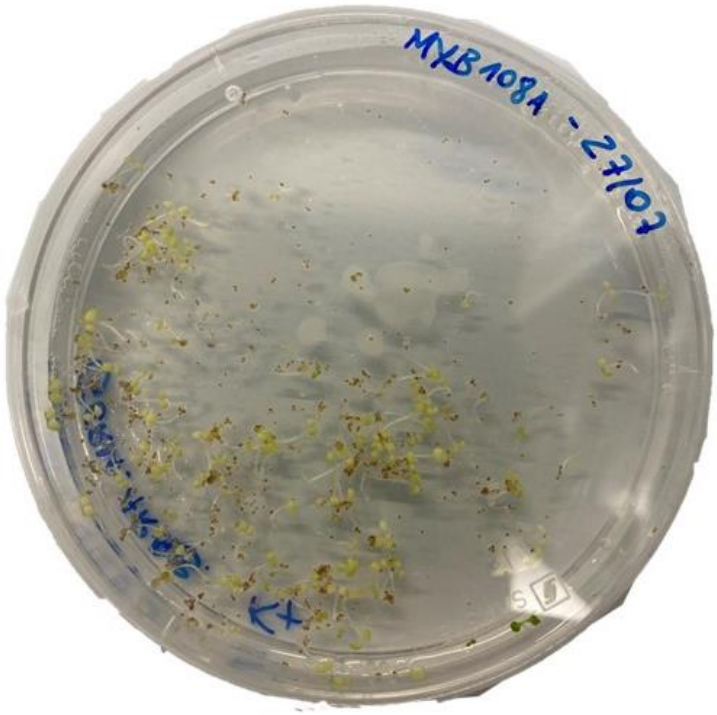


Figure 11.5. Transformant seedlings of *Arabidopsis thaliana* germination on Kanamycin selective Solid MS medium

5.2.10 *Arabidopsis* mutant lines complementation

The seeds of mutant lines of *Arabidopsis thaliana* SALK_056061 and SALK_024059 were ordered at Nottingham Arabidopsis Stock Centre (NASC; https://arabidopsis.info/BasicForm). As reported in user guide, before to proceed with the complementation, it was necessary to genotype the lines in order to be sure to use only homozygous individuals with an insertion in both the chromosomes. To this purpose, it was necessary to design a genomic-specific couple of primers (LP and RP) for each line and a T-DNA insertion specific forward primer (LB1.3). In this regard, for SALK_056061 line the primers are LP: 5'-TGTCTATTCCCTTCATCGCAC-3' and RP: 5'-CAGGGCGGAGATAGTTTAACC-3'. For SALK_024059 line, the primers are LP: 5'-ATTGTGCGATTTTTGACCATC-3' and RP: 5'-AGGTCGCTGGAACTCTCTTTC-3'. The forward primer specific for the T-DNA insertion is LBb1.3: 5'-ATTTTGCCGATTTCGGAAC-3' and it is suitable for both the lines (Fig. 12.5). Primers were designed by using the SIGNAL primer design online tool (http://signal.salk.edu/tdnaprimers.2.html). To distinguish wild type, homozygous insertions and heterozygous ones, for each individual, it was necessary to extract gDNA and to set up two paired PCR reactions: one with LP + RP primers and the

other with LBb1.3 + RP. For wild type or heterozygous individual, it is expected a PCR product in the LP + RP reaction or a blank for homozygous, while it is expected a band in the LBb1.3 + RP for homozygous or heterozygous plants. Taking advantage of the SIGNAL online tool, it was possible to predict the sizes of the PCR fragments. SALK_056061: LP + RP = 1218 bp and LBb1.3 + RP = 607-907 bp. SALK_024059: LP + RP = 1271 bp and LBb1.3 + RP = 583-883 bp. Some results of this procedure are shown in Fig. 13.5. After this genotyping analysis, it was proceeded with the complementation of the mutants by using only the homozygous individuals, following again the floral dip method as described in paragraph "5.2.9 *Arabidopsis thaliana* floral dip".

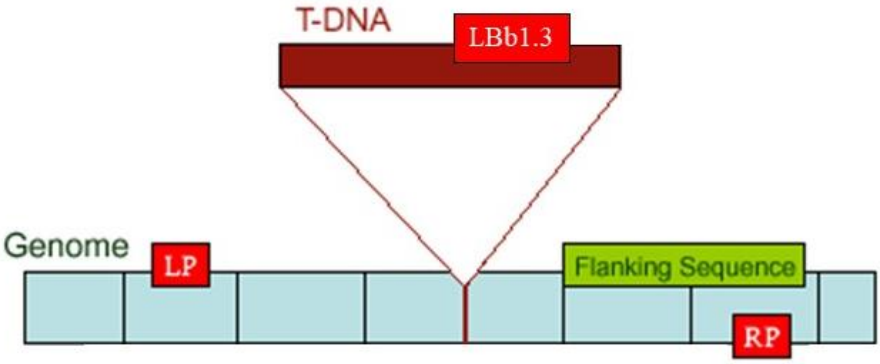


Figure 12.5. Localization of LP, RP and LBb1.3 primers along the genomic region in which is supposed to be the T-DNA insertion (picture modified from SIGNAL website)

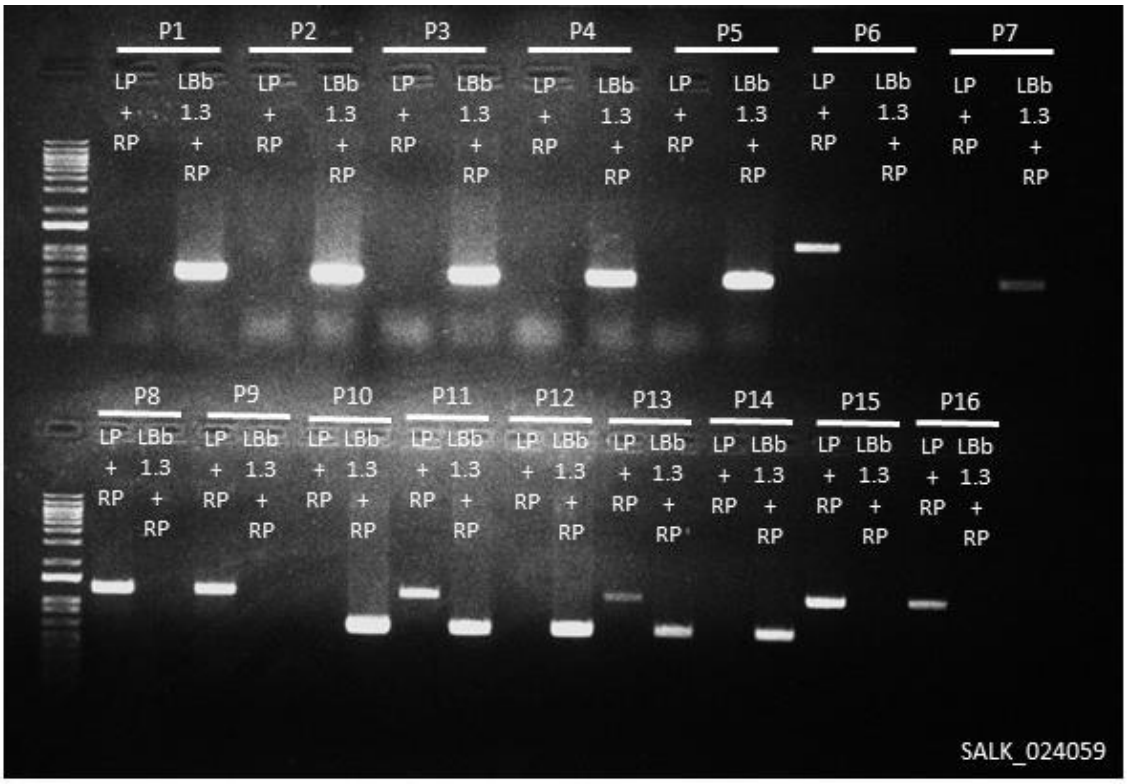


Figure 13.5. Electrophoretic agarose gel showing some of the results achieved with genotyping analysis on SALK_024059 line. The samples are named with a code composed by "P" (for "plant") and a progressive number from 1 to 16. Each sample includes two wells, one for LP + RP PCR product and the other for LBb1.3 + RP. Taking advantages of what explained in paragraph "5.2.10 *Arabidopsis* mutant lines complementation", from this picture it is evincible that the plants P1, P2, P3, P4, P5, P7, P10, P12 and P14 are homozygous and can be used for complementation purposes. On the other hand, the plants P11 and P13 are heterozygous, and the plants P6, P8, P9, P15 and P16 are wild type.

5.2.11 Transient overexpression in *Nicotiana benthamiana* (modified from Espley et al., 2007)

REAGENTS

Infiltration buffer (pH 5.6)	
Reagent	**Concentration**
MES	10 mM
$MgCl_2$	10 mM

Acetosyringone (ACS)*	200 µM
milliQ H_2O	up to volume

*Add just before use

PROTOCOL

I. Grow *Nicotiana benthamiana* with long photoperiod conditions for 5 weeks: the best developmental stage for the agroinfiltration od the plant is the 5-6 leaves one (Fig. 14.5)
II. Following the paragraph "5.2.4 *Agrobacterium tumefaciens* transformation", grow a recombinant *Agrobacterium* C58 suspension. Always check the culture with a colony PCR before to use
III. Centrifuge the suspension at 4000 rpm for 10 minutes
IV. Remove supernatant and resuspend the cellular pellet in 1 ml of Infiltration buffer
V. Measure the OD_{600} and dilute the culture with Infiltration buffer to obtain 0.2 values
VI. Incubate the suspension at room temperature for 2 hours without shaking
VII. With a needle, lightly scratch the lower leaf surface
VIII. Place the needleless tip of an insulin syringe on the scratch and inject the *Agrobacterium* suspension until all the leaf mesophyll is completely impregnated (Fig. 15.5)
IX. Put the plants for 48 hours in long photoperiod conditions
X. Sample the infiltrated leaves and put a piece of it on a microscope slide to investigate the subcellular localization with a confocal microscopy approach

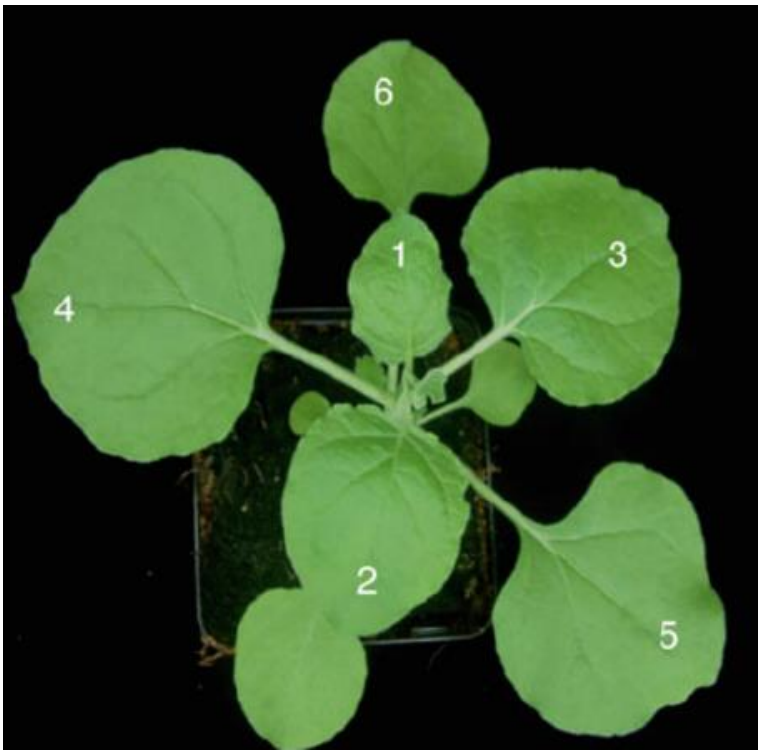


Figure 14.5. Optimal growth stage of *Nicotiana benthamiana* for agroinfiltration purposes (Photograph property: TOMSBiolab)

Figure 15.5. A just infiltrated leaf of *Nicotiana benthamiana*

5.3 Results

5.3.1 Subcellular localization

To investigate the localization of *VviMYB108A* and *VviMYB108B* in cells, we fused GFP to the N-terminus of the ORFs of both the TFs, and the construct was agroinfiltrated into leaves of wild type *N. benthamiana*. The expression of the GFP:*VviMYB108A-B* in leaf epidermal cells was examined by confocal microscopy by exciting with a laser line at 488 nm. Results showed that the GFP:*VviMYB108A-B* localized exclusively to the nucleus. The green fluorescent signal from the GFP:*VviMYB108A-B* was overlapped with the red fluorescent signal from the chloroplasts and the bright field picture in order to give back a complete overview of the cell (Fig. 16.5).

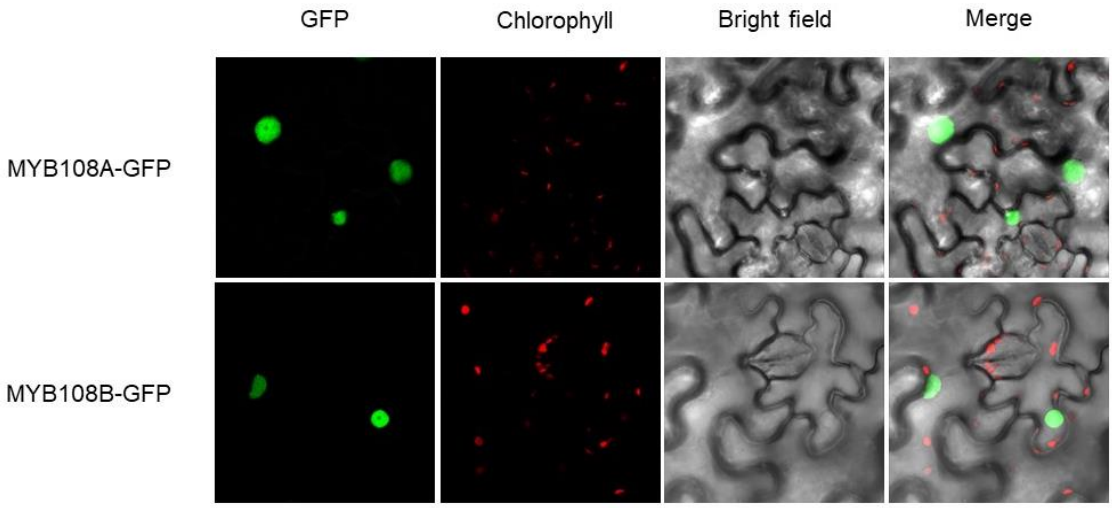


Figure 16.5. Subcellular localization of green fluorescent protein (GFP) fused to *VviMYB108A* and *B* ORFs under the control of strong promoter CaMV 35S, transiently overexpressed in *Nicotiana benthamiana* leaves (as shown in Fig. 1.5). Images in the first column show cells with GFP signal. Images in the second column show the same cells but with the red signal naturally emitted by chlorophyll at 488 nm wavelength. Images in the third column show the bright-field view of the same cells and the images in the fourth column are the overlays of the bright-filed and fluorescent images.

5.4 Future developments

Since the time for plant growth ad regeneration was not compatible with the established duration of PhD activities, here the results of the several functional characterization assays performed have not been presented, apart from the subcellular localization and a few concerning the complementation of the mutant lines of *Arabidopsis thaliana*. The subsequent activities will regard the phenotyping analysis of the stable overexpressing lines of tomato and *Arabidopsis*, and of this last also considering the complemented mutants. Since the mutant lines in first instance seemed of loss-of-function type (Cui et al., 2013) and exactly in the middle of our functional characterization analysis it was finally assessed the opposite (Cui et al., 2022), the expected results should be different from the previously predicted. In this regard, if a gain-of-function mutation showed reduced male fertility associated with delayed anther dehiscence, reduced pollen viability and a global decreased fecundity compared to the wild type (Mandaokar and Browse, 2009), these phenotypic traits – nevertheless at theorical level - are expected to be exacerbated in the complemented mutants, which more than "complemented mutants" are most suitably definable "double overexpressing lines". A middle situation between these "double overexpressing lines" and the wild type should be observable in the stable overexpressing transformant lines of tomato and *Arabidopsis*, theoretically the most similar to what reported for the gain-of-function mutants described by Mandaokar and Browse (2009). To be honest, the situation is more complex. Indeed – for completeness – is necessary to evaluate the cooccurrence of both the overexpression of the endogenous *AtMYB108* and of the heterologous *VviMYB108A* and *B*, also considering the possibility of plant endogenous (trans-) gene silencing mechanisms (Fagard and Vaucheret, 2000) that would make harder the study of a direct and linear relationship between an overexpression level and a given phenotype. The same reasoning could be extended also to the *Botrytis cinerea* susceptibility issue. Pathogenicity infection tests with *Botrytis cinerea* are already planned and, from a speculative point of view, the "double overexpressing lines" are expected to be the most susceptible to the fungus - showing the highest level of uncontrolled PCD - if compared to the wild type and to the stable transformant lines, which supposedly are going to be like the ones described by Cui et al. (2013). But – again – there are other mechanisms affecting the system that are needed to be considered for the global evaluation. Certainly, a transcriptomic approach - together with a qPCR directed on the overexpressed target – would greatly help to make clearness on this nested and

partially imbricated landscape, but also avoiding in the future the risk of confusion in the definition of the mutant lines as instead happened for the considered case.

BIBLIOGRAPHIC REFERENCES

Cui F., Li X., Wu W., Luo W., Wu Y., Brosché M., Overmyer K. Ectopic expression of *BOTRYTIS SUSCEPTIBLE1* reveals its function as a positive regulator of wound-induced cell death and plant susceptibility to *Botrytis*, *The Plant Cell*, Volume 34, Issue 10, October 2022, Pages 4105–4116, https://doi.org/10.1093/plcell/koac206

Cui F., Brosché M., Sipari N., Tang S., & Overmyer K. (2013). Regulation of aba dependent wound induced spreading cell death by myb 108. New Phytologist, 200(3), 634-640.

Espley R.V., Hellens R.P., Putterill J., Stevenson D.E., Kutty-Amma S. and Allan A.C. (2007). Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. Plant J. 49 : 414–427.

Fagard M. & Vaucheret H. (2000). (Trans) gene silencing in plants: how many mechanisms?. Annual review of plant biology, 51, 167.

Gambino G., Ruffa P., Vallania R., & Gribaudo I. (2007). Somatic embryogenesis from whole flowers, anthers and ovaries of grapevine (Vitis spp.). Plant cell, tissue and organ culture, 90(1), 79-83.

Ishizaki K., Nishihama R., Ueda M., Inoue K., Ishida S., Nishimura Y., Shikanai T., Kohchi T. Development of Gateway Binary Vector Series with Four Different Selection Markers for the Liverwort *Marchantia polymorpha*. PLoS One. 2015 Sep 25;10(9):e0138876. doi: 10.1371/journal.pone.0138876. PMID: 26406247; PMCID: PMC4583185.

Kikkert J. R., Striem M. J., Vidal J. R., Wallace P. G., Barnard J., & Reisch B. I. (2005). Long-term study of somatic embryogenesis from anthers and ovaries of 12 grapevine (*Vitis* sp.) genotypes. *In Vitro Cellular & Developmental Biology-Plant*, *41*(3), 232-239.

McElver J., Tzafrir I., Aux G., Rogers R., Ashby C., Smith K., Thomas C., Schetter A., Zhou Q., Cushman M. A., Tossberg J., Nickle T., Levin J. Z., Law M., Meinke D., Patton D. Insertional mutagenesis of genes required for seed development in Arabidopsis thaliana. Genetics. 2001 Dec;159(4):1751-63. doi: 10.1093/genetics/159.4.1751. PMID: 11779812; PMCID: PMC1461914.

Mandaokar A. and Browse A. MYB108 Acts Together with MYB24 to Regulate Jasmonate-Mediated Stamen Maturation in Arabidopsis, *Plant Physiology*, Volume 149, Issue 2, February 2009, Pages 851–862, https://doi.org/10.1104/pp.108.132597

Mengiste T., Chen X., Salmeron J., Dietrich R. The *BOTRYTIS SUSCEPTIBLE1* Gene Encodes an R2R3MYB Transcription Factor Protein That Is Required for Biotic and Abiotic Stress Responses in Arabidopsis, *The Plant Cell*, Volume 15, Issue 11, November 2003, Pages 2551–2565, https://doi.org/10.1105/tpc.014167

Perrin M., Gertz C. & Masson J. E. (2004). High efficiency initiation of regenerable embryogenic callus from anther filaments of 19-grapevine genotypes grown worldwide. *Plant Science*, *167*(6), 1343-1349.

Stamp J. A. & Meredith C. P. (1988). Somatic embryogenesis from leaves and anthers of grapevine. Scientia Horticulturae, 35(3-4), 235-250.

Wong D. C. J., Schlechter R., Vannozzi A., Höll J., Hmmam I., Bogs J., ... & Matus J. T. (2016). A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family uncovers new insights into the regulation of stilbene accumulation. *DNA research*, *23*(5), 451-466.

Xu R. and Qingshun L. Q. Protocol: Streamline cloning of genes into binary vectors in *Agrobacterium* via the Gateway® TOPO vector system. *Plant Methods* **4**, 4 (2008). https://doi.org/10.1186/1746-4811-4-4

Zhang X., Henriques R., Lin SS. et al. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. Nat Protoc 1, 641–646 (2006). https://doi.org/10.1038/nprot.2006.97

Paper published in Journal of Experimental Botany (https://doi.org/10.1093/jxb/erac487)

CHAPTER VI: Past, present and future of genetic strategies to control tolerance to the main fungal and oomycete pathogens in grapevine

Carlotta Pirrello[1#], Gabriele Magon[1#], Fabio Palumbo[1,2], Silvia Farinati[1], Margherita Lucchin[1,2], Gianni Barcaccia[1,2], Alessandro Vannozzi[1,2*]

[1] Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Campus of Agripolis, Viale dell'Università 16, 35020 Legnaro (PD), Italy.

[2] Centro Interdipartimentale per la Ricerca in Enologia e Viticultura (CIRVE), University of Padova, Via XXVIII Aprile 14, 31015 Conegliano (TV), Italy.

[*]Correspondence: alessandro.vannozzi@unipd.it

[#]Equal contribution as first author

6.1 Abstract

The production of high-quality wines is strictly related to the correct management of the vineyard, which guarantees good yields and grapes with the right characteristics required for subsequent vinification. Winegrowers face a variety of challenges during the grapevine cultivation cycle: the most notorious are fungal and oomycete diseases such as downy mildew, powdery mildew, and gray mold. If not properly addressed, these diseases can irremediably compromise the harvest, with disastrous consequences for the production and wine economy. Conventional defense methods used in the past involved the use of chemical pesticides. However, such approaches are in conflict with the growing attention on environmental sustainability and shifts from the uncontrolled use of chemicals to the use of integrated approaches for crop protection. Improvements in genetic knowledge and the availability of novel biotechnologies have created new scenarios for possibly producing grapes with a reduced, if not almost zero, impact. Here, the main approaches used to protect grapevines from fungal and oomycete diseases are reviewed, starting from conventional breeding, which allowed the establishment of new resistant varieties, followed by biotechnological methods, such as transgenesis, cisgenesis, intragenesis and genome editing, and ending with more

recent perspectives concerning the application of new products based on RNA interference (RNAi) technology. Evidence of their effectiveness, as well as potential risks and limitations based on the current legislative situation, are critically discussed.

Keywords: *Cisgenesis, Downy mildew, dsRNA, Genome editing, Gray mold, Powdery mildew, Vitis.*

6.2 Introduction

Currently, grapevine cultivation has become enormously important both at economic and food production levels on a global scale. In 2019, according to the economic note of the world wine sector drawn up by the International Organization of Vine and Wine (OIV), the world area under vines, i.e., the total area planted with vineyards for all uses (wine, table grapes and raisins), was estimated at 7.4 million hectares. In the following year, the area covered by grapes for use in worldwide wine production was estimated at approximately 258 million hectoliters (OIV, 2020), with approximately 159 million hectoliters produced in Europe and 99 million in the rest of the world. Italy is the world's leading wine producer, with a share of 18%, equal to 47.2 million hectoliters. Wine, understood as the main product of vine cultivation, is inextricably linked to the complex interactions of the vine with the surrounding environment. Moreover, the high quality and excellence of a wine often reflect the genetic characteristics typical of a single genotype, making it necessary to avoid the recombination of characters and therefore sexual reproduction. The vine, consequently, is identified as a static genetic entity, basically represented by one or a few genotypes perfectly adapted to a territory and multiplied vegetatively by means of cuttings or clones. The large global cultivation area and the need for genetic standardization within one or more territories make the vine particularly susceptible to attack by pathogens and abiotic stresses. In addition, the threat of pathogens in a purposely static genetic context can be intensified with the phenomenon of climate change, leaving room for emerging phytosanitary problems, as well as increasingly frequent and highly unfavorable atmospheric extremes affecting crop fitness. Next-generation viticulture will need to consider environmental trends toward high temperature. In addition to reviewing the traditional "suitability" of different territories, the geographical redistribution of the vines and their adaptation to cultivation techniques, it will be necessary to intensify the contribution that genetics and biotechnologies can provide to solve phytosanitary and qualitative problems linked to defense. At the end of the nineteenth century, mildews and phylloxera (*Daktulosphaira vitifoliae*) from North America put the entire European supply chain in jeopardy. Still today, some of the most disruptive grape diseases in the Eurasian region are downy mildew caused by the oomycete *Plasmopara viticola* and powdery mildew , and gray mold caused by the ascomycetes *Erysiphe necator* and *Botrytis cinerea*, respectively (Elad *et al.*, 2016; Buonassisi *et al.*, 2017; Pirrello *et al.*, 2019). Genetic systems have already been widely exploited to confer desired characteristics and to counter impending problems. The most emblematic example is the development and use of rootstocks. The root system of the European grapevine, in fact, was found to be highly vulnerable to infestation by

phylloxera, while the wild species of American grapevine showed resistance to aphid attacks (Forneck *et al.*, 2001). Initially, attempts were made to combat the insect by hybridizing European and American species to obtain varieties resistant to phylloxera. However, the obtained hybrids did not offer total resistance and did not adapt well to the calcareous, poor, and dry soils that characterize viticulture. Furthermore, the product obtained from these hybrids showed a lower organoleptic quality compared to European varieties. Therefore, the best solution was to select rootstocks derived from interspecific crosses between some *non-vinifera* species, mainly *V. berlandieri*, *V. riparia* and *V. rupestris*, or between these and *V. vinifera*. This intense hybridization work led to the selection of many rootstocks that are still used today and guarantee resistance of the root system to phylloxera as well as greater adaptability to all soil and climatic conditions. The fight against phylloxera represented, in the context of vine cultivation, the trigger toward a growing importance in the selection of rootstocks, which had to confer resistance not only to the aphid but also to the main biotic agents, especially nematodes (Wilcox *et al.*, 2015) and related viruses (Walker *et al.*, 1994), and also needed to show good adaptability to stressful abiotic conditions, such as deficiencies in the absorption and transport of water and mineral elements (Marguerit *et al.*, 2012; Meggio *et al.*, 2014; Corso and Bonghi, 2014; Vannozzi *et al.*, 2017). The need for rootstocks, and more generally for resistant cultivars, also stemmed from a problem directly related to phytosanitary defense or to improvements in plant fitness, i.e., the use of insecticides, nematicides, soil fumigants and chemical fertilizers, the indiscriminate use of which can lead to water contamination and environmental toxicity. For these reasons, chemical strategies for disease management are increasingly regulated by the EU to alleviate the potential negative effects of pesticides on workers' health and field biodiversity and to promote initiatives that fall within the broader framework of integrated pest management (IPM; Pertot *et al.*, 2017).

In this regard, grapevine breeding remains a valuable tool aimed not only at scavenging and exploiting natural resistance traits but also at coupling them with the market-dictated fruit quality and agronomical characteristics of elite varieties (Topfer *et al.*, 2011). In fact, the cultivation of disease-resistance genotypes has made it possible to limit the number of pesticide treatments in viticulture in many areas of the world (Topfer and Hausmann, 2021). Nonetheless, the increasingly widespread development of these cultivars raises several important issues related to their qualitative potential and marketing and, more importantly, to the management of the durability of resistance, with several cases of the erosion or breakdown of resistance already reported in Europe (Guimier *et al.*, 2019). In recent decades, the development of new biotechnological approaches such

as cisgenesis, genome editing, and RNA interference has expanded the range of tools that can be used to improve the tolerance to pathogens in difficult crops such as grapevine. However, it should be noted that, similar to conventional breeding, these new biotechnological techniques also have advantages and disadvantages that must be taken into consideration. In 2017, the EU Explanatory Note (EU Commission and Directorate-General for Research and Innovation, 2017) introduced the definition of conventional breeding technologies, consisting of traditional breeding and marker (MAB) and genomic assisted breeding (GAB), to be considered separate from these new breeding biotechnologies (NBTs).

This study aims to review the contribution of conventional and new breeding techniques to the genetic improvement of grapevine, starting with MAB, moving on to the use of transgenesis, cisgenesis, and posttranscriptional silencing mechanisms and ending with genome editing. Parallel to the description of the main results obtained from the research in this area, an attempt was also made to provide the regulatory framework at a national and European level.

6.3 Conventional genetic improvement of grapevine for fungal and oomycete diseases

Cross breeding exploits the genetic segregation that occurs through controlled sexual reproduction to obtain progeny with broad genetic diversity. It was first used in grapevine in the 18th century, when unsuccessful attempts to cultivate *V. vinifera* subsp. *Vinifera* in America finally resulted in the development of "American hybrids" adapted to local climatic conditions with better characteristics for winemaking than the local wild species (Topfer *et al.*, 2011). In Europe, cross breeding for resistance started at the beginning of the 20th century to cope with emerging pests coming from the New World, such as downy mildew, powdery mildew, and gray mold. The first interspecific cross attempts were conducted by private French breeders leading to so-called "first-generation hybrids" or "direct producers" that were ungrafted. Given their process of coevolution with pathogens, American wild species were considered resistant sources (Toffolatti *et al.*, 2018), but unfortunately, they were not suitable for wine production in terms of quality (Eibach *et al.*, 2010) (Box 1). In fact, although this technique proved effective for rootstocks, the unsatisfactory viticultural properties of the French hybrids forced the breeding programs in France to stop and led to the advent of laws that still rigorously regulate hybrids in the EU wine market today (Meloni and Swinnen, 2013). Nonetheless, subsequent breeding activities in other countries resulted in cultivars being successfully grown and vinified throughout Europe (Topfer *et al.*, 2011).

Given the need to make the most of genetic resources for resistance from wild species, gene introgression has been employed. This process consists of an interspecific hybridization between a wild individual carrying the resistance trait (donor plant) and an individual from a commercial variety (recipient plant). The recipient plant is then used for subsequent backcrosses coupled with the recurrent selection of the resistance trait. Unfortunately, as described by Dalla Costa *et al.* (2017), this strategy is labor- and time-consuming for a woody crop with a long juvenile phase such as grapevine, without the guarantee of a satisfactory result in agronomical terms.

MAB can partially overcome the aforementioned drawbacks and accelerate the process of selection of resistant cultivars obtained from interspecific crosses by tracing the introgression of the loci of interest (Di Gaspero and Cattonaro, 2010). In the '00s, a large collection of molecular markers associated with quantitative trait loci (QTLs) was developed to allow the construction of genetic maps (Hvarleva *et al.*, 2009). The first to be used were the simple sequence repeats (SSRs), which, despite their high mutation rate, are still the best selective markers for their stability and codominance (Adam-Blondon *et al.*, 2004). Over the years, many other highly polymorphic markers have been exploited (Vezzulli *et al.*, 2019*a*), such as random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP; Fischer *et al.*, 2004) and sequence characterized amplified regions (SCAR; Welter *et al.*, 2007). Given their ease of use, reproducibility, and abundance in the grapevine genome (Velasco *et al.*, 2007), single nucleotide polymorphisms (SNPs) have been used as integration tools for SSRs (Emanuelli *et al.*, 2013) and to create dense genetic maps (Vezzulli *et al.*, 2008).

As a consequence of the development and application of new molecular markers, in recent years, there has been a significant increase in the number of identified *R* loci as dominant sources of resistance. In this regard, North American species have contributed most to the collection of *R* loci thus far (Table 1). In recent years, the investigation of QTLs for resistance has expanded from *Muscadinia* through new American and Asian *Vitis* species up to *V. vinifera*. However, despite the large updated list of *R* loci (Table 1) against downy mildew, powdery mildew and gray mold, the exploitation of resistance traits in unexplored wild species is rapidly proceeding, but the search for new resistance genes is also needed in cultivated resistant varieties (e.g., *Rpv27* identified in Norton by Sapkota *et al.*, 2019). Moreover, the recent loosening of the regulation on the cultivation of hybrids, which now allows them to be classified with a Protected Geographical Indication (PGI; Council Regulation (EC), 2009), has given a new impetus to the development of resistant species from interspecific crossbreeding. These new cultivars are considered *V. vinifera* varieties and are

defined as pilzwiderstandsfähig (PIWI), meaning “fungal disease resistant”. The same definition is used by an international working group aimed at the promotion of fungus-resistant grape varieties via different breeding programs in Europe and North America (PIWI International). Currently, resistant varieties are widely cultivated in many regions of the world (Figure 1, Table S1). Among the top 30 resistant wine grape varieties, Concord is the most widespread and covers 0.24% of the worldwide area cultivated with vines; it is mostly cultivated in Brazil, the USA, and Japan. Similarly, Bianca covers 0.22% and is concentrated in Russia, Hungary, and Moldova. Seyval Blanc, Regent, Chambourcin, Villard Noir and Baco Noir follow with 0.06-0.02% of the worldwide share and covering countries on all five continents (Anderson and Nelgen, 2020) (Figure 1, Table S1, Table 2). Not much is known about the genetic traits that determine resistance in these cultivars. While Norton and Golden Muscat have only one *R* locus against *P. viticola,* Seyval blanc, Regent and Solaris carry five *Rpv* and *Ren* loci. Four loci for resistance to *E. necator* and *P. viticola* are carried by Cabernet Cortis, with three by Phoenix and Johanniter, and two by Bianca and Chambourcin (Table 2); no data are available for the other cultivars.

Regarding the wild and cultivated sources of resistance, one of the main issues of their investigation is that not all individuals belonging to a species show the same resistant phenotype, raising the need to test every single germplasm with local strains of the pathogen (Dry *et al.*, 2019). Moreover, as proven by Peressotti *et al.* (2010), resistance conferred by major QTLs cannot be considered a long-term solution since it is highly race-specific and can be overcome by more virulent strains of the pathogen; on the other hand, partial resistance is the result of an additive effect of many genes and is more durable but still prone to erosion (Stuthman *et al.*, 2007). The ability and time to overcome resistance by a pathogen are determined by several factors, such as its morphological and reproductive characteristics and the permanence of the crop in the field (Dry *et al.*, 2019). Stam and McDonald, (2018) calculated the combination of four *R* genes as a minimum to guarantee total impermeability of the resistance barrier in herbaceous crops, but it is not yet possible to establish whether these data are also valid for perennial crops such as grapevine and how much they should be combined with other control strategies (Feechan *et al.*, 2015; Zini *et al.*, 2019). In this regard, gene stacking (pyramiding) is proposed to improve breeding efficiency because it takes advantage of known resistance QTLs to downy mildew and powdery mildew (Miedaner, 2016; Mundt, 2018). Grapevine breeding programs aimed at stacking *R* loci are widespread throughout the world. In Europe, the collaboration between INRAE, Julius Kühn-Institut (JKI), Staatliches Weinbauinstitut (WBI) and Agroscope in the INRA-ResDur program (Merdinoglu *et al.*, 2018) made it possible to

release the four ResDur1 cultivars 'Voltis', 'Artaban', 'Vidoc', and 'Floreal', carrying *Run1, Ren3, Rpv1,* and *Rpv3.* This collaboration is still working on ResDur2 (with *Run1, Ren3.2, Rpv1,* and *Rpv10*) to reach ResDur3 cultivars with six *R* loci. Moreover, the Fondazione Edmund Mach (FEM) recently registered four new varieties ('Termantis', 'Nermantis', 'Valnosia', and 'Charvir') in the Italian grapevine variety catalog (Pecile and Zavaglia), and its breeding program produced progeny with individuals carrying from two to seven *R* loci to downy mildew and powdery mildew (Vezzulli *et al.*, 2019*c*). Furthermore, Foria *et al.* (2019), at the University of Udine, were able to combine two *Rpv* and two *Ren/Run* loci into "elite" parenting offspring, releasing 'Cabernet Volos', 'Fleurtai', 'Merlot Kanthus' and 'Soreli'. Overseas, Riaz *et al.* (2019) are working on combining up to three *Ren/Run* loci to develop advanced lines with a *V. vinifera* background. In Chile, Agurto *et al.* (2017) obtained *Ren1Run1Rpv1* table grape cultivars, and Saifert *et al. (* 2018) were able to stack *Rpv1* and *Rpv3.1* in elite genetic material through a collaboration between Ecuadorian and Brazilian universities.

The ability to move from a genetic map to a genetic sequence was certainly provided by the many genome sequencing works on *Vitis* cultivars (Velasco *et al.*, 2007; Jaillon *et al.*, 2007; Carrier *et al.*, 2012; Gambino *et al.*, 2017; Roach *et al.*, 2018; Minio *et al.*, 2019; Magris *et al.*, 2021; Foria *et al.*, 2022) and American species (Girollet *et al.*, 2019; Cochetel *et al.*, 2021). Since the first SNP detection in grapevine (Owens, 2003), the harnessing of SNPs in breeding has been inevitably coupled with systematically increasing the genetic information obtained through next-generation sequencing (NGS), third-generation sequencing (TGS; Varshney *et al.*, 2009; Kumar *et al.*, 2012), and cost-efficient amplicon sequencing (AmpSeq; Yang *et al.*, 2016; Pirrello *et al.*, 2021). The advent of genome wide association studies (GWAS) based on grouping shared genetic variants led to the development of 9K and 18K SNP chips (Myles *et al.*, 2010; Le Paslier *et al.*, 2013), which allowed extensive studies on population genomics, germplasm genetic diversity and linkage mapping (Myles *et al.*, 2011; Mercati *et al.*, 2016; Riaz *et al.*, 2018). Later, the more genetic diversity-focused restriction-site-associated DNA sequencing (RAD-seq) and genotyping by sequencing (GBS; Peterson *et al.*, 2014; Campbell *et al.*, 2015) techniques were established to separate the analysis from the bias of using a predetermined set of point markers, resulting in high-resolution maps with a large density of detected variant sites (Barba *et al.*, 2014; Hyma *et al.*, 2015; Marrano *et al.*, 2017).

6.4 Cisgenesis, intragenesis and transgenesis approaches for fungal and oomycete pathogen resistance in grapevine

Despite the availability of several resistance genes isolated within *Vitis* spp. (Table 3), the lack of efficient promoters and selection markers made the application of cisgenesis (Box 2) in this genus very difficult (Limera *et al.*, 2017). This would explain why cisgenesis attempts aimed at improving fungal and oomycete tolerance in *V. vinifera* are rare. Based on our literature, only Dhekney *et al.* (2011) claimed the successful application of cisgenesis approaches in grapevine through the transfer of the *VvTL-1* gene (encoding a thaumatin-like protein with broad spectrum antifungal activity) from Chardonnay to Thompson. Some cisgenic lines regenerated from somatic embryos showed a delay of 7-10 days in the development of powdery mildew symptoms in greenhouse conditions. However, although the authors recognize it as the first report of a cisgenic approach for fungal and oomycete disease resistance, the resulting engineered plants were known to contain viral promoters as well as reporter/marker genes. In reality, based on the strictest definitions, no study has successfully developed fungal-resistant cisgenic or intragenic grapevines, and the aforementioned study could be placed halfway between cisgenesis and transgenesis. Similarly, other studies have used a hybrid approach based on both stable and transient transformation with genes derived from the same species or sexually compatible taxa, as well as viral promoters (i.e., CaMV35S), bacterial selection markers (i.e., *nptII* and *hptII*) and reporter genes (i.e., GFP).

For example, full or increased tolerance to powdery mildew was observed by transferring genes isolated from *V. quinquangularis*, *V. vinifera* and *V. pseudoreticulata* into *V. vinifera*. These genes belong either to the stilbene synthase family (*VpSTSgDNA2* and *VqSTS6*; Dai *et al.*, 2015; Cheng *et al.*, 2016) or to the pathogenesis-related (PR) protein pathway (*VpPR4-1*, *VvNPR1.1* and *VpEIFP1*; Le Henanff *et al.*, 2011; Dai *et al.*, 2016; Wang *et al.*, 2017).

Stilbenes represent phenolic compounds that accumulate at the infection site of the main fungal and oomycete pathogens and show antimycotic activity by inhibiting conidial germination (Hasan and Bae, 2017), while PR proteins are a class of soluble proteins that are generally expressed following abiotic and biotic stresses (Ali *et al.*, 2018). PR proteins are also involved in increased resistance to *P. viticola,* as observed by engineering *V. vinifera* with genes isolated from *V. pseudoreticulata* (*VpPR10.1*; Su *et al.*, 2018). Improved resistance to *P. viticola* was also observed in *V. vinifera* transformed with *VaHAESA*, a gene isolated from *V. amurensis* that is highly similar to pattern recognition receptor (PRR)-like kinase 5 (RLK5; Liu *et al.*, 2018). Finally, the only hybrid approach (i.e., combining cisgenesis and transgenesis elements) aimed at reducing the susceptibility

of grapevine to *Botrytis cinerea* included transforming 41B rootstock and cv. Sugraone, with the stilbene synthase coding gene *VST1* under the control of the alfalfa PR 10 and 35S promoters, respectively. A strong reduction in symptoms was linked in both cases to the accumulation of resveratrol (Coutos-Thévenot *et al.*, 2001; Dabauza *et al.*, 2015).

Regarding transgenic approaches, the chitinase family has aroused enormous interest due to its involvement in the response to *E. necator*. Two rice chitinase genes (*Chi11* and *RCC2*) were effective in reducing the effects of this pathogen in transgenic lines of *V. vinifera* (Yamamoto *et al.*, 2000; Nirala *et al.*, 2010). Acquired resistance to the abovementioned fungus was also obtained by engineering grapevine with both an *ech42*-nag70 double gene construct (i.e., a 42 kDa extracellular endochitinase combined with an N-acetyl-β-D-hexosaminidase) isolated from *Trichoderma* spp. and the *RESISTANCE TO POWDERY MILDEW 8* (*RPW8.2*) gene isolated from *Arabidopsis thaliana* (Rubio *et al.*, 2015; Hu *et al.*, 2018). Reduced rates of lesion expansion caused by *B. cinerea* were observed by transforming *V. vinifera* with a gene isolated from *Pyrus* spp. and coding for a polygalacturonase-inhibiting protein (pPGIP; Agüero *et al.*, 2005). Finally, a chitinase gene isolated from scab-infected Sumai-3 wheat and introduced into cv. Crimson Seedless produced enhanced tolerance to *P. viticola* infection (Nookaraju and Agrawal, 2012). The development and application of transgenesis and cisgenesis strategies are linked to several technical and biological advantages and disadvantages. Moreover, these strategies often give rise to biosafety issues and public concerns. All these aspects are summarized in Box 5.

6.5 Application of CRISPR/Cas9 to improve tolerance to the main fungal pathogens in grapevine

A first example of the application of the CRISPR/Cas9 system (Box 3) in grapevine involved the *Agrobacterium*-mediated transformation of Chardonnay cell suspensions with a single plasmid containing a specific sgRNA able to target the *L-idonate dehydrogenase* (*IdnDH*) gene to alter the biosynthetic pathway of tartaric acid (Ren *et al.*, 2016). Later, the gene encoding *phytoene desaturase* (*VvPDS*) was effectively knocked out in embryogenic calli of Neo Muscat (Nakajima *et al.*, 2017), Chardonnay and 41B rootstock by means of CRISPR/Cas9 binary vectors (Ren *et al.*, 2019). The first attempt to edit the grapevine genome to increase resistance against *B. cinerea* was performed on the Thompson Seedless variety, targeting the transcription factor (TF) *VvWRKY52*. Induced mutations in this gene led to a significant reduction in the development of fungal colonies, especially in biallelic mutant lines (Wang *et al.*, 2018). In further studies, CRISPR/Cas9 technology

was used to obtain different mutant lines of *V. vinifera* characterized by a lack of expression of *Downy Mildew Resistance 6* (*DMR6*) and *Mildew Locus O* (*MLO*). *DMR6* and *MLO* represent susceptibility (S) genes toward *P. viticola* and *E. necator*, and their knockout led to a significant increase in resistance to these fungal and oomycete pathogens (Giacomelli *et al.*, 2019, 2022). The efficiency of CRISPR/Cas9-mediated mutagenesis in grapevine depends on factors common to other genetic transformation methods (technical methods, plant genotype, target gene, *in vitro* regeneration, and selective conditions) and on factors specific to this particular type of approach, such as the choice of the upstream Cas9 promoter and the sgRNA sequence (Ma *et al.*, 2015). With reference to this last point, CRISPR/Cas9-mediated genome editing in grape was first reported in 2016 (Ren *et al.*, 2016), and since then, the *Arabidopsis* AtU6 or AtU3 promoters have been used to regulate the expression of sgRNAs (Ren *et al.*, 2016; Nakajima *et al.*, 2017; Wang *et al.*, 2018; Osakabe *et al.*, 2018; Giacomelli *et al.*, 2019). Recently, Ren *et al.* (2021) identified four VvU3 and VvU6 promoters and two ubiquitin (UBQ) promoters in grapevine and demonstrated that the use of the identified VvU3/U6 and UBQ2 promoters could significantly increase the editing efficiency in grape by improving the expression of sgRNA and Cas9, respectively. CRISPR components can be delivered within the cell in different ways: (i) in the form of nucleic acids, with genes coding for Cas9 and gRNA delivered using vectors, (ii) with genes coding for gRNA cloned within vectors and a recombinant Cas9 protein, and (iii) in the form of a ribonucleoprotein complex (RNP), constituted *in vitro* with Cas9 protein and an *in vitro*/*in vivo* transcribed gRNA (Glass *et al.*, 2018). The major bottleneck in grapevine editing is that all the applications described thus far still exploit *Agrobacterium*-mediated genetic transformation using nucleic acids (i.e., DNA/mRNA coding for the entire system). Consequently, these obtained varieties would not overcome the limits highlighted for transgenic varieties (Capriotti *et al.*, 2020) (for a more detailed description of the advantages/disadvantages of genome editing see Box 5). Clearly, a delivery strategy based on the use of the RNP complex represents a valid methodology to edit the genome without the introgression of exogenous DNA. At the same time, however, an approach of this type cannot be accompanied by an *Agrobacterium*-mediated transformation, hence the need to develop or implement new methods to associate the use of RNP, which does not involve integration of nonhost DNA into the edited species, into an efficient, inexpensive, and fast transformation system. Among the best candidate delivery systems are particle gun or polyethylene glycol (PEG)-mediated transfection. A first example of the direct delivery of CRISPR–Cas9 ribonucleoproteins (RNPs) in grapevine protoplasts to obtain efficient DNA-free targeted mutations was achieved by Malnoy *et*

al. (2016) in Chardonnay. In this study, protoplasts obtained from embryogenic calli were edited, and a site-specific mutation was determined at the level of the *MLO7* gene locus, which is one of the susceptibility factors to *E. necator*. A detailed protocol adapting the CRISPR–Cas9 system to grapevine plants using the direct delivery of CRISPR–Cas9 ribonucleoproteins (RNPs) to achieve efficient DNA-free targeted mutations in grapevine protoplasts was described two years later by Osakabe et al. (Osakabe *et al.*, 2018). Clearly, one of the major impediments to the use of protoplasts/RNPs in genome editing approaches lies precisely in the difficulty encountered in the regeneration of plants starting from single cells. Numerous efforts are being made in this field to alleviate this bottleneck to applying genome editing successes to plants. Recently, a successful DNA-free methodology was used to obtain fully edited grapevine plants regenerated from protoplasts obtained from a *V. vinifera* cv. Crimson seedless L. embryogenic callus. The transfected protoplasts were edited on the downy mildew susceptibility gene *VvDMR6-2* (Scintilla *et al.*, 2021). Excluding the protoplast-based approach, in all other cases, the RNP complex needs to cross a substantial barrier, the cell wall, whose structure and composition are complex and tissue-specific and can vary over time, making it a complicated structure to be crossed. Nanoparticles were shown to deliver different molecules (DNA, RNA, and protein) to mammalian cells. Interestingly, increased interest into their use in agriculture occurred with relative success in delivering different molecules to normal plant cells with cell walls. The scientific community currently pays much attention to a particular class of nanoparticles, the cell-wall penetrating peptide (CPP), which seems to be less cytotoxic to plants than other nanoparticles, such as gold or silicone nanoparticles. CPPs consist of short peptide sequences (5 to 30 amino acid residues) that facilitate cargo penetration through the cell membrane. Recent studies have demonstrated their efficacy in penetrating intact plant cells, and they can help deliver large cargo molecules such as proteins, DNA, and RNA (Numata *et al.*, 2018). The development of a methodology for CPP-mediated delivery of a CRISPR/Cas9 system to a regenerable grapevine plant material will be a game changer because it will align with the pursuit of generating edited grapevine material that is GMO-free.

6.6 RNA interference mechanisms and potential applications in grapevine defense

To date, conventional RNAi approaches (Box 4) have relied on the use of weakened recombinant plant viruses (the so-called virus-induced gene silencing, VIGS), *Agrobacterium*-mediated transiently expressed transgenes, and stably transformed transgenic plants that express dsRNAs to silence

specific genes that control target traits (known as host-induced gene silencing, HIGS). Despite their evident success, the use and trading of RNAi-based transgenic crops as genetically modified organisms (GMOs) has aroused widespread concern and criticism from the public surrounding their long-term consequences on human health and the environment. In this regard, the development of low- or nonpathogenic virus expression vectors for the application of VIGS in crops poses technical and safety challenges. The PTGS mechanism is still poorly explored in grapevine compared to other crops; however, these silencing mechanisms have been applied to different sides of the host–pathogen interaction, targeting both host and pathogen sequences. At the pathogen level, *B. cinerea* was observed to silence immune response genes in *V. labrusca* producing sRNAs through the DCL system to establish infection. In this regard, a PTGS strategy was developed based on *BcDCL1* and *BcDCL2* as pathogen target genes, obtaining a "silencing of the silencing", which showed reduced growth of the fungi with consequent pathogenicity mitigation (Wang *et al.*). A similar approach was used for downy mildew management by targeting *PvDCL1* and *PvDCL2*, the genes responsible for the silencing machinery of *P. viticola* during grapevine infection (Brilli *et al.*, 2018).

In addition to the host–pathogen relationship determinants, an infection can also be contained by directly leveraging the molecular mechanisms underlying the vital processes of the pathogenic organism. Indeed, an emblematic case is the one carried out by Nerva *et al.* (2020) in *V. vinifera* cv. Moscato, in which three *B. cinerea* target genes were identified: *BcCYP51* (or *erg11*), *Bcchs1* and *BcEF2b*. *erg11* encodes a lanosterol 14α-demethylase belonging to the cytochrome P450 monooxygenase (CYP) superfamily that which controls a key step in the ergosterol biosynthetic pathway; it is a fungi-specific compound and the target of the main antifungal triazole-based products. Chitin synthase 1 is involved in chitin accumulation in the fungal cell wall, and its inactivation in *B. cinerea* results in stunted growth due to a weakened fungal cell wall (Soulié *et al.*, 2003). Finally, elongation factor 2 catalyzes ribosomal translocation and is the target gene of sordarin-derived commercial products. Independent of the application modality used, a collapse of *B. cinerea* virulence was observed in all the treatments compared to the control. From the host plant point of view, the *MLO* family contains the S genes to *E. necator* (Jorgensen, 1992). In monocots, the S genes phylogenetically belong to the IV clade (Reinstädler *et al.*, 2010) of the MLO family, whereas in dicots, they belong to the V clade (Feechan *et al.*, 2008). *MLO* S genes seem to negatively regulate the vesicle-associated and actin-dependent defense pathways at the *E. necator* penetration site (Panstruga, 2005). Of the four S genes identified in *V. vinifera* (*VvMLO6, VvMLO7,*

VvMLO11 and *VvMLO13*), only the last three are known to be upregulated during powdery mildew (Feechan *et al.*, 2008). A HIGS-based knockdown strategy showed a decrease in powdery mildew symptom severity up to 77% by silencing *VvMLO7* in combination with *VvMLO6* and *VvMLO11* in cv. Brachetto, with no pleiotropic effects observed (Pessina *et al.*, 2016). Another work on S genes involved targeting the *VvLBDIf7* grapevine gene, which encodes an LOB domain (LBD)-containing protein (Canaguier *et al.*, 2017). This gene is the putative ortholog of an LBD TF, a repressor of jasmonate-mediated defense mechanisms in *Arabidopsis* roots during *Fusarium oxysporum* infection. The disrupted function of this gene leads to enhanced resistance to pathogens (Thatcher *et al.*, 2012). A PTGS procedure on *V. vinifera* (cv. Pinot noir) gave satisfying results with a significant decrease in the growth and sporulation of *P. viticola* (Marcianò *et al.*, 2021).

Due to the strong restrictions on the main genetic engineering techniques, research has recently focused on the development of methods that do not involve modification of the host genome and, therefore, do not lead to the production of varieties considered genetically modified by the current legislation. For these reasons, exogenous RNAi-based approaches have emerged as a widely accepted and environmentally friendly strategy relying on the direct exogenous application of RNA molecules (dsRNAs and/or sRNAs) to improve crop qualities, with the potential to trigger RNAi in various plants regardless of their genetic backgrounds (for a more detailed description of the advantages/disadvantages of this approach see Box 5). Following the results that emerged in the application of various RNAi methods for the development of resistance by plants, the evidence of how the exogenous application of polynucleotides can influence the mRNA levels of important virulence-related genes of pathogens or plants has proven to be pivotal in the context of the development of new techniques and strategies for crop protection (Dubrovina and Kiselev, 2019). In detail, spray-induced gene silencing (SIGS) allows the absorption of dsRNA by cells and plant tissues, with the RNAi machinery being created by the host plant or directly conveyed at the cellular level of the pathogen, thus triggering gene silencing through the RNAi machinery of the pathogen itself (Sang and Kim, 2020). An experimental SIGS trial demonstrated how the exogenous application of dsRNA specific for the transcripts of the *B. cinerea DCL1/2* genes, which regulate the expression of Dicer proteins essential for the production of sRNA, on the leaf surface of *V. vinifera* led to effective absorption by the necrotrophic fungus, thus enhancing plant protection. Moreover, local application of *B. cinerea vacuolar protein sorting 51* (*VPS51*), *dynactin* (*DCTN1*), and *suppressor of actin* (*SAC1*)-targeted dsRNA particles to grape berries led to stunted mycelial growth and reduced susceptibility in the plant (Qiao *et al.*, 2021). Using SIGS, encouraging results against *B. cinerea* were

observed by Nerva *et al.* (2020) with postharvesting berry application, as well as in the abovementioned work on *PvDCL1* and *PvDCL2* (Haile *et al.*, 2021) and on *VvML*O (Marcianò *et al.*, 2021).

However, in the development of dsRNA-based RNAi strategies, several aspects must be considered: (i) the choice of the target gene; (ii) the delivery methods; and (iii) the formulation of the exogenous "inoculum". In addition to the abovementioned genes, members of different classes of transcripts can be identified as promising RNAi targets (Table 4). Among them is the ever-growing class of plant *S* genes, which are recessively inherited and less suitable for canonical breeding programs for resistance. S genes belong to different functional classes according to the role they play in plant–pathogen interactions (van Schie and Takken, 2014), such as TFs or genes catalyzing catabolic reactions. The major class of secreted effectors in *P. viticola* are the RxLR cytoplasmic effectors, which act inside the cell of the host plant (Toffolatti *et al.*, 2020). Among them, a few have been identified as PCD inhibitors shown to suppress the immune response in *Nicotiana benthamiana*. Something similar was observed when investigating the role of some *E. necator* RNase-like effector proteins expressed in haustoria (RALPH) effectors. Exploring these gene classes as well as those already scavenged in RNAi-targeted research represents a promising strategy. Moreover, regarding the target tissue, there are some variations in SIGS delivery methods, such as petiole absorption and trunk injection, with each having different efficiencies and resistance persistence; however, they still function and give promising results (Nerva *et al.*, 2020).

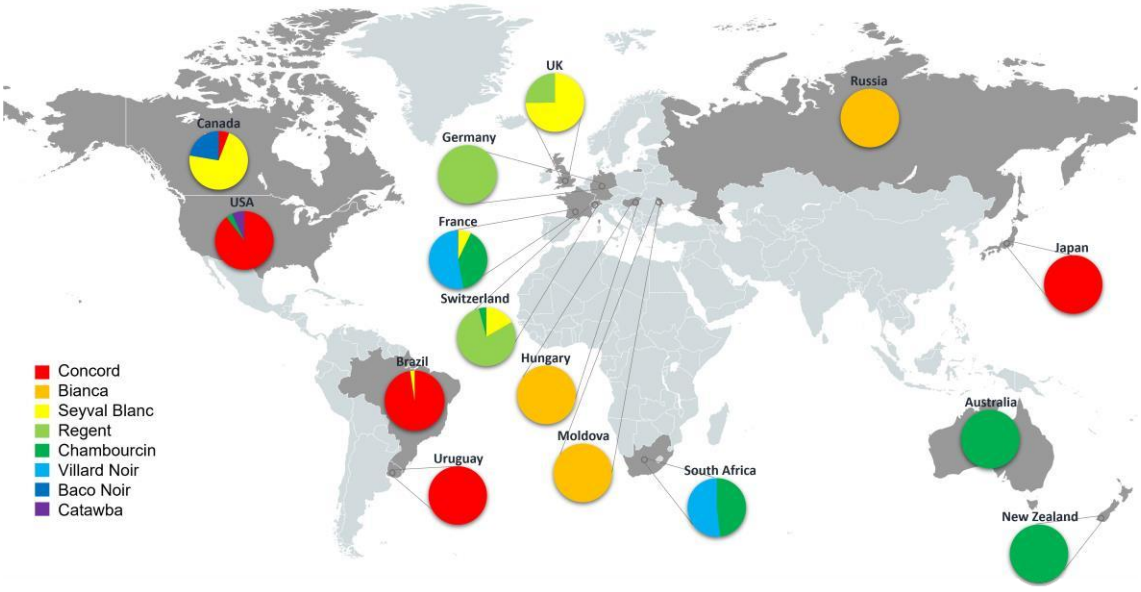


Figure 1.6. Distribution of the top resistant cultivated grapevine varieties. The map shows the distribution of the eight most resistant cultivated wine grape varieties throughout the world. Data were retrieved from Anderson and Nelgen (2020). Detailed information is provided in Table S1.

6.7 Conclusions

The coevolutionary interaction between host and parasite is often characterized by strong reciprocal selective pressures exerted by the two antagonists upon one another. If the interaction is long-lasting, then it may lead to fast evolutionary change (Woolhouse *et al.*, 2002). Although this phenomenon is common in plant–pathogen interactions in wild populations, it is not common in agriculture or viticulture. In fact, while the pathogen is subject to constant evolution, determined by the selective pressure exercised by chemical defense treatments, the plant host is not exposed to any evolutionary process due to strictly correlated commercial and propagative reasons. The absence of coevolution in viticulture led to a phenomenon of genetic stagnation at the level of elite cultivars, making them increasingly subject to attack by pathogens. Consequently, viticulture, while occupying 3% of the entire cultivated area, uses approximately 65% of all fungicides applied in agriculture, thus creating strong issues relating to the environmental, economic and social sustainability of the wine sector. In this scenario, plant breeding for genetic improvement can be extremely helpful for the development of new sustainable viticulture through the development of genotypes resistant to pathogens by exploiting the existence of resistance genes or the presence of genes for susceptibility in the cultivars of interest. However, at least for grapevine, the pursuit of this goal cannot be approached by means of the traditional methods of hybridization and selection alone, not only for purely technological reasons but also for commercial reasons. In fact, although the application of conventional breeding methods has already led to the development of improved cultivars from an oenological point of view that are sometimes characterized as having resistance to certain plant diseases, these represent completely new genotypes that have lost their commercial names and may possibly be subject to important existing regulations. If it is true that the expertise of grapevine breeders and conventional breeding methods will never be totally replaced by grapevine biotechnologists and biotechnological methods, it is also true that next-generation technologies available for molecular selection and genetic improvement will become increasingly important for the development of superior and resilient grapevine cultivars. The potential use of new biotechnological approaches for either loss-of-function and gain-of-function applications, such as cisgenesis (for conspecific gene transfer or introgression), genome editing (for endogenous gene knockout and gene editing or replacement) and RNAi-based methodologies (for targeted gene silencing), will assumes strong relevance in the future of viticulture, as these methods would allow direct intervention at the genomic level in any variety without changing its genetic background. All these genome-wide biotechnological strategies represent rapidly increasing precision breeding

methods that are potentially useful for developing cultivars genetically improved for single traits while preserving all other remaining traits. In addition, these methods have the benefit of reducing the time needed to breed a certain varietal genotype coupled with the advantage of avoiding any genetic recombination or transfer of undesirable genetic material. These are common aspects to all tree species as well as grapevines, and such technologies have been successfully applied in woody crops. Furthermore, in the coming years, it seems possible that biotechnological approaches could allow not only the transfer of cisgenes or the editing of genes of interest but also the control of whole biosynthetic pathways and regulatory networks of certain plant genomes, making the improvement of grapevine cultivars achievable by intervening in the development or composition of specific tissues and organs. Assuming that an accurate risk and environmental impact assessment is carried out for each individual genetic improvement intervention implemented through the application of new plant breeding techniques, the use of these methods in viticulture could concretely solve specific phytosanitary problems and, therefore, guarantee the safeguarding and competitiveness of grape-related products. To conclude, the economic and environmental sustainability of viticulture requires the design and development of programs that include conventional breeding and the most recently advanced breeding technologies in a synergistic way. A fundamental step to achieve this goal is the recognition of the indispensability of new plant breeding techniques by the national institutions responsible for information on organic farming and competent ministerial bodies (Box 6). Since their purpose is to guarantee a productive rural network that is less invasive from an environmental, economic, and social point of view, the potential use of these new technologies is absolutely in line with their goals.

Supplementary material

Supplementary material available at the link: https://doi.org/10.1093/jxb/erac487

Author contributions

AV conceived the review. CP, AV, GM, FP, AM, and SF wrote the manuscript. GB and ML supervised the final manuscript. All authors contributed to editing the manuscript.

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Founding

The work was founded by grants from the Italian Ministry of University and Research (MIUR), PRIN cod. 20178L3P38_004 and by the University of Padova, project BIRD 220029

6.8 Tables

Table 1. Resistance (*R*) loci to *B. cinerea, E. necator* and *P. viticola* in grapevine (Topfer and Hausmann, 2021 implemented) Together with the locus name, the chromosome (Chr), the resistance source (Source), the genotype, the associated marker, the geographical origin (NA: North America, CA: Central Asia, CH: China) and the Type of resistance (To: total, Mj: major, Pa: partial, Mn: minor), are listed.

Pathogen	Locus	Chr	Source	Genotype	Associated Marker	Origin	Type	Ref
B. cinerea	*Unnamed QTL*	2	*V. aestivalis*	Norton	VMC6F1, VMC3B10	NA	Mj	(Sapkota *et al.*, 2019)
E. necator	*Ren1*	13	*V. vinifera*	Kishmish vatkana	UDV020, VMC9h4-2, VMCNg4e10.1	CA	Mj/Pa	(Hoffmann *et al.*, 2008)
	Ren2	14	*V. cinerea*	Illinois 547-1	CS25	NA	Pa	(Dalbó *et al.*, 2001)
	Ren3	15	unknown	Regent	UDV015b, VVIv67 ScORA7-760 VChr15CenGen02 GF15-28 / VVIv67 GF15-42	NA	Pa	(Welter *et al.*, 2007) (Akkurt *et al.*, 2007) (van Heerden *et al.*, 2014) (Zyprian *et al.*, 2016) (Zendler *et al.*, 2017)
	Ren4	18	*V. romanetii*	C166-043 C87-41	VMC7f2 SNPs	CH	Mj/Pa	(Riaz *et al.*, 2012) (Mahanil *et al.*, 2012)
	Ren5	14	*M. rotundifolia*	Regale	VMC9c1	NA	Mj/To	(Blanc *et al.*, 2012)
	Ren6	9	*V. piasezkii*	DVIT2027	PN9-057, PN9-068	CH	Mj/To	(Pap *et al.*, 2016)
	Ren7	19	*V. piasezkii*	DVIT2027	VVIp17.1, VMC9a2.1	CH	Pa	(Pap *et al.*, 2016)
	Ren8	18	unknown		UDV117, SPS_P_SNP632GF	NA	Mn/Pa	(Zyprian *et al.*, 2016)
	Ren9	15	unknown	Regent	CenGen6	NA	Pa	(Zendler *et al.*, 2017)
	Ren10	2	unknown	Seyval blanc	S2_17854965 Haploblock validation	NA	Mn	(Teh *et al.*, 2017)
	Ren11	15	*V. aestivalis*	Tamiami	rh_chr15_15294725 rh_chr15_13822901 rh_chr15_13698923	NA	Pa	(Karn *et al.*, 2021)
	Run1	12	*M. rotundifolia*	VRH3082-1-42	VMC4f3.1, VMC8g9 49MRP1.P2, CB53.54	NA	Mj/To	(Barker *et al.*, 2005) (Feechan *et al.*, 2013)
	Run2.1	18	*M. rotundifolia*	Magnolia	VMC7f2, VMCNg1e3 VVIn16, VMC7f2, VMC7f2	NA	Mj/Pa	(Riaz *et al.*, 2011)
	Run2.2	18	*M. rotundifolia*	Trayshed	VMC7f2	NA	Mj/Pa	(Riaz *et al.*, 2011)
P. viticola	*Rpv1*	12	*M. rotundifolia*	28-8-78	VVIb32	NA	Pa	(Merdinoglu *et al.*, 2003)
	Rpv2	18	*M. rotundifolia*	8624		NA	Mj/To	(Wiedemann-Merdinoglu *et al.*, 2006)
	Rpv3	18	*V. rupestris*	Regent Bianca Regent	UDV112, UDV305. VMC7f2, VMC7f2	NA	Pa	(Welter *et al.*, 2007) (Bellin *et al.*, 2009) (van Heerden *et al.*, 2014)
	Rpv3.1 (=Rpv3$^{299\text{-}279}$)	18	*V. rupestris*	Seibel 4614 Villard blanc	UDV305, UDV737 GF18-06 , GF18-08	NA	Pa	(Di Gaspero *et al.*, 2012) (Zyprian *et al.*, 2016)
	Rpv3.2 (=Rpv3$^{null\text{-}297}$)	18	*V. rupestris* *V. lincecumii*	Munson GF.GA-47-42	UDV305, UDV737 GF18-06 , GF18-08	NA	Pa	(Di Gaspero *et al.*, 2012) (Zyprian *et al.*, 2016)
	Rpv3.3 (=Rpv3$^{null\text{-}271}$)	18	*V. labrusca* *V. riparia*	Noah Merzling	UDV305, UDV737 VVIN16, UDV737	NA	Pa	(Di Gaspero *et al.*, 2012) (Vezzulli *et al.*, 2019*b*)
	Rpv3$^{321\text{-}312}$	18	*V. labrusca* *V. riparia*	Noah	UDV305 UDV737	NA	Pa	(Di Gaspero *et al.*, 2012)
	Rpv3$^{361\text{-}299}$	18	*V. rupestris*	Ganzin	UDV305 UDV737	NA	Pa	(Di Gaspero *et al.*, 2012)
	Rpv3$^{299\text{-}314}$	18	*V. rupestris*	Ganzin	UDV305 UDV737	NA	Pa	(Di Gaspero *et al.*, 2012)
	Rpv3$^{null\text{-}287}$	18	*V. rupestris* *V. labrusca*	Bayard (Couderc 28-112)	UDV305 UDV737	NA	Pa	(Di Gaspero *et al.*, 2012)
	Rpv4	4	unknown	Regent	VMC7h3, VMCNg2e1	NA	Mn	(Welter *et al.*, 2007)
	Rpv5	9	*V. riparia*	G. de Montpellier	VVIo52b	NA	Mn	(Marguerit *et al.*, 2009)
	Rpv6	12	*V. riparia*	G. de Montpellier	VMC8g9	NA	Mn	(Marguerit *et al.*, 2009)
	Rpv7	7	unknown	Bianca	UDV097	NA	Mn	(Bellin *et al.*, 2009)
	Rpv8	14	*V. amurensis*	Ruprecht	Chr14V015	CH	Mj	(Blasi *et al.*, 2011)
	Rpv9	7	*V. riparia*	W63	CCoAOMT	NA	Mn	(Moreira *et al.*, 2011)
	Rpv10	9	*V. amurensis*	Solaris	GF09-46	CH	Pa	(Schwander *et al.*, 2012)

Rpv11	5	unknown	Regent Chardonnay Solaris	VVMD27 CS1E104J11F VCHR05C	NA	Mn	(Fischer *et al.*, 2004) (Bellin *et al.*, 2009) (Schwander *et al.*, 2012)
Rpv12	14	*V. amurensis*	99-1-48 20/3	UDV014 UDV304, rgvvin180, UDV370	CH	Mj	(Venuti *et al.*, 2013)
Rpv13	12	*V. riparia*	W63	VMC1g3.2	NA	Mn	(Moreira *et al.*, 2011)
Rpv14	5	*V. cinerea*	Börner	GF05-13	NA	Mn	(Ochssner *et al.*, 2016)
Rpv15	18	*V. piasezkii*	DVIT2027		CH	Mj	-
Rpv16	9	*V. piasezkii*	DVIT2027		CH	Mn	-
Rpv17	8	unknown	Horizon		NA	Mn	(Divilov *et al.*, 2018)
Rpv18	11	unknown	Horizon		NA	Mn	(Divilov *et al.*, 2018)
Rpv19	14	*V. rupestris*	B38		NA	Mn	(Divilov *et al.*, 2018)
Rpv20	6	unknown	Horizon		NA	Mn	(Divilov *et al.*, 2018)
Rpv21	7	unknown	Horizon		NA	Mn	(Divilov *et al.*, 2018)
Rpv22	15	*V. amurensis*	Shuanghong		CH	Pa	(Fu *et al.*, 2020)
Rpv23	2	*V. amurensis*	Shuanghong		CH	Mn	(Fu *et al.*, 2020)
Rpv24	18	*V. amurensis*	Shuanghong		CH	Mn	(Fu *et al.*, 2020)
Rpv25	15	*V. amurensis*	Shuangyou	Marker561375, Marker549779	CH	Pa	(Lin *et al.*, 2019)
Rpv26	15	*V. amurensis*	Shuangyou	Marker525926, Marker526446	CH	Pa	(Lin *et al.*, 2019)
Rpv27	18	*V. aestivalis*	Norton	VVCS1H077H16R1-1 UDV737	NA	Pa	(Sapkota *et al.*, 2019)
Rpv28	10	*V. rupestris*	B38	VVIH01, UDV-073	NA	Mj	(Bhattarai *et al.*, 2021)
Rpv29	14	*V. vinifera*	Mgaloblishvili	chr14_21613512_C_T	CA	Mj	(Sargolzaei *et al.*, 2020)
Rpv30	3	*V. vinifera*	Mgaloblishvili	cn_C_T_chr3_16229046	CA	Mn	(Sargolzaei *et al.*, 2020)
Rpv31	16	*V. vinifera*	Mgaloblishvili	li_T_C_chr16_21398409	CA	Mn	(Sargolzaei *et al.*, 2020)

Table 2. List of the 30 most cultivated wine grape resistant varieties throughout the World. Classification and distribution data were retrieved from Anderson and Nelgen (2020). Origin, parentage, and *R* loci data were retrieved from Topfer and Hausmann (2021) and Maul (2022).

Rank among hybrids	Rank among varieties	Name	Country of origin	Parent 1	Parent 2	*R* loci	Cultivated area (ha)	World share (%)
1	64	Concord	USA	Catawba	*V. labrusca*		10544	0.24
2	69	Bianca	Hungary	Villard Blanc	Bouvier	*Rpv3, Rpv7*	9766	0.22
3	151	Seyval Blanc	France	Seibel 5656	Rayon D'Or	*Ren3, Ren9, Ren10, Rpv3.2, Rpv3.3*	2699	0.06
4	173	Regent	Germany (PIWI)	Diana	Chambourcin	*Ren3, Ren9, Rpv3.1, Rpv4, Rpv11*	1974	0.04
5	254	Chambourcin	France (PIWI)	Joanne Seyve 11369	Plantet	*Rpv3.1, Rpv3.2*	968	0.02
6	282	Villard Noir	France	Chancellor	Subereux		777	0.02
7	288	Baco Noir	France	Folle Blanche	Riparia Grand Glabre + V. riparia		735	0.02
8	316	Catawba	USA	*V. labrusca*	Semillon		626	0.01
9	358	Couderc 13	France	*V. aestivalis*	Couderc 162- 5		474	0.01
10	401	Norton	USA	*V. aestivalis* var. Lincecumii Buckley	*V. vinifera* subsp. vinifera	*Rpv27*	328	0.01
11	441	Aurore	France	Seibel 788	Seibel 29		255	0.01
12	449	Vignoles	France				241	0.01
13	450	Traminette	USA	Joannes Seyve 23416	Gewürztraminer		239	0.01
14	456	Maréchal Foch	France	Millardet et Grasset 101 O.P.	Goldriesling		229	0.01
15	465	Cayuga White	USA	Seyval blanc	Schuyler		217	<0.01
16	469	Frontenac	USA	Riparia 89	Landot Noir		212	<0.01
17	505	Marquette	USA	Minnesota 1094	Ravat Noir		166	<0.01
18	554	Solaris	Germany (PIWI)	Merzling	Geisenheim 6493	*Ren3, Ren9, Rpv3.3, Rpv10, Rpv11*	118	<0.01
19	564	Johanniter	Germany (PIWI)	Riesling Weiss	Freiburg 589-54	*Ren3, Ren9, Rpv3.1*	111	<0.01
20	576	De Chaunac	France	Seibel 5163	Seibel 793		102	<0.01
21	583	La Crescent	USA	St. Pepin	Elmer Swenson 6-8-25		94	<0.01
22	594	Chardonel	USA	Seyval blanc	Chardonnay		90	<0.01
23	604	Léon Millot	France	Millardet et Grasset 101 O.P.	Goldriesling		85	<0.01
24	687	Rondo	Germany (PIWI)	Zarya Severa	Saint Laurent	*Rpv10*	51	<0.01
25	689	Golden Muscat	USA (PIWI)	Muscat Hamburg	Diamond		50	<0.01
26	706	Phoenix	Germany	Bacchus Weiss	Villard blanc	*Ren3, Ren9, Rpv3.1*	46	<0.01
27	730	Cabernet Cortis	Germany (PIWI)	Cabernet Sauvignon	Solaris	*Ren3, Ren9, Rpv3.3, Rpv10*	38	<0.01
28	734	Chancellor	France	Seibel 5163	Seibel 880		38	<0.01
29	770	Rathay	Austria (PIWI)	Blauburger	Klosterneuburg 1189-9-77		32	<0.01
30	809	La Crosse	USA	Minnesota 78 x Seibel 1000	Seyval		26	<0.01

Table 3. List of the most representative examples of *Vitis* genetic transformation to improve resistance against *E. necator*, *P. viticola* and *B. cinerea*.

Target	Gene Source	*V. vinifera* host	Protein (Gene)	Promoter	Selection and reporter genes	Transformed tissue	Ref
E. necator	*V. vinifera*	Thompson Seedless	Thaumatin-like protein 1 (*Vvtl-1*)	CaMV 35S	nptII (GFP)	Leaf	(Dhekney *et al.*, 2011)
	V. vinifera	Chardonnay	Non-expressor of pathogenesis related 1 (*VvNPR1.1*)	CaMV 35S	nptII (GFP)	Anther	(Le Henanff *et al.*, 2011)
	V. pseudoreticulata	Red Globe	Pathogenesis-related protein 4-1 (*VpPR4-1*)	CaMV 35S	nptII	Anther	(Dai *et al.*, 2016)
	V. pseudoreticulata	Red Globe	F-box/Kelch-repeat protein (*VpEIFP1*)	CaMV 35S	nptII	Anther	(Wang *et al.*, 2017)
	V. pseudoreticulata	Chardonnay	Stilbene synthase (*VpSTSgDNA2*)	CaMV 35S	hptII	Anther, ovary, flower	(Dai *et al.*, 2015)
	V. quinquangularis	Thompson Seedless	Stilbene synthase 6 (*VqSTS6*)	CaMV 35S	nptII	Anther	(Cheng *et al.*, 2016)
	O. sativa	Pusa seedless	Rice chitinase (*Chi11*)	maize ubiquitin promoter	hptII, bar	Leaf	(Nirala *et al.*, 2010)
	O. sativa	Neo Muscat	Rice chitinase (*RCC2*)	CaMV 35S	nptII	Ovule	(Yamamoto *et al.*, 2000)
	A. thaliana	Thompson Seedless	Resistance to powdery mildew 8 locus (*RPW8.2*)	CaMV 35, native promoter	nptII	Anther	(Hu *et al.*, 2018)
	Trichoderma spp.	Thompson Seedless	Endochitinases (*ech42*) + N-acetyl-β-d-hexosaminidase (*nag70*)	CaMV 35S	nptII	Bud	(Rubio *et al.*, 2015)
P. viticola	*V. pseudoreticulata*	Thompson Seedless	Pathogenesis-related protein 10.1 (*VpPR10.1*)	CaMV 35S	nptII (GFP)	Anther	(Su *et al.*, 2018)
	V. amurensis	Thompson Seedless	Thaumatin-like protein (*VaTLP*)	CaMV 35S	nptII	Anther	(He *et al.*, 2017)
	F. graminearium	Crimson Seedless	Chitinase and β-1,3-glucanase	maize ubiquitin	hptII, bar	Leaf	(Nookaraju and Agrawal, 2012)
B. cinerea	*V. vinifera*	Thompson Seedless	Thaumatin-like protein (*Vvtl-1*)	CaMV 35S	nptII (GFP)	Leaf	(Dhekney *et al.*, 2011)
	V. vinifera	Sugraone	Stilbene synthase 1 (*Vst1*)	CaMV 35S	nptII (GFP)	n.s.	(Dabauza *et al.*, 2015)
	V. vinifera	41B (Chasselas × Berlandieri)	Stilbene synthase 1 (*Vst1*)	alfalfa *PR10.1* promoter	nptII	Immature embryo	(Coutos-Thévenot *et al.*, 2001)

Table 4. List of candidate and validated (in bold) target genes for RNS interference conferring resistance towards *E. necator* (PM), *P. viticola* (DM) and *B. cinerea* (GM). Together with pathogen target genes we also reported several grapevine susceptibility target genes.

Gene Name	Gene ID	Organism	Disease	Function	Ref
VviMLO6 [A]	Vitvi10g00509*	*V. vinifera*	PM	Trans-membrane protein, Negative regulator of cell-wall appositions.	(Pessina *et al.*, 2016)
VviMLO7 [A]	Vitvi13g00578*				
VviMLO11 [A]	Vitvi08g01055*				
VviCSN5	Vitvi03g00286*			COP9 signalosome complex subunit 5 (CSN5), Involved in protein degradation via the ubiquitin-proteasome pathway.	(Cui *et al.*, 2021)
VviDMR6-1	Vitvi16g01336*		DM	SA 5-hydroxylase, Involved in SA catabolism.	(Pirrello *et al.*, 2022)
VviDMR6-2	Vitvi13g01119*				(Giacomelli *et al.*, 2022)
VvWRKY40	Vitvi13g00189*			TF, stabilized by *PvRXLR111* effector.	(Ma *et al.*, 2021)
VviLBDIf7 [B]	Vitvi13g00549*			LBD protein, Repressor of JA-mediated defense.	(Marcianò *et al.*, 2021)
BcCYP51 [B]	BcDW1_10539**	*B. cinerea*	GM	Lanosterol 14α-demethylase, Primary target of sterol demethylation inhibitors (DMI) fungicides.	(Nerva *et al.*, 2020)
Bcchs1 [B]	BcDW1_7822**			Chitin synthase, Contributes to cell wall composition.	
BcEF2 [B]	BcDW1_2938**			β 1-tubulin promoter.	
BcDCL1 [B]	BcDW1_481**			Dicer-like (DCL) gene, Effector.	(Wang *et al.*)
BcDCL2 [B]	BcDW1_2215**				
BcVPS51 [B]	BcDW1_6998**			Vesicle-trafficking pathway genes.	(Qiao *et al.*, 2021)
BcDCTN1 [B]	-				
BcSAC1 [B]	-				
PvDCL1 [B]	PVITv1_T038441***	*P. viticola*	DM	Dicer-like (DCL) gene, Effector.	(Figueiredo *et al.*, 2021)
PvDCL2 [B]	PVITv1_T003331***				
PvRXLR159	-			Cell death inhibitor, Effector.	(Lei *et al.*, 2019)
PvRXLR28	-			Cell death inhibitor, Effector.	(Xiang *et al.*, 2016)
PvRXLR131	-			Cell death inhibitor through interaction with *BRI1 kinase inhibitor 1* (*VviBKI1*), Effector.	(Lan *et al.*, 2019)
PvRXLR53	-			INF1-triggered programmed cell death and defense-related genes expression inhibitor, Effector.	(Liu *et al.*, 2021)
EnCSEP56	-	*E. necator*	PM	Candidates for Secreted Effector Proteins (CSEPs) RALPH effector, plant immunity inhibitors.	(Zhang, 2021)
EnCSEP65	-				
EnCSEP115	-				
EnCSEP087	-			Candidates for Secreted Effector Proteins (CSEPs) RALPH effector, Targets *VviADC* suppressing ROS production.	(Mu *et al.*, 2022)

[A] Experiment conducted by means of host-induced gene silencing (HIGS)
[B] Experiment conducted by means of spray-induced gene silencing (SIGS)

* Available at JBrowse https://urgi.versailles.inra.fr/jbrowse/gmod_jbrowse/ (Buels *et al.*, 2016)
** Available at Grapegenomics.com http://www.grapegenomics.com/pages/P_Bcinerea/ (Blanco-Ulate *et al.*, 2013);
*** Available at NCBI https://www.ncbi.nlm.nih.gov/assembly/GCA_003123765.1 (Brilli *et al.*).

6.9 Boxes

Box 1 The genetic diversity of the *Vitis* genus

The members of the *Vitaceae* family are classified into 5 tribes (Wen *et al.*, 2018), consisting of 16 genera and 950 species (Lu *et al.*, 2018). The grape genus *Vitis* L., with 75 species, belongs to the Viteae tribe. The *Vitis* genus includes the *Muscadinia* [(Planch.) Rehder] (2n=40 chromosomes) and *Vitis* (2n=38) subgenera, to which the *V. vinifera* wild subspp. *sylvestris* and cultivated subspp. *Sativa* belong. Species belonging to the *Vitis* subgenus are interfertile. The genetic proximity between *sylvestris* and *sativa* and the occurrence of intermediate genotypes corroborates the widely accepted theory that cultivated subspecies originated from the domestication of wild individuals (Grassi and Arroyo-Garcia, 2020). The autogamous nature of *V. vinifera* determines low levels of genetic polymorphism compared to hybrids and wild species, which, however, results in interspecific and even intervarietal genomic variations (Cardone *et al.*, 2016). This high level of genetic plasticity is a crucial resource for the exploitation of resistance traits to the main fungal pathogens (Vezzulli *et al.*, 2022). Interestingly, the recent awareness that spontaneous gene introgression in wild species initially contributed to the establishment of resistance traits to biotic stresses (Morales-Cruz *et al.*, 2021) led to new studies on the ancestry of *Vitis* members (Foria *et al.*, 2022). Moreover, self-fertilization represents a primal selection tool that, on the one hand, favors the establishment of lethal recessive alleles but, on the other hand, naturally selects surviving homozygous individuals as “good performers”. For these reasons, several *vinifera* and non-*vinifera* core collections have been created in the last few years (Cipriani *et al.*, 2008; Le Cunff *et al.*, 2008; Emanuelli *et al.*, 2013; Migliaro *et al.*, 2019).

Box 2 Transgenesis, intragenesis and cisgenesis in plants

By definition, transgenesis is the genetic alteration of a plant with one (or more) gene(s) derived from any unrelated and sexually incompatible plant taxa or even from non-plant organisms (Schouten *et al.*, 2006). In contrast, cisgenesis and intragenesis represent two possible alternatives that rely on the use of genetic material from the same species or from closely related species capable of sexual hybridization being transformed (Holme *et al.*, 2013).

While cisgenesis involves genetic modification using a complete gene unit, including its regulatory elements (i.e., the entire original transcriptional unit), intragenesis refers to the *in vitro* recombination and transfer of elements (i.e., promotors, introns, exons, and terminators) isolated from different genes and assembled into a chimeric construct. In both cases, the result is a product

free of non-plant DNA sequences, and consequently, any selection markers or residual vector sequences at the level of transformed plants are not expected (Rommens *et al.*, 2007; Holme *et al.*, 2013). In contrast to intragenesis, cisgenesis events also occur spontaneously in nature or by means of conventional breeding. However, cisgenesis has great potential to overcome the two major bottlenecks typical of traditional breeding: much longer times and the risk of linkage drag, i.e., the possibility to transfer genes with deleterious effects that are associated with the gene of interest.

Box 3 Genome editing and CRISPR/Cas9 in plants

The advent of the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) system has been revolutionizing genome editing approaches, allowing the modification of cellular DNA with a much higher level of versatility and precision and with a wide extension of applications with respect to previous technologies (Ren *et al.*, 2021). The system is based on the recognition of a predetermined DNA target site through the action of a complementary RNA sequence, namely, the guide RNA (gRNA), and the Cas endonuclease, followed by the DNA double-strand break at the target site, which determines the onset of insertions, deletions, or nucleotide substitutions (Scintilla *et al.*, 2021). This technique is less expensive, simpler, and faster than other enzyme-based techniques, such as zinc finger nuclease (ZFN) or transcription activator-like effector nuclease (TALEN; Ghogare *et al.*, 2021). Therefore, especially in the case of the genetic improvement of elite varieties of woody plants, such as grapevine, CRISPR–Cas9 technology ensures genotype preservation while resulting in targeted and specific genetic modifications. Although there is enormous scientific interest in applying this technology to grape varieties, there are few examples of successes. On the one hand, the decrease in the application of this technique could be due to the disillusionment caused by the short-sightedness of national and international organizations that still associate genome editing with transgenesis and with the generation of GMOs. On the other hand, especially in grapevines, there are technical impediments to using these processes, including the difficulty and timing of regeneration to create stable edited lines and the difficulty of creating plants free of exogenous DNA through segregation of the induced mutation by the CRISPR/Cas transgenic construct.

Box 4 RNA interference mechanisms in plants

RNA interference (RNAi) is one of the most important natural regulatory and defense mechanisms that is implicated in the control of plant growth, development, and response to biotic or abiotic

stresses. RNAi controls the expression of endogenous protein-coding genes in a sequence-specific manner and regulates the plant response to undesired nucleic acids, transposons, and transgene activity (Dubrovina *et al.*, 2016). The suggested RNA silencing mechanism coordinately involves key components, such as Dicer-Like (DCL) proteins, Argonaute (AGO) proteins, and RNA-dependent RNA Polymerase (RDR) proteins, which are responsible for the genesis of 21-24 nucleotide (nt) small RNAs (sRNAs) and their biological function (Vaucheret, 2006). Although a double-stranded RNA (dsRNA) molecule serves as a precursor for subsequent sRNA production, sRNAs can be classified as small interfering RNAs (siRNAs), microRNAs (miRNAs), or trans-acting small interfering RNAs (ta-siRNAs) based on their origin. A number of DCL, AGO, and RDR genes have been discovered in various plant species, detailing their roles in plant defense systems (Vaucheret, 2008; Nakasugi *et al.*, 2013; Liu *et al.*, 2014). Plant AGO proteins can be divided into two main classes depending on the RNAi pathway in which they are involved: AGO4, 6, and 9 mediate so-called transcriptional gene silencing (TGS), and AGO1, 2, 3, 5, 7 and 10 mediate posttranscriptional gene silencing (PTGS; Mallory and Vaucheret, 2012). In TGS, sRNA loaded on AGO proteins recruits DNA methyltransferases to methylate the cytosine residues of the corresponding target gene in the genome, whereas in PTGS, the sRNA-AGO complex scans the cytoplasm for complementary transcripts for cleavage and degradation. Since its discovery over 20 years ago, RNA interference (RNAi) has been widely used in reverse genetics and functional genomics to downregulate the expression of genes responsible for the control of abiotic stress tolerance, developmental processes, and other plant responses (Niu *et al.*, 2008; Biswal *et al.*, 2015), as well as in crop protection activities for the control of pest and pathogen resistance (as reviewed in Kaur *et al.*, 2021).

Box 5 Technical, biological, and social issues associated with conventional and biotechnological breeding techniques

Conventional breeding. This approach is based on several backcrossing generations, introgressions, induced mutagenesis and somatic hybridization. Genes of interest are found in crossable, sexually compatible organisms. Although this technology has a high consumer acceptance as a non-GMO approach and has no biosafety concerns, it presents several side technical issues. First, the timing required for breeding programs is quite long and requires deep knowledge and the availability of genetic resources. Newly developed genotypes are similar to the original clones but are not exactly the same, a less appreciated factor in a highly conservative discipline such as viticulture. Moreover,

if resistance to diseases is provided by the insertion of dominant R genes, it will be easily overcome by the onset of new pathogen strains (Myles, 2013).

Transgenesis/cisgenesis. Both approaches are based on genetic transformation**.** The main difference between them is the origin of the target gene: transgenesis involves the overexpression of genes from nonsexually compatible organisms and the presence of sequences from noncompatible organisms (promoters, terminators, selective markers), while cisgenesis only considers genes (including introns) and their regulatory regions (promoters, terminator) from sexually compatible plants. These technologies have the advantage of bypassing linkage-drag phenomena and take less time to produce (1-2 years). Nevertheless, they are considered GMOs, which presents several side effects or technical issues, such as the release of genes with different origins into the environment, the expression of new protein products with possible allergen/toxic effects (transgenesis), and the scarce availability of efficient cisgenic selectable marker genes. Furthermore, the expression of unknown proteins and the use of antibiotic- or herbicide-resistant markers make transgenesis unpopular with consumers. In contrast, cisgenic plants may partially solve the current biosafety problems regarding the presence of foreign DNA in the host genome, increasing consumer acceptance.

Genome editing is based on both genetic transformation and plasmid-free protoplast transformation. The advantage of this technique lies in the fact that it is extremely precise and can introduce mutations at the single nucleotide level without changing the genomic background and thus preserving clonal identity. Technical issues are related to the fact that when applying plasmid-based genetic transformation, segregation is required from T0 plants, an extremely disadvantageous procedure in a clonally propagated plant with a highly heterozygous genetic background. In contrast, plasmid-free protoplast transformation bypasses the aforementioned problem but requires *in vitro* regeneration from protoplasts, which can be tricky since many grapevine varieties are recalcitrant to regeneration.

dsRNA-based SIGS has the advantage of not being considered a transgenic approach since it is not based on recombinant DNA technology and can be implemented in a few months. Compared to many of the techniques discussed, the exogenous application of small RNA molecules is characterized by a much higher targeted precision during intervention and even reduced pesticide use following intervention, which fits in an integrated crop protection strategy by being able to downregulate exogenous or endogenous gene expression. On the other hand, it is necessary to evaluate how the effectiveness of gene silencing depends on the efficiency and specificity of the

RNA molecule sequence and the degree of absorption feasible by the plant host and pathogen cells (Gebremichael *et al.*, 2021). Likewise, the RNAi sequence must be carefully selected to avoid any possible off-target and pleiotropic effects. The onset of potential negative effects on the ecosystem and human health from the use of this technique must be investigated in more detail. A multilevel study that includes every single layer and all reciprocal interactions would feedback on the wide complete integrated model to predict the effect of a treatment on every single stakeholder and on the entire system.

Box 6 Rules and disputes relating to the use of transgenic, cisgenic, intragenic and edited plants in agriculture and food in the European framework

Despite the potential ability of the application of genetic transformation techniques to control the main phytopathogens, the use of transgenic varieties in agriculture is at present severely limited in the EU. The community regulatory context, based on the precautionary principle, was first defined by Directive 2001/18/EC of the European parliament (EU Parliament & Council, 2001). This law replaces Council directive 90/220/EEC and stipulates the basic rules for the approval of a new GMO. Second, there are two regulations (EU Parliament & Council, 2003*a*,*b*) that control the authorization and labeling/traceability of food and feed consisting of or derived from GMOs. Third, the Commission Recommendation of 23 July (EU Commission, 2003) indicates the guidelines for the coexistence of GM and conventional crops, to which national and regional regulations must refer. Finally, the EC Directive 2015/412 (EU Parliament & Council, 2015) integrates the previous EC Directive 2001/18 by delegating ownership to national states to allow, limit or prohibit the commercial cultivation of GMO plants. The action taken by EU countries must comply with EU legislation and be reasonable, proportional, and nondiscriminatory. EU countries where GMOs are grown must take measures in border areas to avoid possible contamination of neighboring EU countries where the cultivation of those GMOs is prohibited. Basically, the cornerstone of European standards is the great attention given to the assessment of all potential risks, based on the precautionary principle, and the fact that all authorizations are granted for a limited period of time, during which environmental and health effects must be carefully monitored. The utmost precaution toward the authorization and cultivation of transgenic plants by the EU derives from a series of concerns raised by public opinion related to the transfer of selection markers expressed by transgenic plants to pathogenic microorganisms in the gastrointestinal tract or soil, bringing resistance to herbicides or antibiotics. This problem could be overcome by the cisgenic approach,

which, however, carries with it a series of critical issues: i) insertion of the donor sequence in a random position, which can potentially influence DNA methylation and regulation; ii) potential occurrence of mutations at the insertion site, with the possibility of rearrangements and translocations in the flanking regions; iii) small portions of T-DNA remaining as "junk DNA" and not being able to determine any type of phenotypic effect; and iv) donor sequences not being replaced with an allelic sequence but instead added to the recipient genome (Schouten *et al.*, 2006; Rommens *et al.*, 2007; Jacobsen and Schouten, 2008; Telem *et al.*, 2013). Many of the questions raised by public opinion on cisgenics could also be raised about plants obtained from traditional breeding. Indeed, according to EFSA, the product obtained by means of a cisgenic approach is not dissimilar to that obtained by traditional breeding (EFSA Panel on Genetically Modified Organisms (GMO), 2012). Nonetheless, at present, according to the community legislation in place, cultivars obtained by cisgenesis and intragenesis are considered GMOs. Breeding innovations consist of known, controllable and predictable procedures. With the development of new molecular techniques for crop improvement, new legislative issues need to be resolved. As testified to by the sentence released by the European Court of Justice on 25 July 2018 (Vives-Vallés and Collonnier, 2020), organisms obtained through techniques or methods of mutagenesis must be considered GMOs. The sentence includes in the scope of application of the EC Directive 2001/18 organisms obtained by means of new techniques or new methods that have emerged after the adoption of the Directive in question (i.e., CRISPR/Cas9-driven genome editing) and, consequently, excluding from its scope only organisms obtained through techniques or methods conventionally used in various applications with a long tradition of safety (such as organisms obtained with traditional methods of random mutagenesis by ionizing radiation or exposure to chemical mutagenic agents). On the one hand, this ruling condemns genome editing techniques to confinement in the legislation on GMOs but, on the other hand, it generates paradoxes that require the European legislator to review the entire discipline. Indeed, the current legislation is based on the regulation of organisms mainly on the basis of their production process (e.g., chemical/physical mutagenesis, conventional genetic transformation, genomic editing); however, it is essential to start a discussion that also takes into consideration the nature of the final product, since some induced genetic modifications can be considered equivalent to those obtainable with natural selection mechanisms. For example, the same plant genotype could be derived from a spontaneous random event, from a mutagenesis experiment with ionizing radiation or from site-specific mutagenesis implemented by CRISPR, and by sequencing the genome of the organism, there would be no possibility of determining the source

of the genetic variation, thus making any type of regulatory framework impossible. Releasing the organisms obtained via new technologies from natural/unnatural dualism of this regulation would allow for the simplification and real effectiveness of the procedures for risk assessment, which, at that point, would be applied to individual cases with the intervention of adequate legislation and regulatory bodies (ISAAA, 2021). As a signal that the European legislative bodies and public opinion are moving in this direction, the EC has recently started a public consultation to prepare a policy initiative on plant products aimed to provide an adequate regulation that guarantees the protection of the environment and human health with the contribution to innovation given by the plants obtained with new genomics techniques (EU Commission, 2022).

BIBLIOGRAPHIC REFERENCES

Adam-Blondon A. F., Roux C., Claux D., Butterlin G., Merdinoglu D., This P. 2004. Mapping 245 SSR markers on the Vitis vinifera genome: A tool for grape genetics. Theoretical and Applied Genetics 109, 1017–1027.

Agüero C. B., Uratsu S.L., Greve C., Powell A. L. T., Labavitch J. M., Meredith C. P., Dandekar A. M. 2005. Evaluation of tolerance to Pierce's disease and Botrytis in transgenic plants of *Vitis vinifera* L. expressing the pear PGIP gene. Molecular Plant Pathology 6, 43–51.

Agurto M., Schlechter R. O., Armijo G., Solano E., Serrano C., Contreras R. A., Zúñiga G. E., Arce-Johnson P. 2017. RUN1 and REN1 pyramiding in grapevine (Vitis vinifera cv. crimson seedless) displays an improved defense response leading to enhanced resistance to powdery mildew (erysiphe necator). Frontiers in Plant Science 8, 758.

Akkurt M., Welter L., Maul E., Töpfer R., Zyprian E. 2007. Development of SCAR markers linked to powdery mildew (Uncinula necator) resistance in grapevine (Vitis vinifera L. and Vitis sp.). Molecular Breeding 19, 103–111.

Ali S., Ganai B. A., Kamili A. N., *et al.* 2018. Pathogenesis-related proteins and peptides as promising tools for engineering plants with multiple stress tolerance. Microbiological Research 212–213, 29–37.

Anderson K., Nelgen S. 2020. *Which winegrape varieties are grown where? A global empirical picture*. Adelaide, Australia: University of Adelaide Press.

Barba P., Cadle-Davidson L., Harriman J., Glaubitz J., Brooks S., Hyma K. E., Reisch B. I. 2014. Grapevine powdery mildew resistance and susceptibility loci identified on a high-resolution SNP map. Theoretical and Applied Genetics 127, 73–84.

Barker C. L., Donald T., Pauquet J., Ratnaparkhe M. B., Bouquet A., Adam-Blondon A. F., Thomas M. R., Dry I. 2005. Genetic and physical mapping of the grapevine powdery mildew resistance gene, Run1, using a bacterial artificial chromosome library. Theoretical and Applied Genetics 111, 370–377.

Bellin D., Peressotti E., Merdinoglu D., Wiedemann-Merdinoglu S., Adam-Blondon A. F., Cipriani G., Morgante M., Testolin R., Di Gaspero G. 2009. Resistance to *Plasmopara viticola* in grapevine 'Bianca' is controlled by a major dominant gene causing localised necrosis at the infection site. Theoretical and Applied Genetics 120, 163–176.

Bhattarai G., Fennell A., Londo J. P., Coleman C., Kovacs L. G. 2021. A Novel Grape Downy Mildew Resistance Locus from Vitis rupestris. American Journal of Enology and Viticulture 72, 12–20.

Biswal A. K., Hao Z., Pattathil S., *et al.* 2015. Downregulation of GAUT12 in Populus deltoides by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. Biotechnology for Biofuels 8, 41.

Blanc S., Wiedemann-Merdinoglu S., Dumas V., Mestre P., Merdinoglu D. 2012. A reference genetic map of Muscadinia rotundifolia and identification of Ren5, a new major locus for resistance to grapevine powdery mildew. Theoretical and Applied Genetics 125, 1663–1675.

Blanco-Ulate B., Allen G., Powell A. L. T., Cantu D. 2013. Draft genome sequence of *Botrytis cinerea* BcDW1, inoculum for noble rot of grape berries. Genome Announcements 1, e00252-13.

Blasi P., Blanc S., Wiedemann-Merdinoglu S., Prado E., Rühl E. H., Mestre P., Merdinoglu D. 2011. Construction of a reference linkage map of Vitis amurensis and genetic mapping of Rpv8, a locus conferring resistance to grapevine downy mildew. Theoretical and Applied Genetics 123, 43–53.

Brilli M., Asquini E., Moser M., Bianchedi P. L., Perazzolli M., Si-Ammour A. A multi-omics study of the grapevine-downy mildew (Plasmopara viticola) pathosystem unveils a complex protein coding-and noncoding-based arms race during infection.

Brilli M., Asquini E., Moser M., Bianchedi P. L., Perazzolli M., Si-Ammour A. 2018. A multi-omics study of the grapevine-downy mildew (Plasmopara viticola) pathosystem unveils a complex protein coding-A nd noncoding-based arms race during infection. Scientific Reports 8, 757.

Buels R., Yao E., Diesh C. M., *et al.* 2016. JBrowse: A dynamic web platform for genome visualization and analysis. Genome Biology 17, 66.
Buonassisi D, Colombo M, Migliaro D, Dolzani C, Peressotti E, Mizzotti C, Velasco R, Masiero S, Perazzolli M, Vezzulli S. 2017. Breeding for grapevine downy mildew resistance: a review of "omics" approaches. Euphytica 213, 103.

Campbell N. R., Harmon S. A., Narum S. R. 2015. Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. Molecular Ecology Resources 15, 855–867.

Canaguier A., Grimplet J., Di Gaspero G., *et al.* 2017. A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). Genomics Data 14, 56–62.

Capriotti L., Baraldi E., Mezzetti B., Limera C., Sabbadini S. 2020. Biotechnological approaches: Gene overexpression, gene silencing, and genome editing to control fungal and oomycete diseases in grapevine. International Journal of Molecular Sciences 21, 5701.

Cardone M. F., D'Addabbo P., Alkan C., *et al.* 2016. Inter-varietal structural variation in grapevine genomes. The Plant Journal 88, 648–661.

Carrier G., Le Cunff L., Dereeper A., Legrand D., Sabot F., Bouchez O., Audeguin L., Boursiquot J. M., This P. 2012. Transposable elements are a major cause of somatic polymorphism in Vitis vinifera L. PLOS ONE 7, e32973.

Cheng S., Xie X., Xu Y., Zhang C., Wang X., Zhang J., Wang Y. 2016. Genetic transformation of a fruit-specific, highly expressed stilbene synthase gene from Chinese wild Vitis quinquangularis. Planta 243, 1041–1053.

Cipriani G., Marrazzo M. T., Di Gaspero G., Pfeiffer A., Morgante M., Testolin R. 2008. A set of microsatellite markers with long core repeat optimized for grape (Vitis spp.) genotyping. BMC Plant Biology 8, 127.

Cochetel N., Minio A., Massonnet M. L., Vondras A. M., Figueroa-Balderas R., Cantu D. 2021. Diploid chromosome-scale assembly of the Muscadinia rotundifolia genome supports

chromosome fusion and disease resistance gene expansion during Vitis and Muscadinia divergence. G3 Genes|Genomes|Genetics 11, jkab033.

Corso M., Bonghi C. 2014. Grapevine rootstock effects on abiotic stress tolerance. Plant Science Today 1, 108–113.

Coutos-Thévenot P., Poinssot B., Bonomelli A., Yean H., Breda C., Buffard D., Esnault R., Hain R., Boulay M. 2001. In vitro tolerance to *Botrytis cinerea* of grapevine 41B rootstock in transgenic plants expressing the stilbene synthase Vst1 gene under the control of a pathogen-inducible PR 10 promoter. Journal of experimental botany 52, 901–910.

Cui K. C., Liu M., Ke G. H., Zhang X. Y., Mu B., Zhou M., Hu Y., Wen Y. Q. 2021. Transient silencing of VvCSN5 enhances powdery mildew resistance in grapevine (Vitis vinifera). Plant Cell, Tissue and Organ Culture 146, 621–633.

Le Cunff L., Fournier-Level A., Laucou V., Vezzulli S., Lacombe T., Adam-Blondon A.-F., Boursiquot J-M, This P. 2008. Construction of nested genetic core collections to optimize the exploitation of natural diversity in Vitis vinifera L. subsp. sativa. BMC Plant Biology 8, 31.

Dabauza M., Velasco L., Pazos-Navarro M., Pérez-Benito E., Hellín P., Flores P., Gómez-Garay A., Martínez M. C., Lacasa A. 2015. Enhanced resistance to Botrytis cinerea in genetically-modified Vitis vinifera L. plants over-expressing the grapevine stilbene synthase gene. Plant Cell, Tissue and Organ Culture (PCTOC) 120, 229–238.

Dai L., Wang D., Xie X., Zhang C., Wang X., Xu Y., Wang Y., Zhang J. 2016. The novel gene VpPR4-1 from vitis pseudoreticulata increases powdery mildew resistance in transgenic vitis vinifera L. Frontiers in Plant Science 7, 695.

Dai L., Zhou Q., Li R., Du Y., He J., Wang D., Cheng S., Zhang J., Wang Y. 2015. Establishment of a picloram-induced somatic embryogenesis system in Vitis vinifera cv. chardonnay and genetic transformation of a stilbene synthase gene from wild-growing Vitis species. Plant Cell, Tissue and Organ Culture (PCTOC) 121, 397–412.

Dalbó M. A., Ye G. N., Weeden N. F., Wilcox W. F., Reisch BI. 2001. Marker-assisted selection for powdery mildew resistance in grapes. Journal of the American Society for Horticultural Science 126, 83–89.

Dalla Costa L., Malnoy M., Gribaudo I. 2017. Breeding next generation tree fruits: Technical and legal challenges. Horticulture Research 4, 17067.

Dhekney S. A., Li Z. T., Gray D. J. 2011. Grapevines engineered to express cisgenic Vitis vinifera thaumatin-like protein exhibit fungal disease resistance. In Vitro Cellular and Developmental Biology - Plant 47, 458–466.

Divilov K., Barba P., Cadle-Davidson L., Reisch B. I. 2018. Single and multiple phenotype QTL analyses of downy mildew resistance in interspecific grapevines. Theoretical and Applied Genetics 131, 1133–1143.

Dry I., Riaz S., Fuchs M., Sosnowski M., Thomas M. 2019. Scion breeding for resistance to biotic stresses. In: Cantu D,, In: Walker M, eds. The grape genome. Cham: Springer, 319–347.

Dubrovina A. S., Aleynova O. A., Kiselev K. V. 2016. Influence of overexpression of the true and false alternative transcripts of calcium-dependent protein kinase CPK9 and CPK3a genes on the growth, stress tolerance, and resveratrol content in Vitis amurensis cell cultures. Acta Physiologiae Plantarum 38, 78.

Dubrovina A. S., Kiselev K. V. 2019. Exogenous RNAs for Gene Regulation and Plant Resistance. International Journal of Molecular Sciences 20, 2282.

EFSA Panel on Genetically Modified Organisms (GMO). 2012. Scientific opinion addressing the safety assessment of plants developed through cisgenesis and intragenesis. EFSA Journal 10, 2561.

Eibach R., Töpfer R., Hausmann L. 2010. Use of genetic diversity for grapevine resistance breeding. Mitteilungen Klosterneuburg 60, 332–337.

Elad Y., Vivier M., Fillinger S. 2016. *Botrytis*, the Good, the Bad and the Ugly. In: Fillinger S,, In: Elad Y, eds. Botrytis - The fungus, the pathogen and its management in agricultural systems. Cham: Springer, 1–15.

Emanuelli F., Lorenzi S., Grzeskowiak L., *et al.* 2013. Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. BMC Plant Biology 13, 39.

EU Commission. 2003. *Commission Recommendation of 23 July 2003 on guidelines for the development of national strategies and best practices to ensure the coexistence of genetically modified crops with conventional and organic farming (notified under document number C(2003) 262*.

EU Commission. 2022. Legislation for plants produced by certain new genomic techniques. Proposal for a regulation.

EU Commission, Directorate-General for Research and Innovation (Unit RTD.01 – Scientific Advice Mechanism). 2017. *New techniques in agricultural biotechnology - Explanatory note 02/2017*.

EU Council. 2009. *Council Regulation (EC) 491/2009 of 25 May 2009 amending Regulation (EC) 1234/2007 establishing a common organisation of agricultural markets and on specific provisions for certain agricultural products (single CMO regulation)*.

EU Parliament & Council. 2001. *Parliament and Council Directive 2001/18/EC of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC*.

EU Parliament & Council. 2003*a*. *Parliament and Council Regulation (EC) 1829/2003 of 22 September 2003 on genetically modified food and feed (Text with EEA relevance)*.

EU Parliament & Council. 2003*b*. *Parliament and Council Regulation (EC) 1830/2003 of 22 September 2003 concerning the traceability and labelling of genetically modified organisms and the traceability of food and feed products produced from genetically modified and amending Directive 2001*.

EU Parliament & Council. 2015. *Parliament and Council Directive 2015/412/EC of 11 March 2015 amending Directive 2001/18/EC as regards the possibility for the Member States to restrict or prohibit the cultivation of genetically modified organisms (GMOs) in their territory*.

Feechan A., Anderson C., Torregrosa L., *et al.* 2013. Genetic dissection of a TIR-NB-LRR locus from the wild North American grapevine species Muscadinia rotundifolia identifies paralogous genes conferring resistance to major fungal and oomycete pathogens in cultivated grapevine. Plant J 76, 661–674.

Feechan A., Jermakow A. M., Torregrosa L., Panstruga R., Dry I. B. 2008. Identification of grapevine MLO gene candidates involved in susceptibility to powdery mildew. Functional Plant Biology 35, 1255–1266.

Feechan A., Kocsis M., Riaz S., Zhang W., Gadoury D. M., Walker M. A., Dry I. B., Reisch B., Cadle-Davidson L. 2015. Strategies for RUN1 deployment using RUN2 and REN2 to manage grapevine powdery mildew informed by studies of race specificity. Phytopathology 105, 1104–1113.

Figueiredo A., Atanassov A. I., Baraldi E., *et al.* 2021. Double-Stranded RNA Targeting Dicer-Like Genes Compromises the Pathogenicity of Plasmopara viticola on Grapevine.

Fischer B. M., Salakhutdinov I., Akkurt M., Eibach R., Edwards K. J., Töpfer R., Zyprian E. M. 2004. Quantitative trait locus analysis of fungal disease resistance factors on a molecular map of grapevine. Theoretical and Applied Genetics 108, 501–515.

Foria S., Magris G., Jurman .I, Schwope R., De Candido M., De Luca E., Ivanišević D., Morgante M., Di Gaspero G. 2022. Extent of wild–to–crop interspecific introgression in grapevine (Vitis vinifera) as a consequence of resistance breeding and implications for the crop species definition. Horticulture Research 9, uhab010.

Foria S., Monte C., Testolin R., Di Gaspero G., Cipriani G. 2019. Pyramidizing resistance genes in grape: A breeding program for the selection of elite cultivars. Acta Horticulturae 1248, 549–554.

Forneck A., Walker M., Blaich R., Yvon M., Leclant F. 2001. Interaction of phylloxera (*Daktulosphaira vitifoliae* fitch) with grape (Vitis spp.) in simple isolation chambers. American Journal of Enology and Viticulture 52, 28–34.

Fu P., Wu W., Lai G., *et al.* 2020. Identifying *Plasmopara viticola* resistance loci in grapevine (*Vitis amurensis*) via genotyping-by-sequencing-based QTL mapping. Plant Physiology and Biochemistry 154, 75–84.

Gambino G., Dal Molin A., Boccacci P., *et al.* 2017. Whole-genome sequencing and SNV genotyping of 'Nebbiolo' (*Vitis vinifera* L.) clones. Scientific Reports 7, 17294.

Di Gaspero G., Cattonaro F. 2010. Application of genomics to grapevine improvement. Australian Journal of Grape and Wine Research 16, 122–130.

Di Gaspero G., Copetti D., Coleman C., *et al.* 2012. Selective sweep at the Rpv3 locus during grapevine breeding for downy mildew resistance. Theoretical and Applied Genetics 124, 277–286.

Gebremichael D. E., Haile Z. M., Negrini F., Sabbadini S., Capriotti L., Mezzetti B., Baraldi E. 2021.

Rna interference strategies for future management of plant pathogenic fungi: Prospects and challenges. Plants 10, 650.

Ghogare R., Ludwig Y., Bueno G. M., Slamet-Loedin I. H., Dhingra A. 2021. Genome editing reagent delivery in plants. Transgenic Research 30, 321–335.

Giacomelli L., Zeilmaker T., Malnoy M., Rouppe van der Voort J., Moser C. 2019. Generation of mildew-resistant grapevine clones via genome editing. Acta Horticulturae 1248, 195–200.

Giacomelli L., Zeilmaker T., Scintilla S., Salvagnin U., Rouppe van der Voort J., Moser C. 2022. Vitis vinifera plants edited in DMR6 genes show improved resistance to downy mildew. bioRxiv Preprint.

Girollet N., Rubio B., Lopez-Roques C., Valière S., Ollat N., Bert P. F. 2019. De novo phased assembly of the Vitis riparia grape genome. Scientific Data 6, 127.

Glass Z., Lee M., Li Y., Xu Q. 2018. Engineering the Delivery System for CRISPR-Based Genome Editing. Trends in Biotechnology 36, 173–185.

Grassi F., Arroyo-Garcia R. 2020. Editorial: Origins and Domestication of the Grape. Frontiers in Plant Science 11, 1176.

Guimier S., Delmotte F., Miclot A. S., Fabre F., Mazet I., Couture C., Schneider C., Delière L. 2019. OSCAR, a national observatory to support the durable deployment of disease-resistant grapevine cultivars. Acta Horticulturae 1248, 21–33.

Haile Z. M., Gebremichael D. E., Capriotti L., Molesini B., Negrini F., Collina M., Sabbadini S., Mezzetti B., Baraldi E. 2021. Double-stranded RNA targeting Dicer-like genes compromises the pathogenicity of *Plasmopara viticola* on grapevine. Frontiers in Plant Science 12, 667539.

Hasan M. M., Bae H. 2017. An overview of stress-induced resveratrol synthesis in grapes: perspectives for resveratrol-enriched grape products. Molecules 22, 294.

He R., Wu J., Zhang Y., Agüero C. B., Li X., Liu S., Wang C., Walker M. A., Lu J. 2017. Overexpression of a thaumatin-like protein gene from *Vitis amurensis* improves downy mildew resistance in *Vitis vinifera* grapevine. Protoplasma 254, 1579–1589.

van Heerden C. J., Burger P., Vermeulen A., Prins R. 2014. Detection of downy and powdery mildew resistance QTL in a 'Regent' × 'RedGlobe' population. Euphytica 200, 281–295.
Helms Jorgensen J. 1992. Discovery, characterization and exploitation of Mlo powdery mildew resistance in barley. Euphytica 63, 141–152.

Le Henanff G., Farine S., Kieffer-Mazet F., Miclot A. S., Heitz T., Mestre P., Bertsch C., Chong J. 2011. Vitis vinifera VvNPR1.1 is the functional ortholog of AtNPR1 and its overexpression in grapevine triggers constitutive activation of PR genes and enhanced resistance to powdery mildew. Planta 234, 405–417.

Hoffmann S., Di Gaspero G., Kovács L., Howard S., Kiss E., Galbács Z., Testolin R., Kozma P. 2008. Resistance to Erysiphe necator in the grapevine 'Kishmish vatkana' is controlled by a single locus through restriction of hyphal growth. Theoretical and Applied Genetics 116, 427–438.

Holme I. B., Wendt T., Holm P. B. 2013. Intragenesis and cisgenesis as alternatives to transgenic crop development. Plant Biotechnology Journal 11, 395–407.

Hu Y., Li Y., Hou F., *et al.* 2018. Ectopic expression of *Arabidopsis* broad-spectrum resistance gene RPW8.2 improves the resistance to powdery mildew in grapevine (*Vitis vinifera*). Plant science : an international journal of experimental plant biology 267, 20–31.

Hvarleva T., Bakalova A., Rusanov K., Diakova G., Ilieva I., Atanassov A., Atanassov I. 2009. Toward marker assisted selection for fungal disease resistance in grapevine. Biotechnology & Biotechnological Equipment 23, 1431–1435.

Hyma K. E., Barba P., Wang M., Londo J. P., Acharya C. B., Mitchell S. E., Sun Q., Reisch B., Cadle-Davidson L. 2015. Heterozygous mapping strategy (HetMappS) for high resolution genotyping-by-sequencing markers: A case study in grapevine. PLOS ONE 10, e0134880.
ISAAA. 2021. *Breaking barriers with breeding: A primer on new breeding innovations for food security*. Ithaca: The International Service for the Acquisition of Agri-biotech Applications (ISAAA).

Jacobsen E., Schouten H. J.. 2008. Cisgenesis, a new tool for traditional plant breeding, should be exempted from the regulation on genetically modified organisms in a step by step approach. Potato Research 51, 75–88.

Jaillon O., Aury J. M., Noel B., *et al.* 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449, 463–467.

Karn A., Zou C., Brooks S., Fresnedo-Ramírez J., Gabler F., Sun Q., Ramming D., Naegele R., Ledbetter C., Cadle-Davidson L. 2021. Discovery of the REN11 locus from Vitis aestivalis for stable resistance to grapevine powdery mildew in a family segregating for several unstable and tissue-specific quantitative resistance loci. Frontiers in Plant Science 12, 733899.

Kaur R., Choudhury A., Chauhan S., Ghosh A., Tiwari R., Rajam M. V. 2021. RNA interference and crop protection against biotic stresses. Physiology and Molecular Biology of Plants 27, 2357–2377.

Kumar S., Banks T. W., Cloutier S. 2012. SNP discovery through next-generation sequencing and its applications. International Journal of Plant Genomics 2012, 831460.

Lan X., Liu Y., Song S., Yin L., Xiang J., Qu J., Lu J. 2019. Plasmopara viticola effector PvRXLR131 suppresses plant immunity by targeting plant receptor-like kinase inhibitor BKI1. Molecular Plant Pathology 20, 765–783.

Lei X., Lan X., Ye W., Liu Y., Song S., Lu J. 2019. Plasmopara viticola effector PvRXLR159 suppresses immune responses in Nicotiana benthamiana. Plant Signaling and Behavior 14, e1682220.

Limera C., Sabbadini S., Sweet J. B., Mezzetti B. 2017. New biotechnological tools for the genetic improvement of major woody fruit species. Frontiers in Plant Science 8, 1–16.

Lin H., Leng H., Guo Y., Kondo S., Zhao Y., Shi G., Guo X. 2019. QTLs and candidate genes for downy mildew resistance conferred by interspecific grape (V. vinifera L. × V. amurensis Rupr.) crossing. Scientia Horticulturae 244, 200–207.

Liu J., Chen S., Ma T., Gao Y., Song S., Ye W., Lu J. 2021. Plasmopara viticola effector PvRXLR53 suppresses innate immunity in Nicotiana benthamiana. Plant Signaling and Behavior 16, 1846927.

Liu X., Lu T., Dou Y., Yu B., Zhang C. 2014. Identification of RNA silencing components in soybean and sorghum. BMC Bioinformatics 15, 281.

Liu S., Zhang C., Chao N., Lu J., Zhang Y. 2018. Cloning, characterization, and functional investigation of VaHAESA from Vitis amurensis inoculated with Plasmopara viticola. International Journal of Molecular Sciences 2018, Vol. 19, Page 1204 19, 1204.

Lu L.. M., Ickert-Bond S., Wen J. 2018. Recent advances in systematics and evolution of the grape family Vitaceae. Journal of Systematics and Evolution 56, 259–261.

Ma T., Chen S., Liu J., Fu P., Wu W., Song S., Gao Y., Ye W., Lu J. 2021. Plasmopara viticola effector PvRXLR111 stabilizes VvWRKY40 to promote virulence. Molecular Plant Pathology 22, 231–242.

Ma X., Zhang Q., Zhu Q., *et al.* 2015. A Robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. Molecular Plant 8, 1274–1284.

Magris G., Jurman I., Fornasiero A., Paparelli E., Schwope R., Marroni F., Di Gaspero G., Morgante M. 2021. The genomes of 204 Vitis vinifera accessions reveal the origin of European wine grapes. Nature Communications 12, 7240.

Mahanil S., Ramming D., Cadle-Davidson M., Owens C., Garris A., Myles S., Cadle-Davidson L. 2012. Development of marker sets useful in the early selection of Ren4 powdery mildew resistance and seedlessness for table and raisin grape breeding. Theoretical and Applied Genetics 124, 23–33.

Mallory A., Vaucheret H. 2012. Form, function, and regulation of ARGONAUTE proteins. The Plant Cell 22, 3879–3889.

Malnoy M., Viola R., Jung M. H., Koo O. J., Kim S., Kim J. S., Velasco R., Kanchiswamy C. N. 2016. DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 ribonucleoproteins. Frontiers in Plant Science 7, 1904.

Marcianò D., Ricciardi V., Marone Fassolo E., Passera A., Bianco P. A., Failla O., Casati P., Maddalena G., De Lorenzis G., Toffolatti S. L. 2021. RNAi of a putative grapevine susceptibility gene as a possible downy mildew control strategy. Frontiers in Plant Science 12, 667319.

Marguerit E., Boury C., Manicki A., Donnart M., Butterlin G., Némorin A., Wiedemann-Merdinoglu S., Merdinoglu D., Ollat N., Decroocq S. 2009. Genetic dissection of sex determinism, inflorescence morphology and downy mildew resistance in grapevine. Theoretical and Applied Genetics 118, 1261–1278.

Marguerit E., Brendel O., Lebon E., Van Leeuwen C., Ollat N. 2012. Rootstock control of scion transpiration and its acclimation to water deficit are controlled by different genes. New Phytologist 194, 416–429.

Marrano A., Birolo G., Prazzoli M. L., Lorenzi S., Valle G., Grando M. S. 2017. SNP-discovery by

RAD-sequencing in a germplasm collection of wild and cultivated grapevines (V. vinifera L.). PLoS ONE 12, e0170655.

Maul E. 2022. Vitis International Variety Catalogue.

Meggio F., Prinsi B., Negri A. S., *et al.* 2014. Biochemical and physiological responses of two grapevine rootstock genotypes to drought and salt treatments. Australian Journal of Grape and Wine Research 20, 310–323.

Meloni G., Swinnen J. 2013. The political economy of European wine regulations. Journal of Wine Economics 8, 244–284.

Mercati F., De Lorenzis G., Brancadoro L., Lupini A., Abenavoli M. R., Barbagallo M. G., Di Lorenzo R., Scienza A., Sunseri F. 2016. High-throughput 18K SNP array to assess genetic variability of the main grapevine cultivars from Sicily. Tree Genetics and Genomes 12, 59.

Merdinoglu D., Schneider C., Prado E., Wiedemann-Merdinoglu S., Mestre P. 2018. Breeding for durable resistance to downy and powdery mildew in grapevine. OENO One 52, 203–209.

Merdinoglu D., Wiedemann-Merdinoglu S., Coste P., Dumas V., Haetty S., Butterlin G., Greif C., Adam-Blondon A. F., Bouquet A., Pauquet J. 2003. Genetic analysis of downy mildew resistance derived from *Muscadinia rotundifolia*. Acta Horticulturae 603, 451–456.

Miedaner T. 2016. Breeding strategies for improving plant resistance to diseases. In: Al-Khayri J,, In: Jain S,, In: Johnson D, eds. Advances in plant breeding strategies: Agronomic, abiotic and biotic stress traits. Cham: Springer, 561–599.

Migliaro D., De Lorenzis G., Di Lorenzo G. S., De Nardi B., Gardiman M., Failla O., Brancadoro L., Crespan M. 2019. Grapevine non-vinifera genetic diversity assessed by simple sequence repeat markers as a starting point for new rootstock breeding programs. American Journal of Enology and Viticulture 70, 390–397.

Minio A., Massonnet M., Figueroa-Balderas R., Castro A., Cantu D. 2019. Diploid genome assembly of the wine grape Carménère. G3 Genes|Genomes|Genetics 9, 1331–1337.

Morales-Cruz A., Aguirre-Liguori J. A., Zhou Y., Minio A., Riaz S., Walker A. M., Cantu D., Gaut B. S. 2021. Introgression among North American wild grapes (Vitis) fuels biotic and abiotic adaptation. Genome Biology 22, 254.

Moreira F. M., Madini A., Marino R., Zulini L., Stefanini M., Velasco R., Kozma P., Grando M. S. 2011. Genetic linkage maps of two interspecific grape crosses (Vitis spp.) used to localize quantitative trait loci for downy mildew resistance. Tree Genetics and Genomes 7, 153–167.

Mu B., Chen J., Wang H., Kong W., Fan X., Wen X. 2022. An effector CSEP087 from Erysiphe necator targets arginine decarboxylase 2 VviADC to regulate host immunity in grapevine. Social Science Research Network Preprint.

Mundt C. C. 2018. Pyramiding for resistance durability: Theory and practice. Phytopathology 108, 792–802.

Myles S. 2013. Improving fruit and wine: what does genomics have to offer? Trends in Genetics

29, 190–196.

Myles S., Boyko A. R., Owens C. L., *et al.* 2011. Genetic structure and domestication history of the grape. Proceedings of the National Academy of Sciences 108, 3530–3535.

Myles S., Chia J. -M., Hurwitz B., Simon C., Zhong G. Y., Buckler E., Ware D. 2010. Rapid genomic characterization of the genus Vitis. PLoS ONE 5, e8219.

Nakajima I., Ban Y., Azuma A., Onoue N., Moriguchi T., Yamamoto T., Toki S., Endo M. 2017. CRISPR/Cas9-mediated targeted mutagenesis in grape. PLoS ONE 12, e0177966.

Nakasugi K., Crowhurst R. N., Bally J., Wood C. C., Hellens R. P., Waterhouse P. M. 2013. De novo transcriptome sequence assembly and analysis of RNA silencing genes of Nicotiana benthamiana. PLOS ONE 8, e59534.

Nerva L., Sandrini M., Gambino G., Chitarra W. 2020. Double-stranded RNAs (DsRNAs) as a sustainable tool against gray mold (Botrytis cinerea) in grapevine: Effectiveness of different application methods in an open-air environment. Biomolecules 10, 200.

Nirala N. K., Das D. K., Srivastava P. S., Sopory S. K., Upadhyaya K. C. 2010. Expression of a rice chitinase gene enhances antifungal potential in transgenic grapevine (*Vitis vinifera* L.). VITIS - Journal of Grapevine Research 49, 181–187.

Niu X., Tang W., Huang W., *et al.* 2008. RNAi-directed downregulation of OsBADH2 results in aroma (2-acetyl-1-pyrroline) production in rice (Oryza sativa L.). BMC Plant Biology 8, 100.
Nookaraju A, Agrawal DC. 2012. Enhanced tolerance of transgenic grapevines expressing chitinase and β-1,3-glucanase genes to downy mildew. Plant Cell, Tissue and Organ Culture 111, 15–28.

Numata K., Horii Y., Oikawa K., Miyagi Y., Demura T., Ohtani M. 2018. Library screening of cell-penetrating peptide for BY-2 cells, leaves of Arabidopsis, tobacco, tomato, poplar, and rice callus. Scientific Reports 8, 10966.

Ochssner I., Hausmann L., Töpfer R. 2016. Rpvl4, a new genetic source for Plasmopara viticola resistance conferred by Vitis cinerea. Vitis - Journal of Grapevine Research 55, 79–81.

OIV. 2020. 2020 Wine production - OIV first estimates. International Organisation of Vine and Wine, 1–8.

Osakabe Y., Liang Z., Ren C., *et al.* 2018. CRISPR–Cas9-mediated genome editing in apple and grapevine. Nature Protocols 2018 13:12 13, 2844–2863.

Owens C. L. 2003. SNP detection and genotyping in Vitis. Acta Horticulturae 603, 139–140.

Panstruga R. 2005. Serpentine plant MLO proteins as entry portals for powdery mildew fungi. Biochemical Society Transactions 33, 389–392.

Pap D., Riaz S., Dry I. B., Jermakow A., Tenscher A. C., Cantu D., Oláh R., Walker M. A. 2016. Identification of two novel powdery mildew resistance loci, Ren6 and Ren7, from the wild Chinese grape species Vitis piasezkii. BMC Plant Biology 16, 170.

Le Paslier M. -C., Choisne N., Bacilieri R., *et al.* 2013. The GrapeReSeq 18k Vitis genotyping chip. 9th International symposium grapevine physiology and biotechnology. La Serena: International Society for Horticultural Science, 123.

Pecile M., Zavaglia C. Registro Nazionale delle Varietà di Vite.

Peressotti E., Wiedemann-Merdinoglu S., Delmotte F., Bellin D., Di Gaspero G., Testolin R., Merdinoglu D., Mestre P. 2010. Breakdown of resistance to grapevine downy mildew upon limited deployment of a resistant variety. BMC Plant Biology 10, 147.

Pertot I., Caffi T., Rossi V., *et al.* 2017. A critical review of plant protection tools for reducing pesticide use on grapevine and new perspectives for the implementation of IPM in viticulture. Crop Protection 97, 70–84.

Pessina S., Lenzi L., Perazzolli M., Campa M., Dalla Costa L., Urso S., Valè G., Salamini F., Velasco R., Malnoy M. 2016. Knockdown of MLO genes reduces susceptibility to powdery mildew in grapevine. Horticulture research 3, 16016.

Peterson G. W., Dong Y., Horbach C., Fu Y. B. 2014. Genotyping-by-sequencing for plant genetic diversity analysis: A lab guide for SNP genotyping. Diversity 6, 665–680.

Pirrello C., Malacarne G., Moretto M., Lenzi L., Perazzolli M., Zeilmaker T., Van Den Ackerveken G., Pilati S., Moser C., Giacomelli L. 2022. Grapevine DMR6-1 is a candidate gene for susceptibility to downy mildew. Biomolecules 12, 182.

Pirrello C., Mizzotti C., Tomazetti T. C., *et al.* 2019. Emergent Ascomycetes in viticulture: An interdisciplinary overview. Frontiers in Plant Science 10, 1394.

Pirrello C., Zeilmaker T., Bianco L., Giacomelli L., Moser C., Vezzulli S. 2021. Mining grapevine downy mildew susceptibility genes: A resource for genomics-based breeding and tailored gene editing. Biomolecules 11, 181.

PIWI International. PIWI International.

Qiao L., Lan C., Capriotti L., *et al.* 2021. Spray-induced gene silencing for disease control is dependent on the efficiency of pathogen RNA uptake. Plant Biotechnology Journal 19, 1756–1768.

Reinstädler A., Müller J., Czembor J. H., Piffanelli P., Panstruga R. 2010. Novel induced mlo mutant alleles in combination with site-directed mutagenesis reveal functionally important domains in the heptahelical barley Mlo protein. BMC Plant Biology 10, 31.

Ren C., Liu Y., Guo Y., Duan W., Fan P., Li S., Liang Z. 2021. Optimizing the CRISPR/Cas9 system for genome editing in grape by using grape promoters. Horticulture Research 8, 52.

Ren C., Liu X., Zhang Z., Wang Y., Duan W., Li S., Liang Z. 2016. CRISPR/Cas9-mediated efficient targeted mutagenesis in Chardonnay (Vitis vinifera L.). Scientific Reports 31, 32289.

Ren F., Ren C., Zhang Z., Duan W., Lecourieux D., Li S., Liang Z. 2019. Efficiency optimization of CRISPR/CAS9-mediated targeted mutagenesis in grape. Frontiers in Plant Science 10, 612.

Riaz S., Hu R., Walker M. A. 2012. A framework genetic map of Muscadinia rotundifolia. Theoretical and Applied Genetics 125, 1195–1210.

Riaz S., De Lorenzis G., Velasco D., *et al.* 2018. Genetic diversity analysis of cultivated and wild grapevine (Vitis vinifera L.) accessions around the Mediterranean basin and Central Asia. BMC Plant Biology 18, 137.

Riaz S., Tenscher A., Pap D., Romero N., Walker M. A. 2019. Durable powdery mildew resistance in grapevines: Myth or reality. Acta Horticulturae 1248, 595–600.

Riaz S., Tenscher A. C., Ramming D. W., Walker M. A. 2011. Using a limited mapping strategy to identify major QTLs for resistance to grapevine powdery mildew (Erysiphe necator) and their use in marker-assisted breeding. Theoretical and Applied Genetics 122, 1059–1073.

Roach M. J., Johnson D. L., Bohlmann J., van Vuuren H. J. J., Jones S. J. M., Pretorius I. S., Schmidt S. A., Borneman A. R. 2018. Population sequencing reveals clonal diversity and ancestral inbreeding in the grapevine cultivar Chardonnay. PLOS Genetics 14, e1007807.

Rommens C. M., Haring M. A., Swords K., Davies H. V., Belknap W. R. 2007. The intragenic approach as a new extension to traditional plant breeding. Trends in Plant Science 12, 397–403.

Rubio J., Montes C., Castro Á., *et al.* 2015. Genetically engineered Thompson Seedless grapevine plants designed for fungal tolerance: selection and characterization of the best performing individuals in a field trial. Transgenic Research 24, 43–60.

Saifert L., Sánchez-Mora F. D., Assumpção W. T., Zanghelini J. A., Giacometti R., Novak E. I., Dal Vesco L. L., Nodari R. O., Eibach R., Welter L. J. 2018. Marker-assisted pyramiding of resistance loci to grape downy mildew. Pesquisa Agropecuaria Brasileira 53, 602–610.

Sang H., Kim J. -I. 2020. Advanced strategies to control plant pathogenic fungi by host-induced gene silencing (HIGS) and spray-induced gene silencing (SIGS). Plant Biotechnology Reports 14, 1–8.

Sapkota S. D., Chen L. L., Yang S., Hyma K. E., Cadle-Davidson L. E., Hwang C. F. 2019. Quantitative trait locus mapping of downy mildew and *botrytis* bunch rot resistance in a *Vitis aestivalisderived* 'Norton'-based population. Acta Horticulturae 1248, 305–311.

Sargolzaei M., Maddalena G., Bitsadze N., Maghradze D., Bianco P. A., Failla O., Toffolatti S. L., De Lorenzis G. 2020. Rpv29, Rpv30 and Rpv31: Three novel genomic loci associated with resistance to *Plasmopara viticola* in *Vitis vinifera*. Frontiers in Plant Science 11, 562432.

van Schie C. C. N., Takken F. L. W. 2014. Susceptibility genes 101: How to be a good host. Annual Review of Phytopathology 52, 551–581.

Schouten H. J., Krens F. A., Jacobsen E. 2006. Cisgenic plants are similar to traditionally bred plants: International regulations for genetically modified organisms should be altered to exempt cisgenesis. EMBO Reports 7, 750.

Schwander F., Eibach R., Fechter I., Hausmann L., Zyprian E., Töpfer R. 2012. Rpv10: A new locus from the Asian Vitis gene pool for pyramiding downy mildew resistance loci in grapevine. Theoretical and Applied Genetics 124, 163–176.

Scintilla S., Salvagnin U., Giacomelli L., Zeilmaker T., Malnoy M. A., van der Voort J. R., Moser C. 2021. Regeneration of plants from DNA-free edited grapevine protoplasts. bioRxiv Preprint.

Soulié M. C., Piffeteau A., Choquer M., Boccara M., Vidal-Cros A. 2003. Disruption of Botrytis cinerea class I chitin synthase gene Bcchs1 results in cell wall weakening and reduced virulence. Fungal genetics and biology 40, 38–46.

Stam R., McDonald B. A. 2018. When resistance gene pyramids are not durable—the role of pathogen diversity. Molecular Plant Pathology 19, 521.

Stuthman D. D., Leonard KJ, Miller-Garvin J. 2007. Breeding crops for durable resistance to disease. Advances in Agronomy 95, 319–367.

Su H., Jiao Y. T., Wang F. F., Liu Y. E., Niu W. L., Liu G. T., Xu Y. 2018. Overexpression of VpPR10.1 by an efficient transformation method enhances downy mildew resistance in V. vinifera. Plant Cell Reports 37, 819–832.

Teh S. L., Fresnedo-Ramírez J., Clark M. D., Gadoury D. M., Sun Q., Cadle-Davidson L., Luby J. J. 2017. Genetic dissection of powdery mildew resistance in interspecific half-sib grapevine families using SNP-based maps. Molecular Breeding 37, 1.

Telem S. R., Wani H. S., Singh B. N., Nandini R., Sadhukhan R., Bhattacharya S., Mandal N. 2013. Cisgenics - A Sustainable approach for crop improvement. Current Genomics 14, 468–476.

Thatcher L. F., Powell J. J., Aitken E. A. B., Kazan K., Manners J. M. 2012. The lateral organ boundaries domain transcription factor LBD20 functions in Fusarium wilt Susceptibility and jasmonate signaling in Arabidopsis. Plant physiology 160, 407–418.

Toffolatti S. L., De Lorenzis G., Brilli M., *et al.* 2020. Novel aspects on the interaction between grapevine and Plasmopara viticola: dual-RNA-seq analysis highlights gene expression dynamics in the pathogen and the plant during the battle for infection. Genes 11, 261.

Toffolatti S. L., De Lorenzis G., Costa A., *et al.* 2018. Unique resistance traits against downy mildew from the center of origin of grapevine (*Vitis vinifera*). Scientific Reports 8, 12523.

Topfer R., Hausmann L. 2021. Table of loci for traits in grapevine relevant for breeding and genetics. Vitis International Variety Catalogue (VIVC).

Topfer R., Hausmann L., Harst M., Maul E., Zyprian E., Eibach R. 2011. New horizons for grapevine breeding. Fruit, Vegetable and Cereal Science and Biotechnology 5, 79–100.

Vannozzi A., Donnini S., Vigani G., Corso M., Valle G., Vitulo N., Bonghi C., Zocchi G., Lucchin M. 2017. Transcriptional characterization of a widely-used grapevine rootstock genotype under different iron-limited conditions. Frontiers in Plant Science 7, 1994.

Varshney R. K., Nayak S. N., May G. D., Jackson S. A. 2009. Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends in Biotechnology 27, 522–530.

Vaucheret H. 2006. Post-transcriptional small RNA pathways in plants: mechanisms and

regulations. Genes & development 20, 759–771.

Vaucheret H. 2008. Plant ARGONAUTES. Trends in Plant Science 13, 350–358.

Velasco R., Zharkikh A., Troggio M., *et al.* 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS ONE 2, e1326.

Venuti S., Copetti D., Foria S., *et al.* 2013. Historical introgression of the downy mildew resistance gene Rpv12 from the Asian species *Vitis amurensis* into grapevine varieties. PLOS ONE 8, e61228.

Vezzulli S., Doligez A., Bellin D. 2019*a*. Molecular mapping of grapevine genes. In: Cantu D,, In: Walker M., eds. The Grape Genome. Cham: Springer, 103–136.

Vezzulli S., Gramaje D., Tello J., *et al.* 2022. Genomic designing for biotic stress resistant grapevine. In: Kole C., ed. Genomic Designing for Biotic Stress Resistant Fruit Crops. Cham: Springer Nature, 87–255.

Vezzulli S., Malacarne G., Masuero D., *et al.* 2019*b*. The Rpv3-3 haplotype and stilbenoid induction mediate downy mildew resistance in a grapevine interspecific population. Frontiers in Plant Science 10, 234.

Vezzulli S., Troggio M., Coppola G., *et al.* 2008. A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers. Theoretical and Applied Genetics 117, 499–511.

Vezzulli S., Zulini L., Stefanini M. 2019*c*. Genetics-assisted breeding for downy/powdery mildew and phylloxera resistance at fem. BIO Web of Conferences.01020.

Vives-Vallés J. A., Collonnier C. 2020. The judgment of the CJEU of 25 July 2018 on mutagenesis: Interpretation and interim legislative proposal. Frontiers in Plant Science 10, 1813.

Walker M. A., Wolpert J. A., Weber E. 1994. Viticultural characteristics of VR hybrid rootstocks in a vineyard site infected with grapevine fanleaf virus. VITIS - Journal of Grapevine Research 33, 19–23.

Wang X., Tu M., Wang D., Liu J., Li Y., Li Z., Wang Y., Wang X. 2018. CRISPR/Cas9-mediated efficient targeted mutagenesis in grape in the first generation. Plant Biotechnology Journal 16, 844–855.

Wang M., Weiberg A., Lin F. -M., Thomma B., Huang H. -D., Jin H. Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection HHS Public Access.

Wang J., Yao W., Wang L., *et al.* 2017. Overexpression of VpEIFP1, a novel F-box/Kelch-repeat protein from wild Chinese Vitis pseudoreticulata, confers higher tolerance to powdery mildew by inducing thioredoxin z proteolysis. Plant Science 263, 142–155.

Welter L. J., Göktürk-Baydar N., Akkurt M., Maul E., Eibach R., Töpfer R., Zyprian E. M. 2007. Genetic mapping and localization of quantitative trait loci affecting fungal disease resistance and leaf morphology in grapevine (*Vitis vinifera* L). Molecular Breeding 20, 359–374.

Wen J., Lu L. M., Nie Z. L., Liu X. Q., Zhang N., Ickert-Bond S., Gerrath J., Manchester S. R., Boggan

J., Chen Z. D. 2018. A new phylogenetic tribal classification of the grape family (*Vitaceae)*. Journal of Systematics and Evolution 56, 262–272.

Wiedemann-Merdinoglu S., Prado E., Coste P., Dumas V., Butterlin G., Bouquet A., Merdinoglu D. 2006. Resistance to downy mildew derived from *Muscadinia rotundifolia*: genetic analysis and use of molecular markers for breeding. Proc. 5th Int. Workshop on Grapevine Downy Mildew and Powdery Mildew. San Michele all'Adige, .

Wilcox W. F., Gubler W. D., Uyemoto J. K. 2015. *Compendium of grape diseases, disorders, and pests* (W. F. Wilcox, W. D. Gubler, and J. K. Uyemoto, Eds.). St. Paul, Minnesota: The American Phytopathological Society.

Woolhouse M. E. J., Webster J. P., Domingo E., Charlesworth B., Levin B. R. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. Nature Genetics 32, 569–577.

Xiang J., Li X., Wu J., Yin L., Zhang Y., Lu J. 2016. Studying the mechanism of *Plasmopara viticola* RxLR effectors on suppressing plant immunity. Frontiers in Microbiology 7, 709.

Yamamoto T., Iketani H., Ieki H., Nishizawa Y., Notsuka K., Hibi T., Hayashi T., Matsuta N. 2000. Transgenic grapevine plants expressing a rice chitinase with enhanced resistance to fungal pathogens. Plant Cell Reports 19, 639–646.

Yang S., Fresnedo-Ramírez J., Wang M., *et al.* 2016. A next-generation marker genotyping platform (AmpSeq) in heterozygous crops: A case study for marker-assisted selection in grapevine. Horticulture Research 3, 16002.

Zendler D., Schneider P., Töpfer R., Zyprian E. 2017. Fine mapping of Ren3 reveals two loci mediating hypersensitive response against *Erysiphe necator* in grapevine. Euphytica 213, 68.

Zhang R. 2021. RALPH effectors from *Erysiphe necator*, the causing agent of grapevine powdery mildew and chitin perception pathway of grapevine. PhD Thesis.

Zini E., Dolzani C., Stefanini M., *et al.* 2019. R-loci arrangement versus downy and powdery mildew resistance level: A Vitis hybrid survey. International Journal of Molecular Sciences 20, 3526.

Zyprian E., Ochßner I., Schwander F., *et al.* 2016. Quantitative trait loci affecting pathogen resistance and ripening of grapevines. Molecular Genetics and Genomics 291, 1573–1594.