UNIVERSITÀ DEGLI STUDI DI VERONA

DIPARTIMENTO DI BIOTECNOLOGIE

Scuola di dottorato in scienze naturali e ingegneristiche

Dottorato di ricerca in biotecnologie

XXXIV ciclo

Dissection of genetic determinants of phenology and qualitative traits in *Vitis vinifera* through two complementary genetic approaches

S.S.D. AGR/07

Coordinatore: Prof. Matteo Ballottari

Responsabile scientifico: Prof.ssa Diana Bellin

Dottorando: Dott. Riccardo Mora

Matricola VR430500

Solo quando è buio riusciamo a vedere le stelle.

Martin Luther King

Table of contents

Acronyms page 2

Summary page 3

Background

Vitis vinifera L. page 5
Genome and genetics of Grapevine page 7
Phenology of grapevine page 10
Grapevine cultivation in a changing environment page 14

Chapter 1: ***QTL mapping in a segregant population Cabernet Sauvignon x Corvina for the dissection of genetic determinants on phenological and qualitative traits***

1. Abstract page 16
2. Introduction page 17
3. Materials and methods page 21
4. Results page 42
5. Discussion page 88

Chapter 2: ***Validation of association results from a genome-wide association study to target genetic regulators of phenological events in a grapevine collection***

1. Abstract page 97
2. Introduction page 99
3. Materials and methods page 103
4. Results page 114
5. Discussion page 130

Conclusions page 133

Bibliography page 136

Appendix page 148

Acknowledgments page 199

Acronyms

ABA= Abscisic acid

bp= Base pair

GO= gene ontology

IM= Interval Mapping

LOD= logarithm of the odds

MAS= Marker-Assisted Selection

Mbp= mega base

Kb= kilobase

MQM= Multiple-QTL mapping

NGS= next generation sequencing

SNP= single nucleotide polymorphism

SSR= Single sequence repeat

Summary

In the global agroeconomy, grapevine occupies one of the first positions in the ranking of commercial crops with the highest impact on the market, thanks to its many products, but especially for wine production. As for many other crops, also grapevine cultivation is afflicted by environmental changes. Long-term studies on climatology and grapevine phenology revealed that global warming is affecting the onset and duration of phenological events, with an acceleration in their timing in many areas. All these processes depend on environmental conditions and cultivar features. The timing with which phenological stages occurs, especially ripening, is highly relevant and critical for berry development. More in detail, too high temperatures during ripening lead to negative biochemical effects, that all together result in an alteration of fruits development and wines quality. While previous studies have demonstrated the strong influence of environment on these biochemical changes, the genetic control of these processes are for the most part still unknown. Adapting local economically relevant varieties to future environment conditions is gaining relevance to preserve the peculiarities of grapes and wines grown in specific area like *Corvina*, a red berry grape cultivated in Verona's area to produce prestigious and world-wide recognized DOC wines like Valpolicella and Amarone. Therefore, many research teams have attempted to elucidate genetic determinants of veraison time to provide useful information with the aim of developing local varieties better adapted to future climatic conditions.

The final aim of this project is the identification of such candidate genes as genetic determinants for quality and phenological traits in *V. vinifera*, especially veraison time, flowering time and the interval among these. The future perspective is to use this genetic information to start breeding programs aimed to adapt the phenology of such relevant local varieties to new climate through the use of new breeding techniques, like CRISPR-Cas9 (Rao et al., 2021) or genomic prediction (Brault et al., 2022).

In grapevine, as in many other crops, phenotypic traits show a complex quantitative inheritance with small additive or dominance effects of individual genes. By applying a quantitative trait loci (QTLs) analysis it is possible to identify segments of genome linked to observed variation that include candidate genes potentially involved in the traits. In this study, the dissection of genetic determinants has been attempted by combining two genetic approaches: QTL mapping and Genome-Wide Association Study (GWAS).

The thesis is divided in two chapters: in the first chapter is described a QTL analysis performed with a QTL mapping approach, which required phenotypic (traits measurement) and genotypic (molecular markers) information on a bi-parental segregant population to find statistical association between markers and phenotypes and to explain the genetic basis of variation of complex traits. To this purpose, a large segregant population obtained by crossing *Corvina* with the international variety *Cabernet Sauvignon* (CsxC) was considered to perform the QTL mapping. CsxC population was propagated in few copies in two locations, genotyped with *Vitis*18KSNP chip (Le Paslier et al., 2013) and phenotyped for 2 years to assemble a consensus linkage map and a QTL map by using traits measurement collected in two vintages.

In the second chapter, the results of a previous GWAS analysis were validated, as a complementary approach for QTLs identification. GWAS requires a larger fraction of diversity within a species for the analysis. In a previous study of our laboratory, a core collection composed by 132 individuals of the CREA-VE Conegliano (TV, Italy) grapevine germplasm collection was created after genotyping and phenotypic assessment, as minimal number to represent its genetic diversity. Association analysis was then performed by integration of genetic data provided from GrapeReSeq *Vitis*18KSNP Chip and phenotypic data (flowering time, veraison time and interval flowering-veraison) using different software. Fifteen SNPs significantly associated to veraison time and to the interval flowering-veraison were found. In this study, 3 of these 15 SNPs were validated as significantly associated to the phenotypic trait flowering–veraison interval on a subset of 94 individual of the collection with extreme phenotype.

Background

Vitis vinifera L.

Grapevine is a perennial woody plant cultivated all over the world that belongs to *Vitaceae* family, composed by 14 genera and 910 known species (Christenhusz and Byng, 2016). The genus *Vitis* is the most common of the family and contains 79 species largely diffused in Boreal Hemisphere, more specifically in temperate zones such as America, Europe and Asia.

To the genus *Vitis* belongs *Euvitis* (n=19) and *Muscadinia* (n=20), two subgenera not interfertile due to diploidization events (Aradhya et al., 2013; Jaillon et al., 2007). The subgenus *Muscadinia* includes 3 species native of America: *Vitis munsioniana*, *V. rutondifolia* and *V. popenoei*. In the subgenus *Euvitis* are grouped many different species interfertile and classified according to their geographic area of origin (Europe and Western Asia, North America, Oriental Asia). Among those of North American origin, at least 25 species are widely used as rootstocks, as carriers of different resistances to biotic and abiotic stress (This et al., 2006). Among the Asian species we find about 40 different species, including *V. amurensis*.

Vitis vinifera L. is the only *Vitis* species native of Europe and it is composed by two subspecies: *Vitis vinifera* L. subspecies *sylvestris*, the common wild grapevine, and *Vitis vinifera* L. ssp *vinifera* (*V. vinifera* from now on), the worldwide cultivated grapevine (Boss et al., 2003).

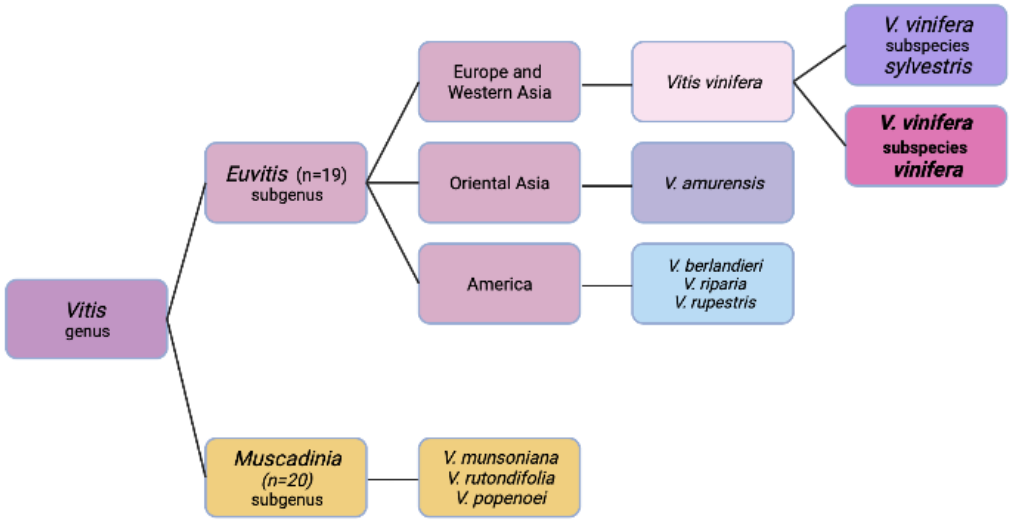


Figure 1. Taxonomy of *Vitis* genus

The first references to the cultivation of vines date back to around 8000 B.C. in Mesopotamia. Grapevine cultivation then spread as the main culture throughout the Mediterranean basin, in Greece, Magna Graecia, France and Spain (McGovern et al., 2003; This et al., 2006; Myles et al., 2011). In later times, it also spread to the rest of Europe, America and Australia and today most of the world's cultivation is concentrated in Europe (Italy and France).

Grapevine is a monoecious plant whose reproduction occurs by autogamous pollination. Being the grapevine highly heterozygous, this type of reproduction produces new allelic combinations of the parental alleles, thus causing phenotypic variability and segregation that could compromise the qualitative characteristics of the progeny. In commercial crops, therefore, seed propagation is avoided but clonal propagation is preferred. However, cross-pollination can be an interesting genetic tool to create new varieties that combine the characteristics of different cultivars, presenting the best qualities of each parental variety.

The maintenance and propagation of genotypes for commercial purposes therefore takes place, as mentioned, vegetatively, by grafting. The practice of grafting has spread since the XIX century, when European grapevines suffered serious damage

due to the spread of phylloxera (*Daktulosphaira vitifoliae*) from America, even though recent findings suggest it was in use even earlier (Madrigal et al., 2019). Cultivars cultivated for commercial purposes are reproduced by means of grafts placed on rootstocks of American varieties that are resistant to phylloxera and which often also allow a better adaptation of the plant to the variable soil climate of the territory. Indeed, market globalization has led to a sharp reduction in the number of varieties used for production purposes in the name of standardization of production (Santos et al., 2014). Most of this genetic diversity is therefore maintained in germplasm collections or varietal collections (Laucou et al., 2011; Cipriani et al., 2010).

Genome and genetics of Grapevine

Grapevine has a genome size of about 475-500 Mbp (Jaillon et al., 2007) and about 42,414 genes annotated (Canguier et al., 2017).

Molecular studies on grapevine genetics began at the end of 90's, when the first genetic map of *V. vinifera* was built by Lohdi et al. (1995), while in 2006 the first integrated genetic map, built from five different populations and composed by 513 SSR marker, was carried out and subsequently used as a reference map to assign a position on the chromosomes to the sequences of the genome. In 2007 a first version of reference genome sequence was made available from Jaillon et al. (2007); the sequence was built from a cultivated clone of Pinot Noir PN40024 with high homozygosity (93%), reached by recurrent selfings. The reference sequence was firstly created with Sanger sequencing assembly, with an average 8X coverage of genome (68,9%). Later, it was updated with the addition of sequences, reaching a new coverage of 12X in 2010, named PN40024-12X.v0 and available on the International Grape Genome Browser Genescope web site (obsolete now). This assembly was annotated more times by Centro di Ricerca Interdipartimentale per le Biotecnologie Innovative (CRIBI) releasing the V1 and V2 annotation (Vitulo et al., 2014). More recently the assembly was improved and updated with mate-pairs and genetic information derived from 6 different genetic maps to the version

12X.v2 by the French-Italian Public Consortium (Canaguier et al., 2017), available on URGI JBrowse. Previous annotations were projected on the new assembly as well as new annotations were released. Currently a new more complete assembly (12X.v4) of grapevine genome is in preparation.

Thanks to the construction of a reference genome, it was possible to start molecular studies aimed to the knowledge and genetic improvement of grapevine.
Genetic markers represent reference points for the exploration of the genome and for the positioning and identification of interesting loci for the trait under investigation or for population structure analysis.

Among the different typologies of molecular markers, SSR and SNP are the most frequently used nowadays for the study of the vine genome.
Simple Sequence Repeats (SSRs, Tautz and Renz 1984), also known as "microsatellites" are small units of non-coding nucleotide motifs (1-6 bp) repeated in tandem and widely distributed along plant genomes. In addition to their co-dominance properties, SSRs are informative markers as they are highly variable between alleles (Thomas and Scott 1993; Bowers et al., 1996, 1999; Di Gaspero et al., 2005). Thanks to numerous sequencings on the grapevine genome over the years, there are currently more than 400 SSR sequences. SSR in grapevine are used mainly in the varietal identification of grapevine cultivars (This et al., 2004, see also the *Vitis* International Variety Catalog (VIVC) website), but also for MAS (Marker-Assisted Selection), analysis of population structure/genetic relationships among cultivars and for mapping (Bacilieri et al., 2013; Imazio et al., 2013; Migliaro et al., 2017; Cipriani et al., 2008, 2010).

Single Nucleotide Polymorphism (SNP) represent the most used molecular marker kind in modern genetics, especially thanks to the advent of Next Generation Sequencing technologies (NGS, Varshney et al., 2009), which, through high-throughput analysis, allow their easy identification and simultaneously automated analysis (Ganal et al., 2009). They are variations of a single base in the DNA sequence between different individuals which recur with a frequency of at least 1%

in the population. SNPs can fall both in non-coding portions of the genome (intronic regions) and in coding regions (exonic regions), causing in some cases changes in the sequences and function of gene products. Furthermore, when multiple SNPs are in close proximity to each other, they can be inherited in blocks forming haplotypes. They are bi-allelic, co-dominant and highly reproducible.

As already mentioned, with the advent of NGS technologies, the study of grapevine genetics has undergone a considerable upgrade. Molecular methodologies based on this sequencing approach allow high-throughput SNP genotyping. For example, the widely used Genotyping by Sequencing (GBS) approach does not require previous knowledge on the genomic regions and allows the identification of new SNPs directly during the genotyping phase. However, this approach returns sequenced regions with low coverage and thus loss of information on some SNPs may occur (Elshire et al., 2011).

As an alternative, for SNPs high throughput genotyping technologies based on the hybridization of oligonucleotides anchored on arrays have been developed with which it is possible to analyze thousands of known SNPs simultaneously (Myles et al., 2010).
The introduction of hybridization-based genotyping on SNP chips has allowed the development of relatively inexpensive and standardized high-throughput genotyping strategies. Compared to GBS, array genotyping allows greater reproducibility and a lower error rate of the results between different populations (Mardis, 2008; Houel et al., 2015). Importantly, in GBS approaches different sets of markers are used in different populations. The use of unique sets of SNP markers positioned on an array allows a much easier comparison of data obtained from different populations. SNP chip arrays have been developed in recent years for different fruit crops, including apple, peach, pear, cherry, strawberry and grapevine. Following the release of the genomic sequence and annotation (Jaillon et al., 2007; Canaguier et al., 2017), a first grapevine SNP array was developed, the *Vitis*9K SNP chip, based on a set of SNPs identified by resequencing 11 cultivars of *V. vinifera* and six wild species (Myles et al., 2010).

Few years later, as part of the GrapeReSeq project, 47 cultivars and 18 species of *V. vinifera* have been re-sequenced and the identified SNPs have been gathered in the final version of grapevine array, *Vitis*18KSNP chip (Le Paslier et al., 2013) which, thanks to the simultaneous genotyping of 18071 SNPs on multiple samples, allows characterization and analysis of genetic variability like population diversity, structure, relatedness and linkage maps construction (De Lorenzis et al., 2015; Sunseri et al., 2018; Mercati et al., 2016; Laucou et al., 2018).
Alternatively, the classic Sanger sequencing can represent a valid option for the characterization of regions containing haplotypes, through amplification with specific primers, purification and sequencing of the same region on multiple individuals.

Phenology of grapevine

In plant biology, the agricultural term phenology indicates the succession of development stages (or phenological phases) of cultivated plants in relation to the environmental conditions. The developmental events are usually classified by using phenological scales that allow the recording and comparison of the dates in which certain developmental events occur in different environments.
Thus, the study of the phenological times of the grapevine plays a fundamental role in understanding the mechanisms that govern them and in genetic improvement.

Vitis vinifera has tree fundamental phenological stages: budbreak, flowering and fruit development.
Budbreak occurs at the beginning of the spring season and represents the beginning of the vegetative phase. In fact, in this phase, the dormant buds formed in the previous season and inhibited by the low winter temperatures are awakened by the milder temperatures.
The phenomenon of budding consists in the enlargement of the buds and the consequent exit of the sprout, controlled by factors external and internal to the plant: environmental factors, such as temperature and rainfall and biotic factors, such as

the position of the buds on the shoot, their vigor and cultivation techniques. After budding, the vegetative apex begins to lengthen at an ever-increasing rate, reaching its maximum in mid-June (2-4 cm/day). During this phase, the sprouts undergo lignification as well as the processing and accumulation of reserve substances in the form of starch. The winter rest period begins with the fall of the leaves, promoted by the accumulation of abscisic acid (ABA) and ends with the vegetative restart of the following season.

Flowering marks the onset of the reproductive stage, which takes place between the end of May and the beginning of June. The flower of cultivated grapevine is autogamous-cleistogamous and its complete development corresponds to the detachment of the calyptra from the floral receptacle. After the development of male and female gametes and the related meiotic recombination, fertilization take place. Usually, in biology this part of the development cycle is crucial to increase the genetic diversity of a species. However, for commercial vineyards, its primary objective is fruit production and clonal propagation remains still preferred way of propagation that allow to maintain all qualities of the cultivar.

The complete formation of the inflorescences and the complete development of each flower are not simultaneous: the process can last from a week to fifteen days, depending on environmental and cultivar characteristics. Environmental conditions that allow a good flowering and therefore a good fertilization are temperatures above 15 ° C and not above 35 ° C, good insolation and photosynthetic activity (Buttrose, 1974; Mullins et al., 1992), control of pathogens, balance between vegetative activity and reproductive activity of the plant and cultivation practices, depending on the grape variety. However, in addition to environmental factors, cultivar-specific characteristics influence the late or early flowering.

Approximately 60 days after flowering (depending on grape cultivar) starts the berry maturation phase. Berry development exhibits a double sigmoid growth pattern (**Fig. 2**). In the first phase, growth occurs under hormonal control of auxins, cytokinins and gibberellins (**Fig. 3**) mostly by cell division and cell expansion, during which the berry is formed and the seed embryos are produced (Coombe 1992; Conde et al., 2007).

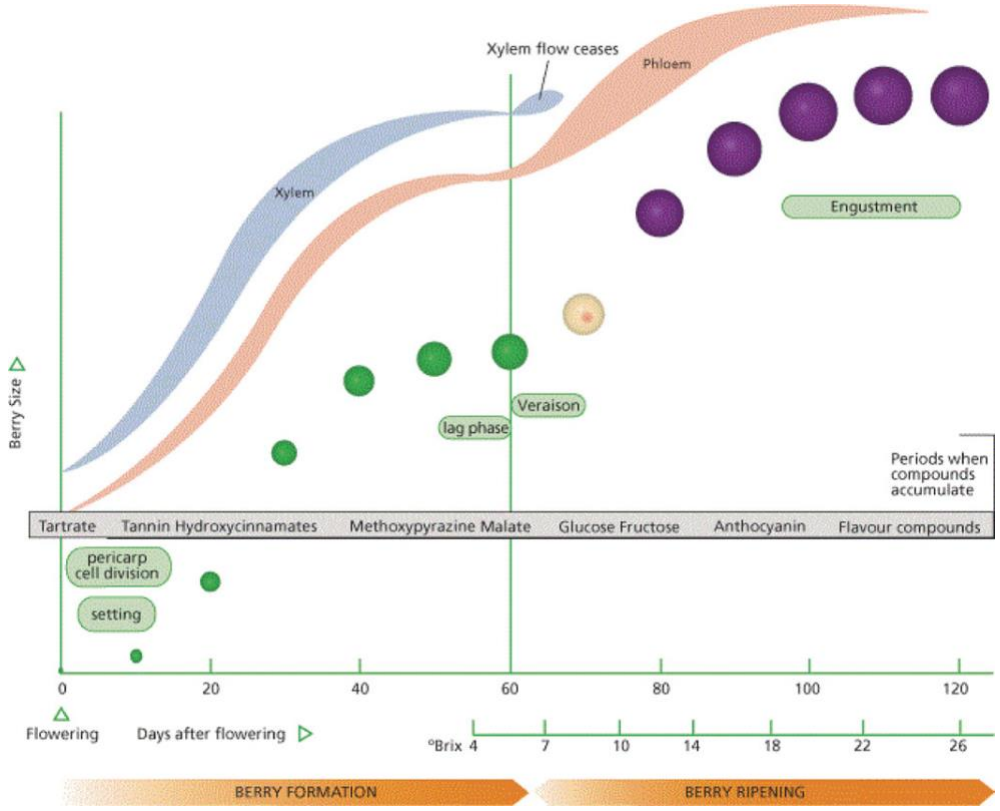


Figure 2. Berry's developmental stages (González-Barreiro et al., 2015)

This phase is followed by a lag phase where no growth occurs and its duration is cultivar-specific. A second growth phase takes place, which coincides with the onset of ripening, and it is known with the French word *véraison*. During this stage many other changes take place: berries switch from small, hard and acidic, with little sugar to a status where they are larger, softer, sweeter, less acidic, strongly flavored and change skin color (Conde et al., 2007; González-Barreiro et al., 2015).

This ripening phase begins usually in August in the north hemisphere and duration depends on environmental conditions and cultivar features. Previous studies have demonstrated that all these biochemical changes, from veraison to harvest, respond to endogenous hormonal control (**Fig. 3**) like ethylene, abscisic acid and brassinosteroids (Davies et al., 1997; Tesnière et al., 2004; Chervin 2008; Symons et al., 2006) and they are strongly influenced by environment, but the genetic control of these processes are for the most part still unknown.

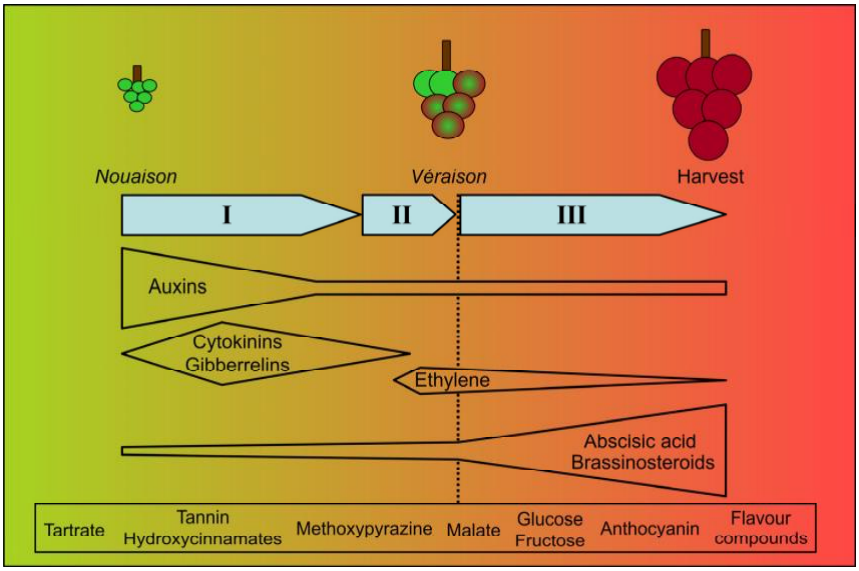


Figure 3. Hormonal control of berry's developmental stages (Conde et al., 2007)

The timing when veraison occurs is highly relevant especially for berry quality, that is sensitive to climate condition along maturation (Mori et al. 2007; Mira de Orduna 2010), in particular to high temperature.

Veraison is usually evaluated by checking the berries for the beginning of softening or the start of anthocyanin pigmentation. However, the latter process results more difficult for white-berried cultivars, that instead of depositing anthocyanins in the skin, lose chlorophyll and start to brighten.

Veraison time is strictly correlated to the other phenological stages and is considered a predictor of maturity and harvest time (Sadras and Petrie, 2011) but it may differ significantly between cultivars.

The complete understanding of the genetic mechanisms that regulate this important process is not yet fully understood. However, previous studies have identified and confirmed several QTLs that explain a portion of the phenotypic variance. For example, several studies have confirmed the presence of loci on chromosome 16 that appear to be involved in the genetic regulation of veraison (Costantini et al., 2008; Duchêne et al., 2012; Zyprian et al., 2016) (see the introduction of Chapter 1 for a more extensive review).

Grapevine cultivation in a changing environment

High temperatures during ripening lead to negative effects on berry quality. Such detrimental effects, include a change of final anthocyanin content in berries (Mori et al., 2007) with alterations in color of fruit and red wines, desynchronization of sugar and organic acid metabolism in ripening fruits, remodeling of berry transcriptome (Rienth et al., 2016; Sweetman et al., 2014), alteration in the content of polyphenolic compounds and incomplete development of flavors.

Beside the negative effect of high temperatures along ripening, extreme environmental events like floods and spring late freezing also negatively affect the cultivation of grapevine. Previous studies have shown that high temperature during bud burst can significantly decrease the number of flowers per inflorescence and of course the number of berries per plant (Keller et al., 2010; Petrie and Clingeleffer, 2005). On the other hand, also cold, floods and freezing temperatures can negatively affect budburst and flowering: cold or rainy weather around flowering can reduce the number of grape clusters formed while frost conditions if occurring during late spring can cause significant crop loss when showing up after bud burst (Mosedale et al., 2015; Mullins et al., 1992).

As conclusion, due to climate changes and rising temperature, grapevine cultivation needs to be updated to the new environment conditions. One of these possibilities is to expand grape cultivation areas to countries with milder weather, like UK. For those varieties used to produced DOC wines like Valpolicella and Amarone, which need to be grown in specific area, adapting these sensible varieties to environment conditions is gaining more and more relevance to preserve the peculiarities of grapes.

CHAPTER 1

QTL mapping in a segregant population Cabernet Sauvignon x Corvina for the dissection of genetic determinants on phenological and qualitative traits

1. Abstract

In grapevine, as in many other crops, phenotypic traits show a complex quantitative inheritance with small additive or dominance effects of individual genes. By applying a quantitative trait loci (QTL) analysis it is possible to identify segments of genome linked to observed variation including a large number of genes.
QTLs mapping, indeed, is a statistical method that links two types of information: phenotypic data (trait measurements) and genotypic data (usually molecular markers), to explain the genetic basis of variation in complex traits.

In this study, QTL mapping was conducted in the bi-parental population Cabernet Sauvignon x Corvina (CsxC), one of the segregant populations developed in our lab, to dissect genetic determinants of quality as well as phenology in grapevine. Genetic and phenotypic data were collected since 2018 on 142 individuals composing CsxC population. Genetic data are represented by SNP profiles and a high-density linkage map developed for the purpose of QTL mapping both obtained from hybridization of populations' genomes on GrapeReSeq *Vitis*18KSNP chip, while phenotypic data were collected in 2020 and 2021 following phenological events under field conditions in productive vineyards of CsxC population. Map QTL 5.0 was finally used to combine this biological information to detect putative QTLs involved in phenological traits. A QTL map was produced and candidates underlying the QTLs and putatively involved in traits are discussed.

2. Introduction

The dissection and understanding of the genetic determinants and mechanisms that regulate the phenological phases in grapevine represent the starting point for the development of breeding programs that lead to the achievement of improved and more stable cultivars. As for abiotic stresses, one of the most desirable targets in genetic improvement is obtaining late varieties for the period of veraison. As already mentioned in *Background* paragraph (page 13), if veraison and ripening take place in very warm periods, an alteration of the biochemical processes may occur which leads to the incorrect formation of aromas and sugar concentrations and alter the quality of grapes and wine (Mori et al., 2007; Mira de Orduña Heidinger 2010; Sadras and Petrie, 2011).

Like for many other species of agricultural interest, even in grapevine many of the target traits for breeding, especially as regards phenology, show a complex and quantitative inheritance, i.e., regulated by the effect of several genes, each with greater or lesser contributions but that control a measurable trait through additivity or dominance; we therefore speak of quantitative trait loci, or QTL. A genetic approach that is commonly used for the identification of QTL regions is represented by QTL mapping, which allows the identification of regions on the genome that most likely contain regulatory genetic elements for the trait in question. Generally, the starting point for a QTL mapping analysis is the development of a bi-parental population by crossing two parents divergent for a given character, in order to obtain a large number of recombinations that allow mapping of variants segregating from the two (Mauricio et al., 2001; Mackay et al., 2009; Duchêne et al., 2009; Garris et al., 2009). Over the years, several studies have demonstrated the effectiveness of this genetic approach also for mapping of phenology related QTLs in grapevine, revealing the localization of some QTLs involved in different phenological phases.

Costantini et al. (2008) identified some QTLs associated with flowering, veraison and ripening located on chromosomes 2, 6 and 16. Later Duchêne et al. (2012)

confirmed the presence of a QTL related to the flowering-veraison interval on chromosome 16, but in a different confidence interval compared to that identified by Costantini et al., thus suggesting the presence of two QTLs in this region. Other QTLs associated with veraison have also been identified on chromosomes 2, 15 and 17 (Grzeskowiak et al., 2013) on a different population. More recently Zyprian et al. (2016) further confirmed the presence of a QTL on chromosome 16 by using two different populations, and identified an additional region involved in veraison time control on chromosome 18. The QTL located on chromosome 2 and identified by the two cited papers seems instead to co-localize with loci encoding two transcription factors, *VvMybA1* and *VvMybA2*, involved in the pigmentation of the berry thanks to the activation of pathways that lead to the biosynthesis and accumulation of anthocyanins (Guo et al., 2019; Kobayashi et al., 2004; Fournier-Level et al., 2009). This QTL could have been identified in these studies since often color is used for detection of the veraison time.

Two data sources are required for the identification of genomic regions through QTL mapping: a genetic source, represented by a linkage map built with molecular markers, and a phenotypic source, consisting of measurements of the traits to be mapped and collected over different years for greater reproducibility.

As for the genetic source, the advent of modern genotyping technologies based on NGS and SNP markers has made the construction of maps much more reliable and precise, thanks to the high number of markers available that increase the resolution of the map and therefore also QTL mapping (Troggio et al., 2007; Salmaso et al., 2008; Moreira et al., 201). In fact, compared to the past and to maps built with few SSR markers, a linkage map built with SNP markers allows for genomic ranges rich in polymorphic markers, which increase the resolution and the probability of finding an association with the trait under examination. As previously discussed in *Background* paragraph (page 9), the introduction of NGS genotyping (such as GBS) and subsequently the development of SNP chip has allowed to increase the throughput of SNP genotyping for many species obtaining thousands of markers in a short time for linkage maps development.

For the QTL mapping, consensus maps built using the crossing population (CP) strategy and coming from the integration of two separate parental maps (P1 and P2) can be used. However, by using this strategy it is possible that loci with additive contributions are neglected and they would not be detectable by the consensus map (Doligez et al., 2013). Moreover, given the bi-allelic nature of SNPs, markers segregating in both populations are usually less informative. This could potentially lead to less precise marker placements in the CP consensus map and increases the QTL mapping computation requirements due to imputation related to missing genotypes.

For all these reasons, in this work it was decided to carry out the mapping of the QTLs on single parental maps (at least for the MQM approach, see later), from which reproducible QTLs between the two maps P1 and P2 were finally extracted. The source of phenotypic data, on the other hand, is represented by the measurements of traits under investigation, carried out over two different vintages, in order to obtain datasets on which to perform statistical analysis for the association between genetic positions of markers and phenotypic data.

The MapQTL 5.0 software was used in this study for QTL mapping purposes (Van Ooijen, 2004). This software was developed specifically for the identification of QTL in diploid experimental populations, to search for the regions of genome responsible for phenotypic variance of investigated traits and to estimate the genetic effect.

MapQTL allows the analysis of QTLs with different functions: Interval Mapping (IM) provides that markers are ordered in a genetic map. The program analyzes the markers in couple in a defined interval, in our case 1 cM, trying to find an association between the genomic region of the interval and the phenotype (Lander and Botstein, 1989; Van Ooijen, 1992). However, if there are multiple QTLs close to each other on the same chromosome, the calculation with IM does not allow a high resolution of close regions, so a second type of function is available. Multiple QTL Mapping (MQM) is a more powerful and precise calculation tool (Jansen, 1993; Jansen and Stam, 1994), especially in the case of regions containing multiple interacting QTLs. This type of analysis searches for maximum LOD significance

values and uses markers as cofactors selected manually or automatically according to a user-defined interval for the research and resolution of QTL concatenated on the same chromosome, thus allowing to separate two neighboring peaks. In this study, both approaches were used to search for QTL regions most probably involved in the regulation of phenological and in quality traits in grapevine.

3. Material and Methods

3.1 Plant material: the Cabernet Sauvignon x Corvina cross population

Plant material used in this study is a population obtained by crossing the Cabernet Sauvignon and Corvina (CsxC) varieties. The cross was performed in 2016, together with further crosses from the local variety Corvina.

In detail, the pollen collected from Corvina (clone 48) was used for the artificial pollination of the emasculated Cabernet Sauvignon (clone ISV5) flowers, at the experimental vineyard of Verona province in San Floriano (VR). Obtained seeds were vernalized and germinated and 667 plants were obtained. When plants reached a size sufficient to be transferred to the open field, they were transported to Vivai Gozzo (Verona) and placed in an experimental vineyard on their own foot and maintained under standard viticulture conditions to produce plant material for subsequent propagation.

Among all 667 plant 142 confirmed as progeny of CsxC cross by SSR genotyping were selected for QTL mapping purposes. These were then propagated for the creation of productive vineyards replicated in multiple copies and in different environments to collect phenotypic data for the QTL analysis.
Individuals, including the parents Cabernet Sauvignon and Corvina, were propagated via cuttings grafted onto SO4 rootstocks and placed in 2019 in productive vineyards. To test the contribution of environment on phenology, vineyards were positioned in two different environments: two copies (left and right CsxC) were positioned at the Vivai Gozzo (Verona) while a third copy was housed in the Valpolicella land belonging to the winery company Guerrieri-Rizzardi (Negrar, Verona) as traditional cultivation area of Corvina, according to the DOC denomination.
The scheme of productive vineyards and locations of the three copies are shown respectively in **figures 4a** and **4b-c.**

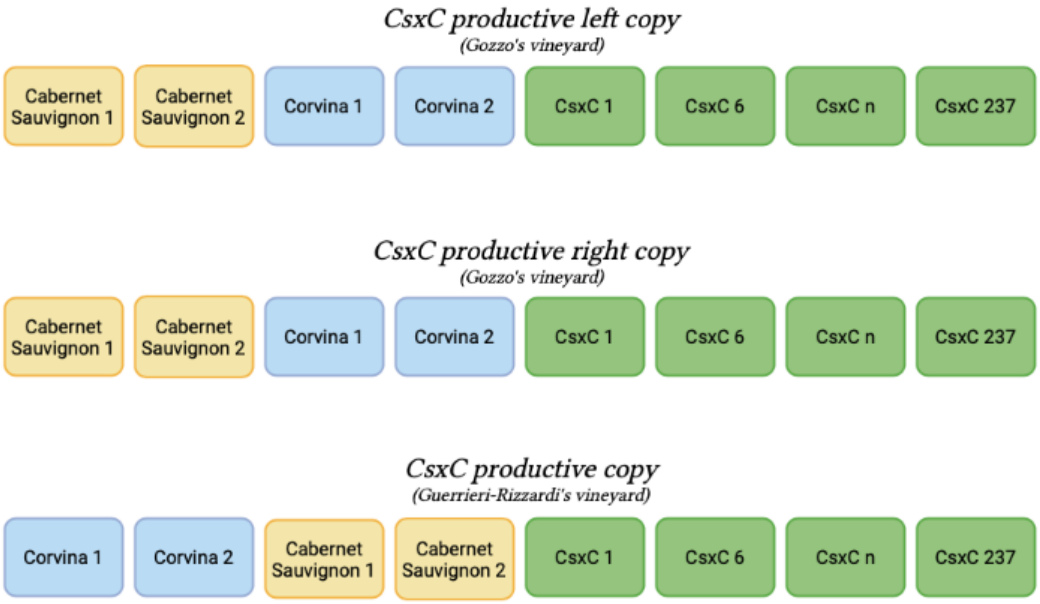


Figure 4a - Schematic representation of field distribution of CsxC productive population copies. Cross individuals are numbered sequentially and are placed in double copy (left and right, 1 plant per genotype in each copy) at Gozzo's vineyard, while only one copy is placed in Guerrieri-Rizzrdi's vineyard.

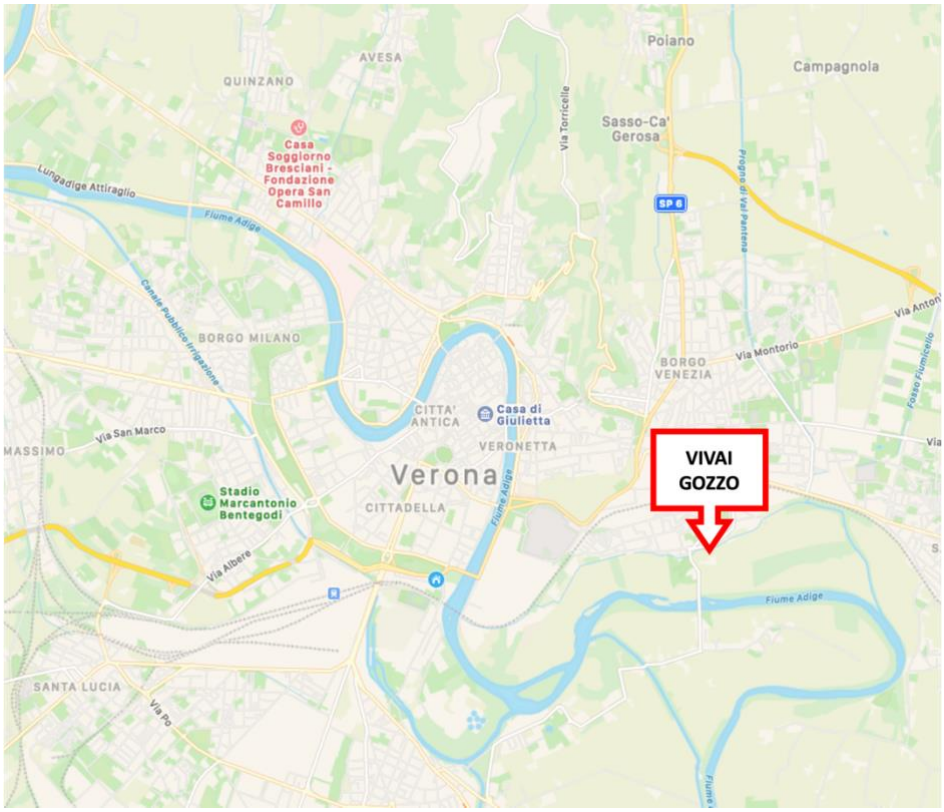


Figure 4b - Geo-localization of Gozzo's productive vineyard

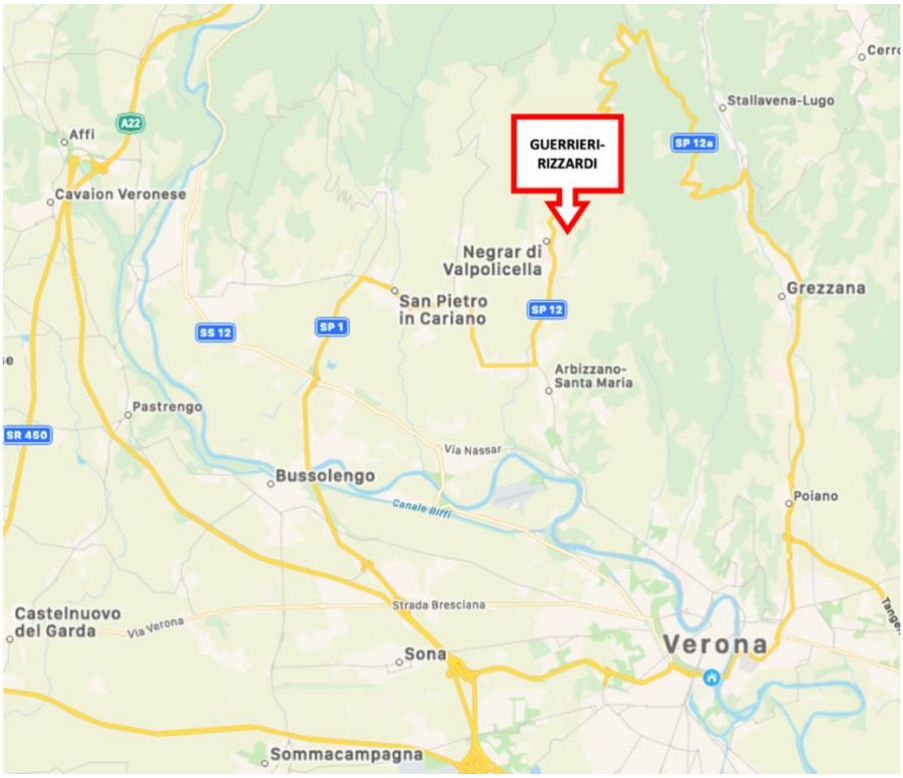


Figure 4c - Geo-localization of Guerrieri-Rizzardi's productive vineyard

3.2 Genomic DNA extraction

For DNA extraction purpose, 2-3 young leaves were taken from each plant, and stored individually frozen in plastic bags. Subsequently, the genomic DNA was extracted using the DNeasy plant mini kit (Qiagen, Germany) according to following protocol:

1) Sterile 1.5 mL Eppendorf tubes are prepared with 1.7 mm glass beads (Roth, Germany) for mechanical grinding of the tissue, marking the cap with the name of the genotype.
2) For each genotype, 300 mg of leaf tissue are sampled from the still frozen sample and placed in the tube with glass beads. Tubes are then immediately placed in liquid nitrogen and processed with mechanical grinding of the tissues through the Silamat S5 grinder, taking care not to thaw the sample. Tubes are then removed from liquid nitrogen with pliers and quickly placed in the grinder. Each sample is grinded for about 6 seconds and then frozen

again in nitrogen. The operation is repeated 3 times until a powder is obtained.

3) In the meantime, prepare an equal number of 1.5 ml labeled Eppendorf tubes and aliquot 500 µL of the buffer AP1 pre-heated to 65° C in a water bath. Tubes with buffer are then placed in the water bath to maintain the temperature
4) After grinding, all the powder is poured into the tubes containing the lysis buffer and careful mixing is carried out, so that all the powder meets the buffer.
5) Aliquot 5 µL of RNase A in each sample and mix thoroughly.
6) The samples are incubated in water bath at 65 ° C for 10 minutes, inverting the tubes every 3 or 4 minutes.
7) After incubation, 163 µL of buffer P3 are aliquoted in each sample; the solution is mixed and incubated on ice for 5 minutes.
8) The lysate is centrifuged for 5 minutes at 14.000 rpm.
9) The supernatant is then carefully removed without touching the pellet and placed in the previously marked QIAshredder column. The column is centrifuged for 2 minutes at 14.000 rpm.
10) The eluate is then transferred to a new marked 1.5 mL Eppendorf tube without touching the pellet at the bottom of the column (if present).
11) 1.5 volumes of buffer AW1 are aliquot to the solution and mixed by pipetting.
12) 650 µL of the obtained solution are transferred into the DNeasy Mini spin column and centrifuged for 1 minute at a speed of at least 8.000 rpm.
13) The eluate is discarded and the step is repeated with the remaining solution always using the same column.
14) The column is then moved to a new 2 mL elution tube and 500 µL of the AW2 buffer is added. Column is centrifuged for 1 minute at a speed of at least 8.000 rpm.
15) Another 500 µL of AW2 buffer are added and column is centrifuged for 2 minutes at 14.000 rpm.

16) Transfer the column to a new sterile 1.5 ml labeled Eppendorf tube and aliquot 100 µL of sterile MilliQ water exactly on the column filter, but without touching it, to allow the elution of the DNA from the silica matrix.
17) Tubes are then incubated for 5 minutes at room temperature and centrifuged for 1 minute at a speed of at least 8.000 rpm
18) Repeat step 17. Keep the extracted DNA on ice.

Each DNA extracted was then quantified with NanoDrop One/OneC spectrophotometer (Thermo Fisher, USA) and analyzed by loading in 0,8% agarose gel.

3.3 SSR genotyping

3.3.1 SSR loci amplification

To confirm the successful outcome of the crossing, two informative SSR microsatellite markers were chosen according to profile parents.
We selected the VVMD25 and VVMD28 markers which were amplified and analyzed through capillary electrophoresis with Sanger sequencer. These data as well as the data for marker VVMD5 were also used for the genetic map building. Primer sequences are shown in **table 1.**

SSR	Primer Forward (5'-3')	Primer Reverse (5'-3')	Annealing Temperature	References
VVMD25	NED-TTCCGTTAAAGCAAAAGAAAAAGG	TTGGATTTGAAATTTATTGAGGGG	56°C	Bowers et al., 1999
VVMD28	6FAM-AACAATTCAATGAAAAGAGAGAGAGAGA	TCATCAATTTCGTATCTCTATTTGCTG	56°C	Bowers et al., 1999

Table 1 – Primer sequences and annealing temperatures of SSR markers used for genotyping

The first step for the SSR analysis is the amplification using primers marked with fluorophores which will subsequently allow the exact determination of the allelic length with Sanger sequencer reading.

PCR reactions were prepared in final 25 µL, as reported in **table 2**, and amplified with a SimplyAmp thermal cycler (Applied Biosystems, USA) by running the program reported in **table 3**.

	VVMD25	**VVMD28**
Reagents	Final concentration	Final concentration
Buffer	1x	1x
$MgCl_2$	1.5 mM	1.5 mM
dNPTs	200 uM	200 uM
Primer F	0.4 uM	0.4 uM
Primer R	0.4 uM	0.4 uM
GoTaq Flexi	0.5 U	0.5 U
genomic DNA	1ng/uL	1ng/uL

Table 2- Master mix composition for SSR markers amplification

Initial Denaturation		94°C	2 minutes
35X	Denaturation	94°C	45 seconds
	Annealing	56°C	45 seconds
	Extension	72°C	60 seconds
Final Extension		72°C	20 minutes
		4°C	∞

Table 3 - Thermocycler program for SSR markers amplification.

Once amplification reaction was completed, amplicons were verified on 1.5% agarose gel by loading, for each sample, a mix composed of 7 µL of amplicon, 3

µL of sterile water and 2 µL of Loading Buffer 6x. The gel was run at 80 V for approximately 30 minutes.

3.3.2 Determination of SSR sizes in seedlings

Amplicons were prepared for analysis with the Sanger sequencer by capillary electrophoresis.

The first step in sample preparation was the removal of all reagents and salts left over from the PCR reaction. For this purpose, Ampure XP magnetic beads (Beckman Coulter, USA) were used, which electrostatically capture nucleotide fragments with a length greater than 100 bp and allow the elimination of all superfluous components that would disturb the capillary electrophoresis reaction.

For each sample, the purification protocol was performed as follows:

1. Vortex and resuspend the Ampure XP beads
2. Add 1.8 uL of beads every 1 uL of each PCR product (45ul of Ampure for 25ul) in a 96-well plate.
3. Mix by pipetting slowly 10 times and incubate at room temperature for 10 minutes.
4. Move the plate to the appropriate support with magnets and incubate for 2 minutes. The solution must become clear.
5. While holding the plate on magnets, discard the supernatant without disturbing the beads and leave 5 uL of solution to keep the beads from drying out.
6. Still holding the plate on magnets, add 200 uL of fresh 70% ethanol solution and incubate for 30 seconds. Aspirate all the ethanol and discard it.
7. Repeat the 70% ethanol wash and allow the ethanol residues to evaporate without drying the beads too much.
8. Remove plate from magnets and elute with 36 uL of sterile MilliQ water. Incubate at room temperature for 5 minutes.
9. After that, pipette gently to resuspend the beads and incubate for 10-15 minutes, pipetting a couple more times during this interval.

10. Place the plate back on the magnets and incubate for 2 minutes to separate the beads from the DNA.
11. Keeping the plate on the magnets, transfer the eluate into new tubes, taking care not to carry any bead residues.

The purified PCR products were checked on 1.5% agarose gel to ensure no material was lost during purification.
The products were prepared for capillary electrophoresis with the Sanger sequencer.

For each product, a denaturation mix was prepared by adding 9.5 uL of HiDi formamide (Applied Biosystems, USA) and 0.05uL of GeneScan 400HD Rox Size Standard (Applied Biosystems, USA).
Just prior to load the plate onto the sequencer, a denaturation cycle was performed at 94°C for 3 minutes.
For the SSR analysis, it was used the Abi 3130XL Genetic Analyzer (Applied Biosystems, USA) located at Pathological Anatomy Laboratory of Prof. Aldo Scarpa (University of Verona).
Output files were then processed with Peak Scanner™ Software v1.0 (Thermo Fisher, USA) to extract peaks sizes.

3.4 High-throughput SNP genotyping by *Vitis*18KSNP chip

In 2013 the GrapeReSeq Consortium (Le Paslier et al., 2013) produced a 18K SNP array, the *Vitis*18KSNP chip (Illumina, San Diego), containing 13 561 SNPs identified in 47 *Vitis vinifera* cultivars and 4 510 SNPs identified in 13 genotypes from five *Vitis* species and five *Muscadinia rotundifolia* genotypes, for a total of 18071 SNPs. Nowadays, this array is used as high throughput method to assay genetic variability among cultivars

In this work, the *Vitis*18KSNP chip was used for hybridizing and genotyping each individual of the mapping population and obtaining a large number of SNP markers for the creation of a high-density genetic map.

Hybridization on the chip requires 200 ng of good quality DNA at a concentration of about 50 ng/uL for a successful genotyping. Each of the extracted DNA was therefore checked for quality, eventually repeating the extraction if necessary. Then samples with lower concentration were subjected to a concentration process by placing the samples with open tube under a biological hood for 12-16 hours, to allow the water to evaporate and increase the concentration.

DNAs samples were then sent to Sequencing Platform Laboratory at Edmund Mach Foundation (FEM, San Michele all'Adige, TN) for the hybridization on SNP chip, that was performed according to standard Illumina's protocol. Raw intensity data were visualized and clustered in GenomeStudio v2011.1 (Illumina Inc, San Diego, CA, USA) providing SNP (manifest file from Illumina) and sample information (sample sheet file). SNP were then called using a GenCall threshold of 0.20.

Subsequently, SNPs were re-filtered using ASSiST v1.01 (Di Guardo et al., 2015), which allows for the recovering of SNPs with poor clustering (such as SNPs containing null alleles or additional polymorphisms appearing as multi-allelic clustering) based on GenomeStudio outputs and a file providing pedigree information. The following thresholds were applied to filter the data: SNPs with more than 5% missing data, individuals with a call rate differing more than 10% from the population mean, individuals displaying more than 0.3% unexpected genotypes and SNP markers displaying more than 5% unexpected genotypes or segregation distortion at $P< 0.001$. ASSiST was also used to recode the genotype dataset into JoinMap® input format based on the segregation of the markers in the population (*ef* x *eg*, *hk* x *hk*, *lm* x *ll* or *nn* x *np*).

3.5 Population maps building

Genotype data of both SNP and the 2 SSR markers VVMD25 and VVMD28 were used to construct population maps in JoinMap® v5 (Van Ooijen, 2019). Marker

data were again filtered in JoinMap® to exclude individuals or markers with more than 5% missing data (with exception of SSR markers) and markers that showed severe segregation distortion ($\chi 2 \geq 15$). Loci were also tested for similarity in JoinMap®.

Within each population, loci that showed 100% similarity (matching genotypes across all individuals that would map to the same position) were also excluded and represented by a single representative marker. A previous effort to build a highly reliable integrated genetic marker order based on *Vitis*18KSNP chip and SSRs data developed from this population and two additional populations (Vervalle et al., in press) was available. We took advantage of such information for our genetic map building process.

Therefore, markers previously identified as a poor fit either due to creating large distances on the map, or to poor map quality criteria (high –log10P, high NN Fit or high NN Stress values) were removed from the dataset. Remaining markers were grouped into linkage groups using a LOD score of 4.0 or higher. They were ordered using the Maximum Likelihood (ML) mapping algorithm, which uses the Haldane mapping function (Haldane, 1919). The ML parameters were adapted to allow the algorithm to run longer until convergence and were set as follows: chain length = 3000 for S1, S2 and S3, 10 000 for S4, 15 000 for S5 and 30 000 for T0; stop criterion = 5000 for S1, S2 and S3, 20 000 for S4, 25000 for S5 and 50 000 for T0. The highly reliable integrated genetic marker order of representative markers previously established (Vervalle et al., in press) was provided as fixed order. The quality of the obtained map was then inspected re-checking the –log10P, Nearest Neighbour (NN) Fit and NN Stress values.

The consensus cross population as well as parental maps data (P1 for Corvina map and P2 for Cabernet Sauvignon map) were then extracted and used to get statistic data and draw maps with the software MapChart (Voorrips, 2002). These maps were then used for QTL mapping.

3.6 Plant phenotyping

Phenotypic data for QTL mapping were collected in 2020 and 2021 vintages on productive populations located in the two experimental vineyards of Gozzo and Guerrieri-Rizzardi (see also *paragraph 3.1* for further details). In detail, the first population present at Vivai Gozzo was phenotyped in 2020 and 2021. The copy in Guerrieri-Rizzardi's vineyard was phenotyped only in 2021 vintage.
Phenotypic data were collected for the tree main phenological stages (budbreak, flowering and veraison) and for qualitative parameters of grapes at maturity (brix, color, polyphenol and anthocyanin content).

3.6.1 Phenological traits: budbreak time

Budbreak is the first phenological phase that occurs during the vegetative development of the grapevine (*Background* paragraph, page 10). The measurement of this phenological stage involved observing the opening of the buds (twice a week) and measuring the stages of development through a scale of values and the annotation of the corresponding days.
For each plant, the four terminal buds (from the apex) of the first shoot that branches off from the main trunk were considered. If one or more buds are missing or damaged, the ones immediately following have been considered.
To obtain the values, the scale of Eichhorn-Lorenz system modified by Coombe (1995) was used. The scale includes different stages of development of the winter bud, each corresponding to a number:

- Winter (or dormant) bud = stage 1 (**Fig. 5a**)
- Swollen bud = stage 2 (**Fig. 5b**)
- Cotton bud = stage 3 (**Fig. 5c**)
- Green tip = stage 4 (**Fig. 5d**)
- Bud opening = stage 5 (**Fig. 5e**)
- First leaf not fully extended = stage 6
- First extended leaf = stage 7

- Two to three spread leaves = stage 8
- Four spread leaves = stage 9
- Five spread leaves, visible inflorescences = stage 10
- Six spread leaves = stage 11
- Seven spread leaves = stage 12
- Eight spread leaves = stage 13
- Ten spread leaves = stage 14 (**Fig. 5f**)

Budbreak time was defined when at least two of the four buds reached stage 5.

5a *5b* *5c*

5d *5e* *5f*

Figure 5a-f. a) Winter bud. b) Swollen bud. c) Cotton bud. d) Green tip. e) Bud opening. f) Ten spread leaves.

3.6.2 Phenological traits: flowering time

As for budbreak, the modified Eichhorn-Lorenz scale was used to determine the flowering time. Similarly, the same first four buds were observed (twice up to three days/week), recording, according to the scale, the flower stage and the corresponding day:

- 0%: all flowers are closed and the so-called "floral buttons" are observed (**Fig. 6a**)
- 25%: opening of the first flowers is observed, with the detachment of calyptra from the floral receptacle (**Fig. 6b**)
- 50%: late flowering stage, most of the flowers are open (**Fig. 6c**)
- 75%: full bloom: all flowers are open and the senescence of the anthers is observed
- 100%: end of flowering after pollination (**Fig. 6d**)
- Pea phase: all the flowers have detached, small berries are visible as “pea” (**Fig. 6f**)

6a

6b

6c

6d

6e

6f

Figure 6a-f. a) Floral buttons. b) Opening of first flowers. c) Late flowering. d) Full bloom. e) End of flowering. f) Pea phase.

As for Budbreak, a plant is defined as flowered when at least two of the four buds reach the "advanced flowering" stage (50%).

3.6.3 Phenological traits: veraison time

The softening of the berry was used to determine the veraison stage. The use of tactile sensation is used internationally to assess the veraison status of both white and red berries. In fact, for white berries, the evaluation of the color change, which could be used for red berries, is not feasible. Therefore, the softening of the berries, typical of this phenological stage, is evaluated.

For every plant that presents well developed clusters, four of them were analyzed (one time to twice a week) by testing the softness of ten berries, distributed on 3 different heights of the cluster and noting a percentage score finally converted into a number:

- 0% = 0
- 25% = 2
- 50 % = 3
- 75% = 4
- 100% = 5

The veraison stage is considered completed when the average of the four clusters reaches at least stage 3 (50%).

3.6.4 Technological traits: maturity

Once veraison process is complete, ripening of the berries was followed by measuring Brix degree. After veraison, in fact, berries continue their ripening process by accumulating sugars. The quantification of the sugar content was carried out with a portable field refractometer (**fig. 7**), which exploits the refraction of light produced by sugars. By squeezing the berry juice on the special lens, it is then possible to read the Brix degree reached on the scale of the instrument.

Figure 7 – Portable refractometer used to measure Brix grade

From the moment in which veraison process is complete, ripening was monitored by collecting data once a week and testing 3 berries randomly chosen from the 4 clusters considered for each plant. When a plant reached Brix values close to 17, reliefs were increased to 2-3 per week by testing 6 berries, to accurately follow the evolution of ripeness and the date on which Brix 19 is reached, a value that we considered as mark for the maturity of the grapes and harvest.

Once each genotype reaches an average Brix degree equal to or greater than 19, the grapes are harvested, taking 4 clusters per plant (if available), labeling them and noting the harvest date.

The grapes harvested in each relief were then transported to the Laboratory of Ecophysiology of Grapevine (Prof. G.B. Tornielli – Department of Biotechnology, San Floriano, VR) for subsequent technological analyzes.

Furthermore, the derived maturity date was also calculated using an interpolation line constructed with the Brix values collected during the reliefs, to obtain the exact ideal date in which the value 19 was reached by each plant.

3.6.5 Technological traits: berry color

We also wanted to provide some very preliminary information about berry maturation, color and further quality parameters related to relevant metabolites and their segregation.

Clusters were immediately photographed and analyzed after transportation to the lab.

The color of clusters of each plant was noted, *b* (white, **fig. 8a**), *r2* (light red, **fig. 8b**) and *r3* (dark red, **fig. 8c**), and converted into numerical format based on the following correspondence:

- *b* = 0
- *r2* = 2
- *r3* = 3

a) b) c)

Figure 8a-c. Three classes of berry color converted into numerical format: a) white grape=0 b) light red grape=2. c)red grape=3.

3.6.6 Technological traits: polyphenols and anthocyanin content

From the clusters collected for each genotype, part of the berries was also taken for the subsequent extractions and quantifications of polyphenols and anthocyanins. At harvest day, 55 grams of berries were randomly collected from clusters of each plant and stored in plastic bags at -20 ° C until the moment of extractions.

The extraction protocol was carried out on two replicates for each sample and included the following steps:

- 20 grams of berries were weighed and thawed for each sample
- The material was placed in a blender in which 40ml of extraction solution (consisting of 847.7 ml H_2O, 2.3 ml of 37% hydrochloric acid and 150 ml of EtOH was added, for 1 liter of solution)
- Once the solution was added, berries were blended at low speed, so as not to mince the seeds and taking care to proceed with the same blending time for each sample.

- The mixture was transferred to a 250 ml flask, covered with parafilm and put under moderate stirring for 60 minutes at room temperature.
- At the end of the incubation, 50 ml was taken and placed in a tube and centrifuged at 4000 rpm for 15 minutes.
- After centrifugation, the supernatant was transferred into a new tube.

The spectrophotometric quantification of polyphenols and anthocyanins required also the dilution of samples obtained with a reading solution, composed of 700 ml of EtOH, 300 ml of H_2O and 10 ml of 37% hydrochloric acid (per liter of solution). The dilution performed for each sample was decided on colorimetric intensity of the sample, in order to obtain values that did not exceed those of the calibration line of the spectrophotometer. Quantification of polyphenols and anthocyanins was measured by reading the absorbance of each sample at 280 nm for polyphenols and 540 nm for anthocyanins, finally averaging values obtained from the replicates of each sample.

3.7 Statistical analysis of phenotypic data

Data obtained from phenological and technological surveys for the different population copies and vintages were recorded at the end of each season in excel matrices and processed in graphics with the distributions of the values expressed in Julian day of the year for phenological traits and with different scales or parameters for technological data. Furthermore, the Pearson correlation coefficients were calculated on the same data. During the statistical analysis performed by CORRELATION function of Excel, data collected in 2020 and 2021 between the two copies of the Gozzo's field (right and left) and their averages were compared. The population copy present at Guerrieri-Rizzardi's field allowed the collection of data only for the 2021 vintage, which was therefore compared with the average values of Gozzo 2021.

Since this type of correlation is not able to consider climatic differences between vintages or between cultivation areas, correlation coefficients of the Spearman

ranks have also been calculated, in which the absolute values of the matrices are not used. In fact, by using RANK.MEDIA function, the values are converted into ranks: the matrix is sorted in ascending or descending order and a position in the matrix is associated with each value. Finally, the correlation between the two rank matrices is calculated.

3.8 QTL mapping

The identification of QTL regions putatively involved in the genetic regulation of the traits under examination was carried out using the MapQTL 5.0 software (Van Ooijen, 2004) that combines suitably formatted phenotypic data with genotypic data to search for a statistical correlation between phenotypes and markers.

3.8.1 MapQTL 5.0: input files

The software, according to the manual, requires the preparation of three input files: a *locus genotype file* (or loc.file) which contains the genotype codes of each marker in the segregating population, a *quantitative data file* (or qua.file) which contains the data of quantitative traits of all individuals present in the map and, finally, a *map file* (map.file) that contains the genetic positions of all the genotyped loci.

The *qua.file* was formatted according to the manual, reporting all quantitative data of each trait for 142 genotype of the CsxC population. In the case of phenological traits, the relevant dates were converted to Julian days. For analysis of the population in Gozzo's vineyard were used both data of left copy, of right copy and the medium between them. For the population present in Guerrieri-Rizzardi's vineyard, data available for the only copy present were used.
On the other hand, for the technological traits was reported the average values of the analysis carried out in two replicates.
The *map.files* were represented by three genetic maps: the consensus map of crossing population (CP map) and two parental maps, P1 (Corvina) and P2 (Cabernet Sauvignon).

Similarly, the *loc.file* has been extracted from JoinMap for each map (P1, P2 and CP maps).

3.8.2 MapQTL 5.0: Interval Mapping (IM)

Once all input files were correctly formatted, a first analysis was conducted with the IM function. In IM, a QTL likelihood map is calculated every 1 centiMorgan (cM) and the presence of a segregating QTL and its genetic effect are determined. For each trait and for each map, an IM analysis was then run, providing as output a list of genomic regions most likely involved in the explanation of the phenotype in exam. To set a significance threshold for the selection of the aforementioned regions, a LOD threshold value was calculated for each trait and for each chromosome, and a permutation test was conducted (see paragraph 3.8.3).
Once the LOD significance threshold was obtained for each linkage group and for each trait, it was possible to select, among the results of IM runs, all the genomic regions that showed LOD peaks above the calculated threshold. These regions were brought together in a table as a first overview of confidence intervals in which possible QTL regions involved in the examined traits are included.

3.8.3 MapQTL 5.0: LOD score and permutation test

The IM function compares the likelihood under the alternative hypothesis (H1) to the likelihood for the situation when a locus with zero genetic effect would segregate (likelihood under the null-hypothesis H0 or no segregating QTL). This comparison uses a likelihood ratio statistic called LOD score, which is the 10-base logarithm of the quotient of the two respective likelihoods. A segregating QTL is detected when LOD score exceeds a predefined significance threshold on a linkage group; the position with the higher LOD score on a linkage group represents the estimated position of the QTL on the map. For this purpose, it is important to select the correct LOD score and the relative threshold above which a QTL is detected. The LOD score depends on genome size of the plant species, number of chromosomes pairs and their average size and number of markers. Van Ooijen (1999) proposed a formula to identify the most suitable LOD threshold for each

case. However, for very large datasets, such as the one in this study, it is advisable to identify the LOD threshold more precisely.

To this aim Map QTL apply a permutation test function, which is a statistical re-sampling test that allows to obtain the significance thresholds for each linkage group (chromosome wide) or for the entire genome (genome wide). In this test it is possible to select the number of permutations (1.000 in our case), in which the quantitative values are randomly rearranged, keeping the markers fixed.
After each permutation, an IM is performed and the maximum LOD value for each linkage group and genome is considered. Once the entire set of permutations has been performed, the distribution frequency of the maximum values of LOD is obtained under the null hypothesis, i.e., the absence of QTL.
The output is a table of absolute, relative and cumulative frequencies for each linkage group and Genome Wide
As indicated by the manual, in this study the LOD threshold with P-value of 0.05 was considered by choosing from the table obtained the LOD value in correspondence of 1-0.05, or 0.95. In our case, a chromosome wide approach was used, and therefore the thresholds of all 19 linkage groups were detected for each trait.

3.8.4 MapQTL 5.0: multiple QTL model mapping (MQM)

MQM was carried out using only the two parental maps P1 and P2 separately for the analysis of phenological and technological traits, with the same datasets and quantitative files used for IM. Indeed, the high number of markers composing the CP map and their typology (*hk* marker kind, which are less informative and require more imputation) did not allow the MQM analysis by MapQTL 5.0 software. Therefore, MQM results are provided only for parental maps.
Similarly to IM, also for MQM the *loc.file* relating to the map P1 or P2, the *qua.file* (unique for both maps) and the *map.file* P1 or P2 have been loaded into the software. For each trait and for each map, the backward MQM approach was followed, which provides for an initial automatic selection of the cofactors for each linkage group. With this function, it is possible to automatically select markers with ideally higher

LOD values as cofactors, which are then used by the program to identify further putative QTL regions.

In some cases, some QTLs may be left out even though the LOD value remains higher than others, or they may be clustered into a single peak while represent separate QTL regions. In these cases, a manual selection of the cofactors was made to improve QTL detection, by analyzing the specific region and performing different tests, testing one or more markers close to the LOD peak as putative cofactor.

4. Results

4.1 Establishing of the Cabernet Sauvignon x Corvina mapping population

Our aim was to explore the genetic control for grapevine phenology focusing on the local cultivar Corvina. In detail, we wanted to exploit one of the genetic crosses recently developed at the University of Verona starting from this variety, to map QTL controlling phenological stages and identify potential candidate genes.
The first task was therefore to establish a mapping population from the segregating progenies obtained by crossing Cabernet Sauvignon and Corvina (CsxC) to be characterized both genetically and phenotypically for QTL analysis.

We focused on the 664 crosses obtained in 2016 from emasculated Cabernet Sauvignon flowers artificially pollinated with Corvina pollen taken from an experimental vineyard in San Floriano (Verona) and that are now grown under open field at Vivai Gozzo.
In order to verify the successful outcome of the crossing, the first 167 plant progenies were subjected to genotyping by SSR molecular markers, to exclude self-cross individuals. Confirmed cross-pollinated genotypes were therefore chosen for the production of a map population on which to carry out high throughput genotyping by SNP chip and phenotyping.

For the kinship analysis VVMD25 and VVMD28 SSR markers were amplified and amplicons analyzed through capillary electrophoresis with Sanger sequencer.
Allelic sizes obtained made it possible to follow the inheritance of the loci and to identify genotypes that have inherited one allele from the father (C) and one from the mother (Cs). An example of these SSR genotyping results is shown in **table 4a**, while the complete list of 167 genotypes analyzed is shown in *Appendix* (table 26, page 147).

In **table 4b** are reported the final 142 CsxC individuals used in this study as confirmed cross-progeny and selected for clonal propagation, genotypic and phenotypic characterizations.

	VVMD25		VVMD28		
Sample	A1 (bp)	A2 (bp)	A1 (bp)	A2 (bp)	
Corvina	241	263	258	258	
Cabernet Sauvignon	239	249	234	236	
CsxC 1	239	241	234	258	*Cross*
CsxC 5	239	241	236	258	*Cross*
CsxC 6	239	263	234	258	*Cross*
CsxC 7	249	263	234	258	*Cross*
CsxC 8	241	249	236	258	*Cross*
CsxC 9	241	249	234	258	*Cross*
CsxC 10	239	263	236	258	*Cross*
CsxC 11	241	249	234	258	*Cross*
CsxC 12	239	263	234	258	*Cross*
CsxC 13	239	241	234	258	*Cross*
CsxC 14	239	241	234	258	*Cross*
CsxC 18	239	263	236	258	*Cross*
CsxC 19	239	263	236	258	*Cross*
CsxC 20	239	239	236	236	*Self*
CsxC 21	241	249	234	258	*Cross*
CsxC 22	249	263	236	258	*Cross*
CsxC 31	239	263	236	258	*Cross*
CsxC 32	239	241	234	258	*Cross*

Table 4a - Example of results from SSR genotyping on first 167 individuals out of 664 seedlings obtained from CsxC cross. In yellow, a case of self-pollination.

List of 142 CsxC genotypes used for QTL mapping					
CsxC	1	CsxC	75	CsxC	140
CsxC	6	CsxC	76	CsxC	141
CsxC	7	CsxC	77	CsxC	142
CsxC	8	CsxC	80	CsxC	144
CsxC	9	CsxC	81	CsxC	145
CsxC	10	CsxC	82	CsxC	146
CsxC	11	CsxC	83	CsxC	147
CsxC	12	CsxC	84	CsxC	149
CsxC	13	CsxC	85	CsxC	150
CsxC	14	CsxC	86	CsxC	156
CsxC	18	CsxC	87	CsxC	159
CsxC	19	CsxC	88	CsxC	163
CsxC	21	CsxC	89	CsxC	165
CsxC	22	CsxC	90	CsxC	169
CsxC	31	CsxC	91	CsxC	171
CsxC	32	CsxC	92	CsxC	172
CsxC	33	CsxC	93	CsxC	173
CsxC	34	CsxC	95	CsxC	177
CsxC	35	CsxC	96	CsxC	179
CsxC	36	CsxC	99	CsxC	181
CsxC	37	CsxC	100	CsxC	184
CsxC	38	CsxC	101	CsxC	192
CsxC	39	CsxC	102	CsxC	194
CsxC	40	CsxC	104	CsxC	197
CsxC	41	CsxC	105	CsxC	198
CsxC	42	CsxC	106	CsxC	199
CsxC	43	CsxC	107	CsxC	200
CsxC	44	CsxC	108	CsxC	201
CsxC	46	CsxC	111	CsxC	202
CsxC	47	CsxC	112	CsxC	203
CsxC	49	CsxC	113	CsxC	205
CsxC	50	CsxC	117	CsxC	206
CsxC	51	CsxC	118	CsxC	207
CsxC	52	CsxC	120	CsxC	208
CsxC	53	CsxC	121	CsxC	210
CsxC	54	CsxC	122	CsxC	211
CsxC	55	CsxC	123	CsxC	212
CsxC	56	CsxC	124	CsxC	213
CsxC	57	CsxC	125	CsxC	214
CsxC	58	CsxC	126	CsxC	228
CsxC	59	CsxC	128	CsxC	229A
CsxC	61	CsxC	129	CsxC	230
CsxC	62	CsxC	131	CsxC	232
CsxC	63	CsxC	134	CsxC	234
CsxC	66	CsxC	137	CsxC	235
CsxC	68	CsxC	138	CsxC	236
CsxC	72	CsxC	139	CsxC	237
CsxC	74				

Table 4b - List of confirmed cross individuals of breeding Cabernet Sauvignon x Corvina by SSR genotyping and propagated in 3 copies of productive vineyards.

4.2 CsxC high throughput genotyping by *Vitis*18KSNP chip

Selected 142 individuals of the CsxC population were used for high throughput genetic characterization by hybridization on GrapeReSeq *Vitis*18KSNP chip.
The genomic DNA of each of the 142 individuals was extracted from the plants grown in Gozzo's field. The extraction was carried out with DNeasy plant mini kit protocol, as described in *paragraph 3.2* (pages 23-25). In **table 5** are reported data on extracted DNAs before and after the drying procedure aimed to increase DNA concentration.

CONCENTRATION AND QUALITY OF EXTRACTIONS				CONCENTRATION AND QUALITY AFTER DRYING		
Sample	***ng/uL***	***260/280***	***260/230***	***ng/uL***	***260/280***	***260/230***
Cs	33.8	1.51	1.27	45.2	1.77	1.81
C	39.7	1.88	1.59	63.1	1.78	1.28
184	33.7	1.94	1.5	53.8	1.84	1.47
229A	52.4	1.86	1.37	52.2	1.82	1.27
228	36.4	1.88	1.49	54.3	1.83	1.46
237	36.2	1.9	1.72	57.3	1.86	1.59
173	22.8	2.16	1.98	53.2	1.89	1.48
232	24.6	1.91	2	59.9	1.82	1.5
181	31.5	1.91	1.45	60.4	1.82	1.31
236	38.8	1.87	1.4	54.8	1.87	1.59
235	38.1	1.84	1.43	55.5	1.82	1.24
197	35.3	1.89	1.61	57.8	1.85	1.4
172	40.8	1.9	1.7	54.2	1.81	1.72
214	33.6	1.84	1.26	44.6	1.84	1.41

Table 5 - List of some DNA extraction of 142 CsxC genotypes used for SNP chip hybridization. On the left are reported the name of sample, concentration (ng/uL) and absorbance 260/280-260/230. On the right are reported values after concentration procedure.

Finally, concentrated samples for 142 progenies and the two parents were sent to Sequencing Platform Laboratory at FEM (San Michele all'Adige, TN) for the hybridization on SNP chip according to standard Illumina's protocol. One sample, Cabernet Sauvignon, was hybridized twice as control.

After hybridization on *Vitis*18KSNP chip, raw data images were manipulated with Genome Studio (Illumina) and Assist (Di Guardo et al., 2015) software to extract and filtering polymorphic SNP data.

Out of 18.071 markers tested (**table 6**), 10610 were discarded as monomorphic, failed, null or with distorted segregation while 7461 markers were approved and used for subsequent construction of dataset. Interestingly, re-clustering using Assist v1.01 (Di Guardo et al., 2015) which allows for the recovering of SNPs with poor clustering, such as SNPs containing null alleles or further polymorphisms (multi-allelic clustering), successfully recovered 227 additional markers.

Approved	**7461**
Robust	7234
AB_2_sub-clusters	27
Null_2_cluster	172
Null_4_cluster	28
Discarded	**10610**
Monomorphic	8594
Failed	1586
Distorted	424
Null allele-failed	6
Total	**18071**

Table 6 - SNPs scoring summary

4.3 Development of the genetic map CsxC

The 7461 SNP markers genotyped by *Vitis*18KSNP chip on the 142 progenies were combined together with available SSR genotyping data to prepare an input dataset to build a genetic map. Assist software already converted SNP data to JoinMap compatible format. Concerning SSRs, the data for VVMD25 and VVMD28 markers were available on all progenies (*paragraph 4.1*, page 41). Furthermore, some of the progenies were also genotyped with the SSR marker VVMD5 and all

SSRs data were converted according to JoinMap input requirements and included in the dataset. In general, segregation was evenly spread between maternal (lmxll) and paternal (nnxnp) segregation types (not shown).

Loci were tested for similarity in JoinMap and those showing 100% similarity (matching genotypes across all individuals that would map to the same position) were represented by a single representative marker for linkage analysis and integration. High-density linkage map of crossing population CsxC (CP map) was built (**fig. 9**), constituted of 3467 final SNP markers, mapped at 1 795 unique positions distributed among 19 linkage groups. The map had a length of 1284.30 cM (**table 7**) with an average inter-locus distance of 0.74 cM. Gaps larger than 10 cM were present in just two chromosomes: chr9, with a gap of 16.54 cM and chr19, with a gap of 10.99 cM.

Genome coverage was estimated by mapping the first and last marker on each chromosome to the PN40024 12X.v2 assembly and determining the percentage of the physical sequence covered by the mapped markers (Tello et al., 2019); CsxC map covered 98.26% of the reference genome assembly.

Chromosome	Number of markers	Length (cM)	Biggest gap	Average gap
1	178	72.05	3.014	0.405
2	161	50.97	5.287	0.317
3	135	56.98	3.664	0.422
4	199	76.17	2.159	0.383
5	213	65.58	3.032	0.308
6	143	66.11	9.255	0.462
7	312	87.88	2.539	0.282
8	172	70.78	3.456	0.412
9	85	52.85	16.543	0.622
10	226	72.80	2.595	0.322
11	160	59.97	4.069	0.375
12	180	59.08	2.942	0.328
13	184	73.32	5.235	0.398
14	203	75.29	2.922	0.371
15	178	78.12	2.661	0.439
16	177	53.39	3.329	0.302
17	162	58.54	2.540	0.361
18	253	91.87	4.337	0.363
19	146	62.55	10.989	0.428
Total	**3467**	**1284.30**		
Average per chromosome	*182.47*	*67.59*	*4.77*	*0.38*

Table 7 – Summary table of linkage map of CP population.

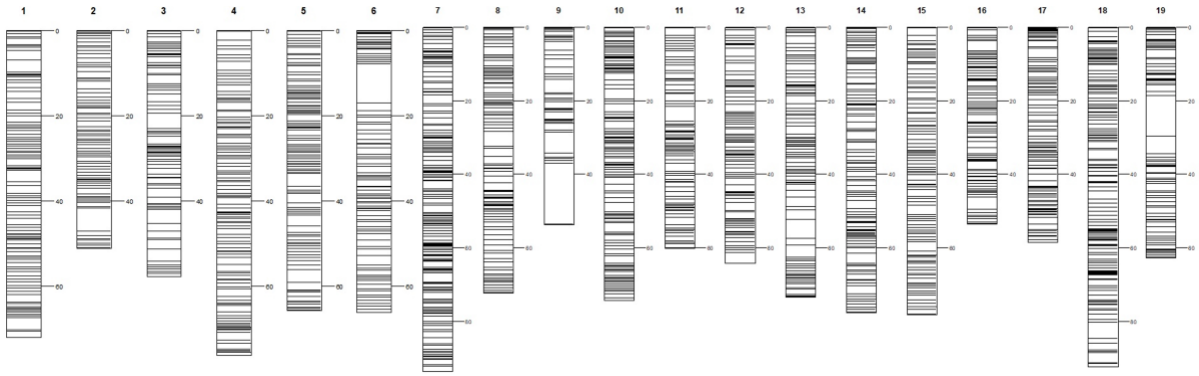


Figure 9 – Graphical representation of 19 linkage groups composing CP (crossing population) linkage map.

In addition, also parental linkage maps were assembled. Corvina linkage map (P1 map) constituted of 2440 final SNP markers distributed among 19 linkage groups (**fig. 10**) for a total length of 1274.459 cM (**table 8**). Three biggest gaps are present: on chr6 (11.866 cM), on chr13 (16.543 cM) and chr18 (13.684).

Chromosome	Number of markers	Length (cM)	Biggest gap	Average gap
1	137	87.418	5.190	0.643
2	113	48.203	2.900	0.430
3	106	58.462	7.590	0.557
4	135	62.083	2.900	0.463
5	149	68.101	4.415	0.460
6	79	66.036	11.866	0.847
7	224	85.726	3.652	0.384
8	128	66.109	5.977	0.521
9	54	35.622	6.777	0.672
10	165	65.611	3.653	0.400
11	113	57.902	2.900	0.517
12	126	60.722	3.651	0.486
13	118	75.623	16.543	0.646
14	150	89.854	5.977	0.603
15	123	88.804	4.415	0.728
16	126	50.437	3.651	0.403
17	118	58.550	2.899	0.500
18	165	88.064	13.684	0.537
19	111	61.132	6.777	0.556
Total	**2440**	**1274.459**		
Average per chromosome	*128.42*	*67.08*	*6.07*	*0.54*

Table 8 – Summary table of linkage map P1 of parental cultivar Corvina.

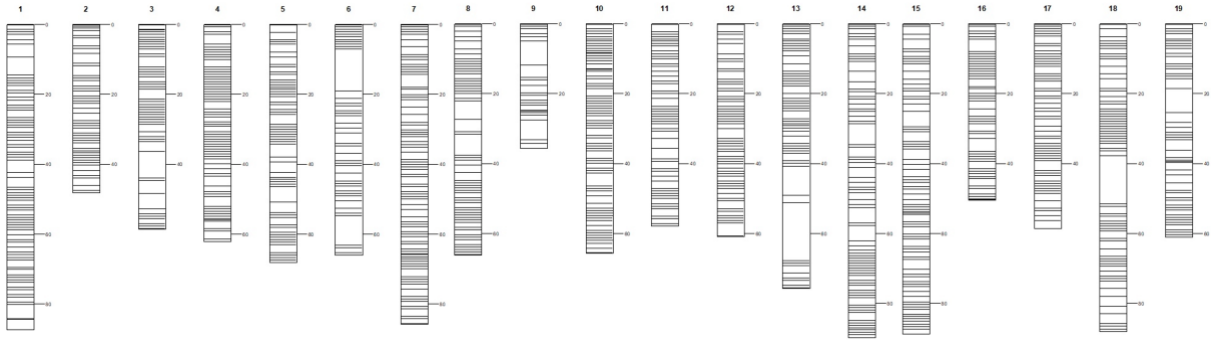


Figure 10 – Graphical representation of 19 linkage groups composing P1 (Corvina) linkage map.

Cabernet Sauvignon linkage map (P2 map) was constituted of 2379 final SNP markers distributed among 19 linkage groups (**fig. 11**) for a total length of 1271.688 cM, (**table 9**). The biggest gap (21.703 cM) is present on chr19, chr 11 (10.112 cM), chr9 (16.543 cM) and chr 3 (10.982 cM).

Chromosome	Number of markers	Length (cM)	Biggest gap	Average gap
1	98	56.608	7.591	0.584
2	121	53.710	8.417	0.448
3	79	56.937	10.982	0.730
4	117	74.617	5.189	0.643
5	156	63.057	5.190	0.407
6	107	62.633	8.417	0.473
7	235	99.443	3.652	0.425
8	109	75.445	8.417	0.699
9	85	55.010	16.543	0.655
10	167	82.254	7.592	0.496
11	105	62.035	10.112	0.596
12	114	64.636	4.414	0.572
13	116	71.001	8.417	0.617
14	132	56.549	4.415	0.432
15	113	65.306	5.190	0.583
16	127	54.201	5.977	0.430
17	128	57.821	3.651	0.455
18	177	96.457	6.778	0.548
19	93	63.968	21.703	0.695
Total	**2379**	**1271.688**		
Average per chromosome	**125.211**	**66.931**	**8.034**	**0.552**

Table 9 – Summary table of linkage map P2 of parental cultivar Cabernet Sauvignon.

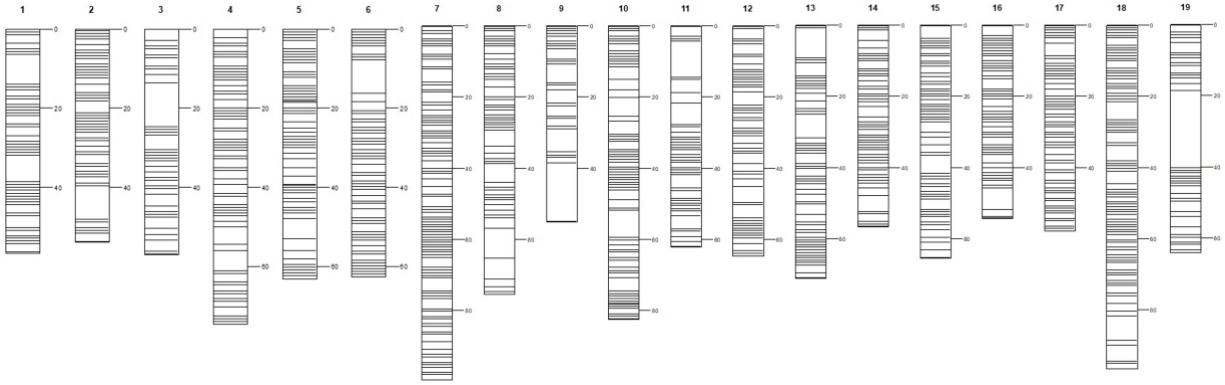


Figure 11 – Graphical representation of 19 linkage groups composing P2 (Cabernet Sauvignon) linkage map.

4.4 Phenological traits measurements

Phenotypic data collected for the budbreak, flowering and veraison time, the intervals among them and finally the ripening were collected in the 2020 and 2021 vintages on the two copies located in Gozzo's vineyard (Verona) and in 2021 for the only copy present in Guerrieri-Rizzardi vineyard's (Negrar di Valpolicella). Below, for each trait, are presented the graphics with mediated distributions for the two copies of Gozzo's vineyard (Left and Right) and for Guerrieri-Rizzardi copy. The abscissas show intervals of Julian days, while the ordinates show the number of plants that fall within the interval of the histogram. Each copy of the CsxC population is composed of 142 individuals but given the young age of the plants it was not always possible to collect the phenotypic data on the total number of plants. The black and orange arrows indicate the behavior of the two parents, Corvina (C) and Cabernet Sauvignon (Cs) respectively.

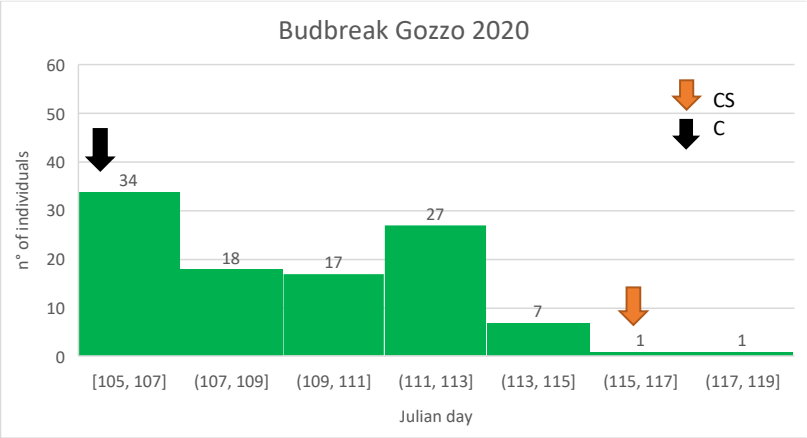
4.4.1 Budbreak time distribution

Budbreak time was recorded according to the criteria described in *paragraph 3.6.1* (page 31). **Graphics 1a-b** show average distributions of the Gozzo's vineyard; parents show a reproducible behavior in the two years.

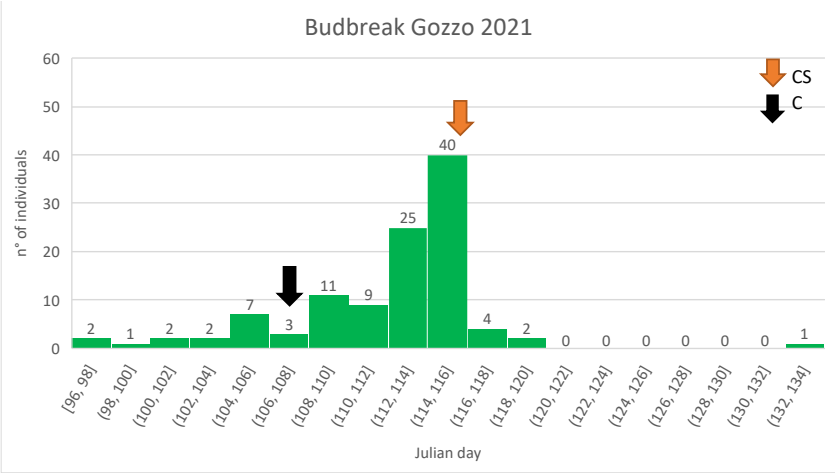
In Guerrieri-Rizzardi's vineyard (**graphic 1c**), on the other hand, they seem to have an opposite behavior compared to that in Gozzo's vineyard, with a precocity in the budding period of Cs compared to C. As regards the CsxC populations of Gozzo's

vineyard, in **graphics 1a-b**, we observed a delay in the budding period 2021 compared to the previous year and a greater extension of this phenological process. Instead, population of Guerrieri-Rizzardi's vineyard shows a slight anticipation compared to Gozzo's vineyard, probably due to the climatic difference of cultivation area and with a tendentially transgressive behavior of the population compared to the parents.

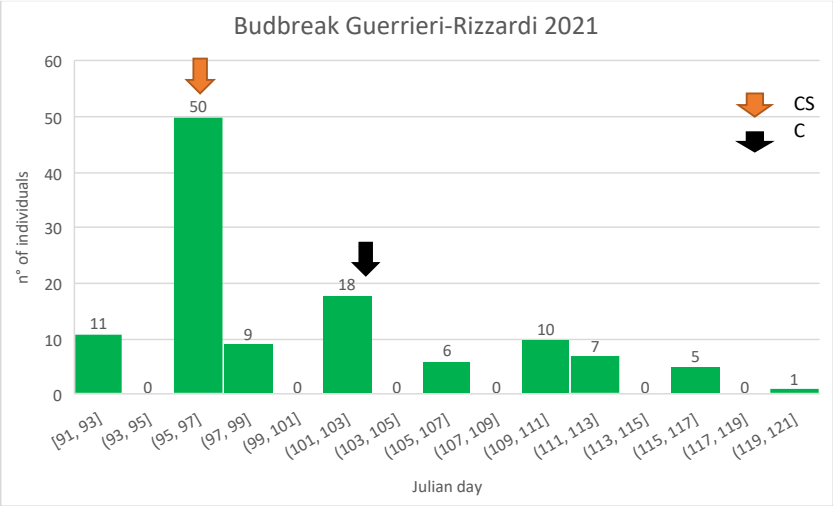
Except for the 2020 vintage, the surveys collected in 2021 for both vineyards show a normal trend in distribution.



Graphic 1a – Distribution of averaged budbreak time (left-right) in CsxC population in Gozzo's vineyard 2020



Graphic 1b - Distribution of averaged budbreak time (left-right) in CsxC population in Gozzo's vineyard 2021



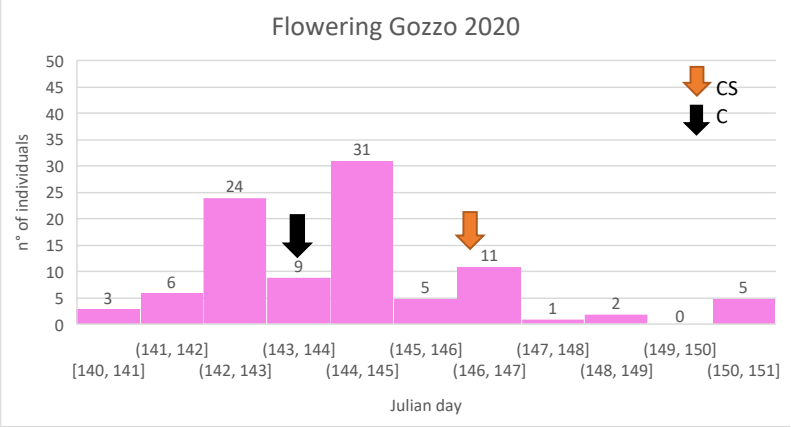
Graphic 1c - Distribution of budbreak time in CsxC population in Guerrieri-Rizzardi's vineyard 2021

4.4.2 Flowering time distribution

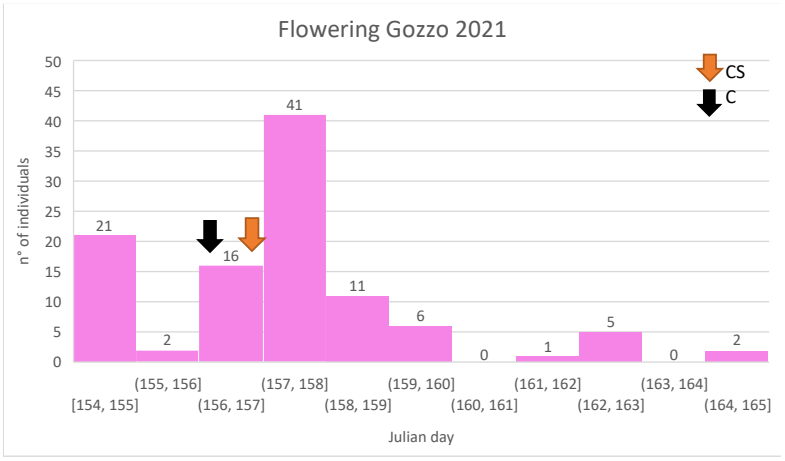
Similarly, to budbreak, data collected for the copies in the Gozzo's vineyard were averaged for flowering and graphed in histograms. Observing the behavior of the parents (**graphics 2a-b**) C flowers earlier compared to Cs, but this is less evident in Guerrieri-Rizzardi's vineyard (**graphic 2c**).

Overall, we observed a delay in the 2021 flowering time compared to the previous year, both for Gozzo's and Guerrieri-Rizzardi's vineyards. This delay is probably caused by late frosts which occurred in spring 2021 that, in addition to damaging buds and flowers of various plants, postponed the entire phenology.

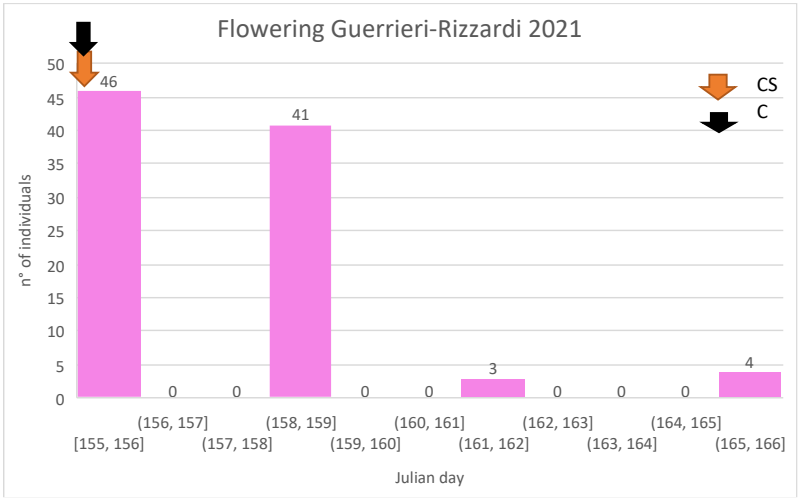
The distributions in Gozzo's vineyard tends to a normal trend, while in Guerrieri-Rizzardi's vineyard it is more difficult to interpret due to the low number of surveys. In any case, it is possible to note a transgressive behavior in the populations.



Graphic 2a - Distribution of averaged flowering time (left-right) in CsxC population in Gozzo's vineyard 2020



Graphic 2b - Distribution of averaged flowering time (left-right) in CsxC population in Gozzo's vineyard 2021



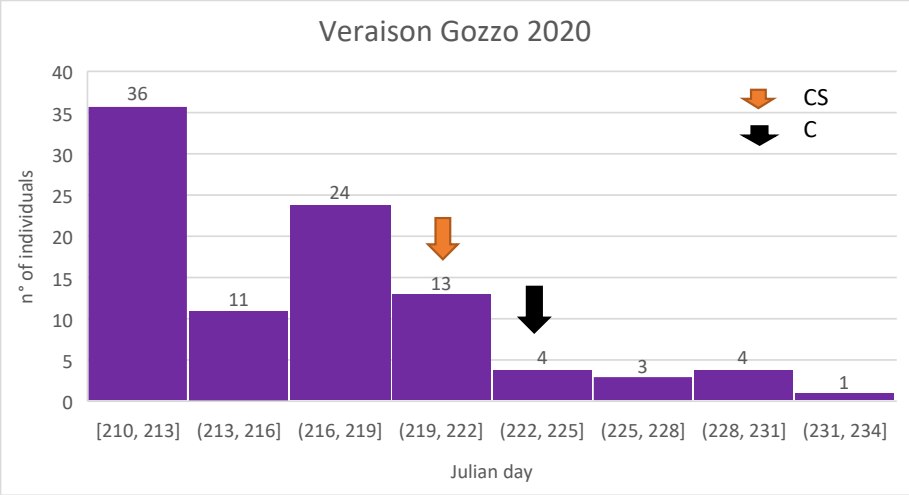
Graphic 2c - Distribution of flowering time in CsxC population in Guerrieri-Rizzardi's vineyard 2021

4.4.3 Veraison time distribution

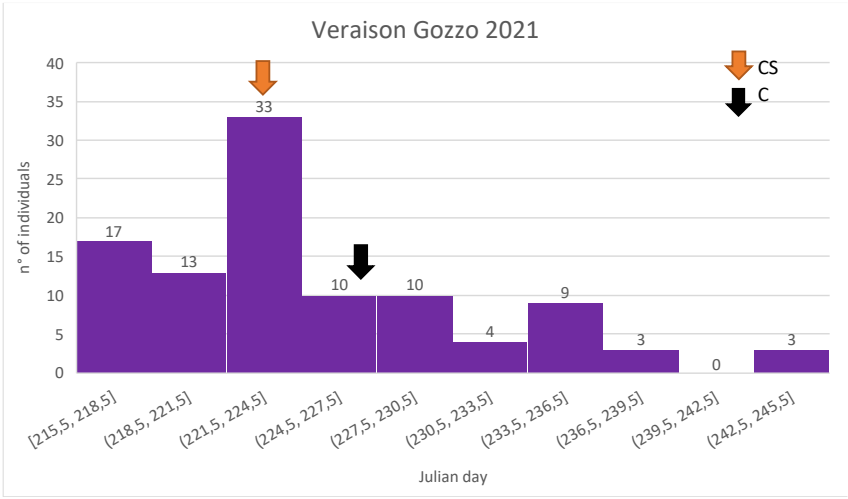
Veraison time was analyzed and recorded on the clusters produced by the shoots and flowers characterized in the previous phenological observations.

Graphics 3a-b of Gozzo's populations shows that the veraison times of C and Cs are reproducible in the two vintages. Furthermore, these are also reproducible in Guerrieri-Rizzardi's vineyard (**graphic 3c**).

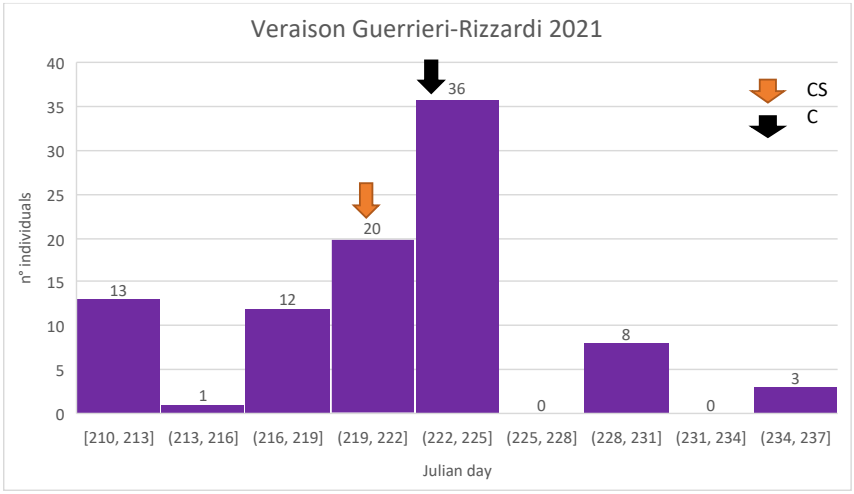
On the other hand, observing the trend of Gozzo's populations, in 2021 it is possible to detect a shift of about 6 days in veraison period (**graphics 3a-b**) compared to 2020 and to Guerrieri-Rizzardi's vineyard, probably due to the same delay already shown from the other phenological phases in this vintage. Also in this case, it is visible a transgressive behavior for all populations with respect to parents.



Graphic 3a - Distribution of averaged veraison time (left-right) in CsxC population in Gozzo's vineyard 2020



Graphic 3b - Distribution of averaged veraison time (left-right) in CsxC population in Gozzo's vineyard 2021



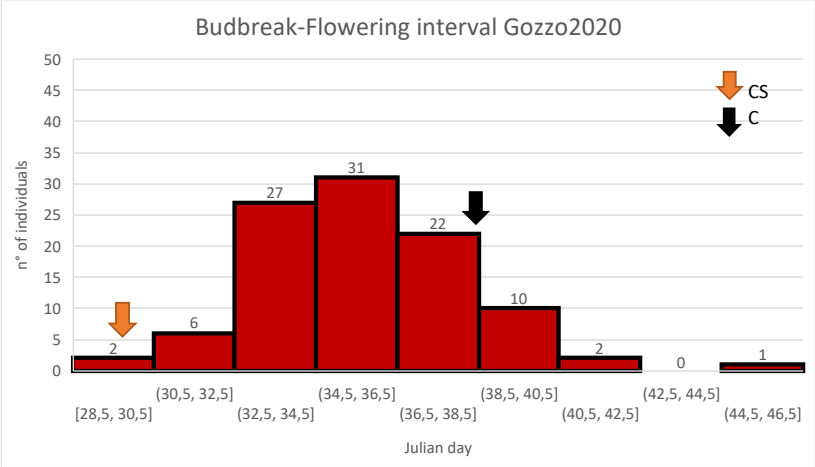
Graphic 3c - Distribution of veraison time in CsxC population in Guerrieri-Rizzardi's vineyard 2021

4.4.4 Budbreak-Flowering interval distribution

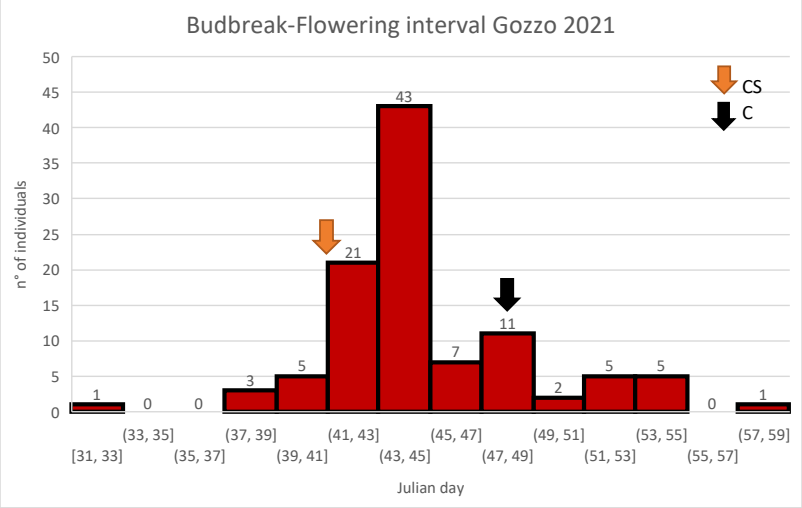
The intervals between the various phenological phases were calculated for each genotype between the single periods recorded in each vintage. Also in this case, for the two copies of Gozzo's population the average values between the two are reported, while for the copy of Guerrieri-Rizzardi's only 2021 data is reported.

In the two vintages in Gozzo's vineyard (**graphics 4a-b**) we observed a divergent behavior of parents, that was less obvious in Guerrieri-Rizzardi's copy. As regards to the extension of the interval, we can see different behaviors: the 2020 interval in Gozzo's vineyard is the shortest (**graphic 4a**), Guerrieri-Rizzardi's vineyard

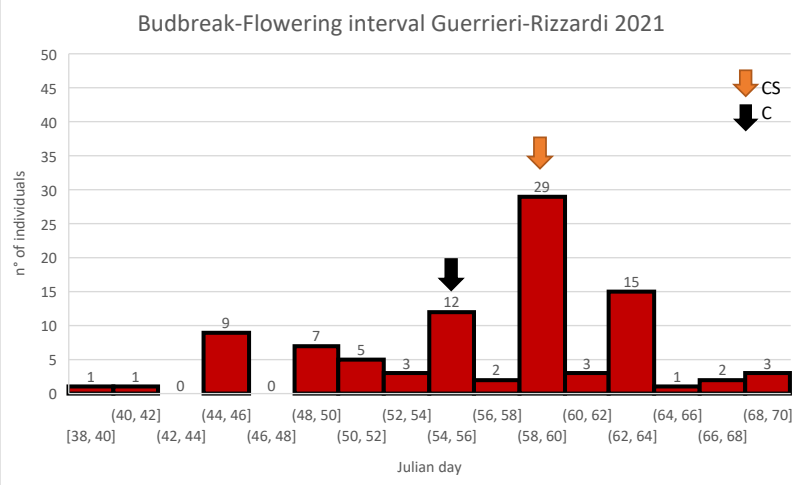
instead presents the interval longer (**graphic 4c**) while 2021 in Gozzo's is intermediate (**graphic 4b**). However, a generally transgressive normal distribution with respect to that of parents C and Cs is generally noted, with the exception of **graphic 4a.**



Graphic 4a - Distribution of the trait "interval budbreak-flowering" in CsxC population in Guerrieri-Rizzardi's vineyard 2020



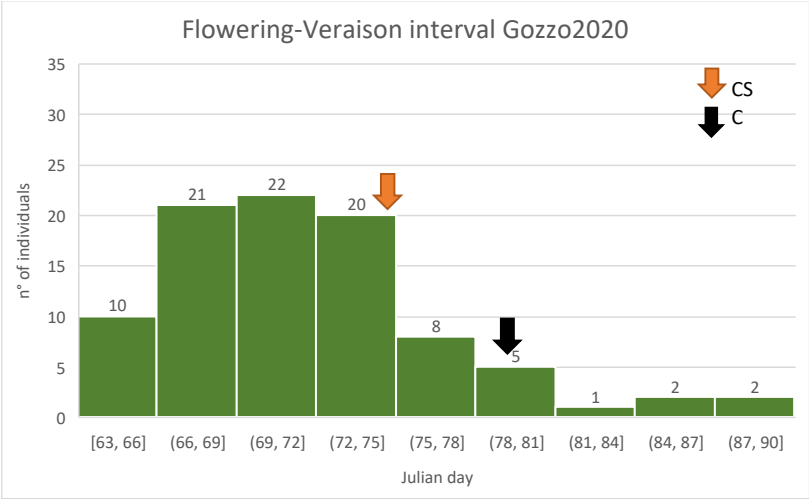
Graphic 4b - Distribution of the trait "interval budbreak-flowering" in CsxC population in Guerrieri-Rizzardi's vineyard 2021



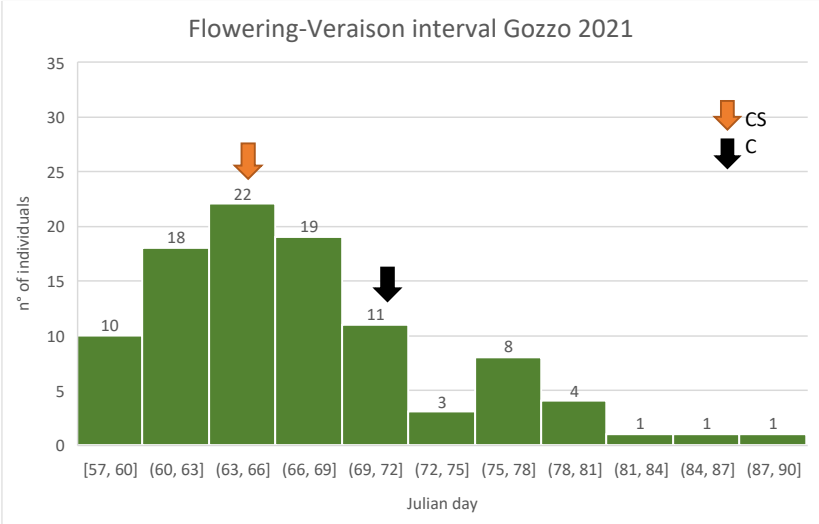
Graphic 4c - Distribution of the trait "interval budbreak-flowering" in CsxC population in Guerrieri-Rizzardi's vineyard 2021

4.4.5 Flowering-Veraison interval distribution

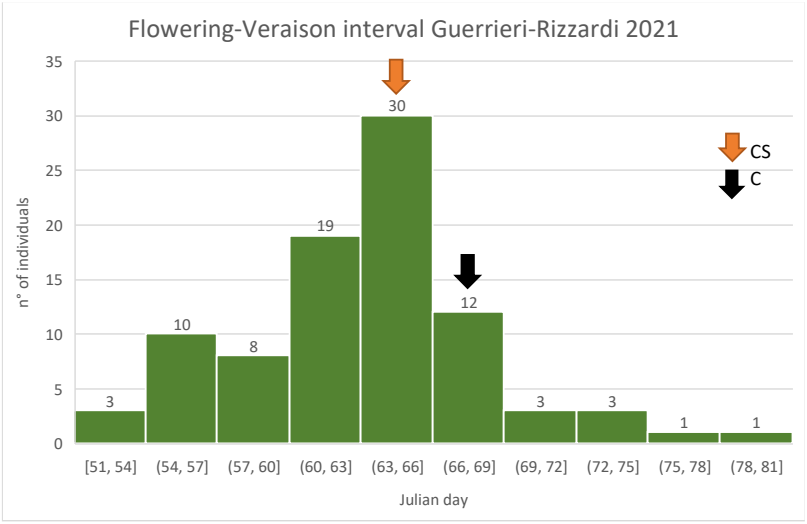
Parents show an overall good reproducibility between vintages and populations. Population distributions, on the other hand, show similar trend in the two Gozzo's vintages (**graphics 5a-b**), while Guerrieri-Rizzardi's vineyard seems to have a slightly narrower range (**graph 5c**). In any case, most individuals show an interval between 63-70 days in all graphics, with a marked transgressive behavior respect to the parents.



Graphic 5a - Distribution of the trait "interval flowering-veraison" in CsxC population in Gozzo's vineyard 2020



Graphic 5b - Distribution of the trait "interval flowering-veraison" in CsxC population in Gozzo's vineyard 2021

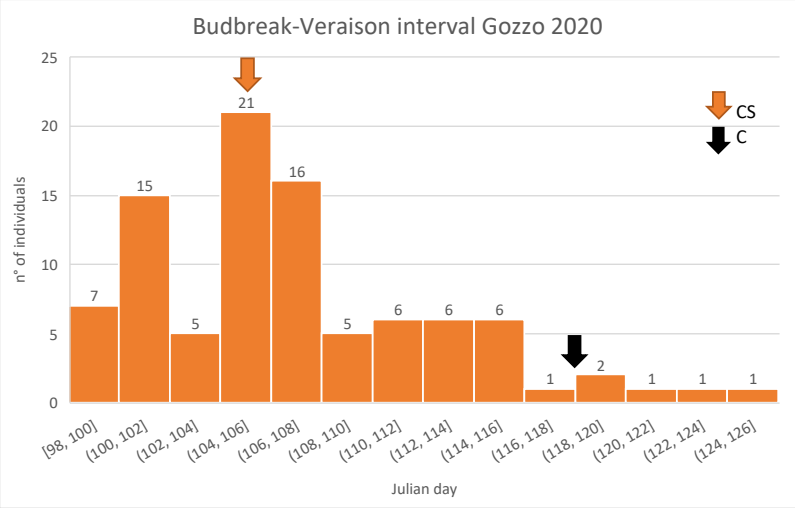


Graphic 5c - Distribution of the trai "interval flowering-veraison" in CsxC population in Guerrieri-Rizzardi's vineyard 2021

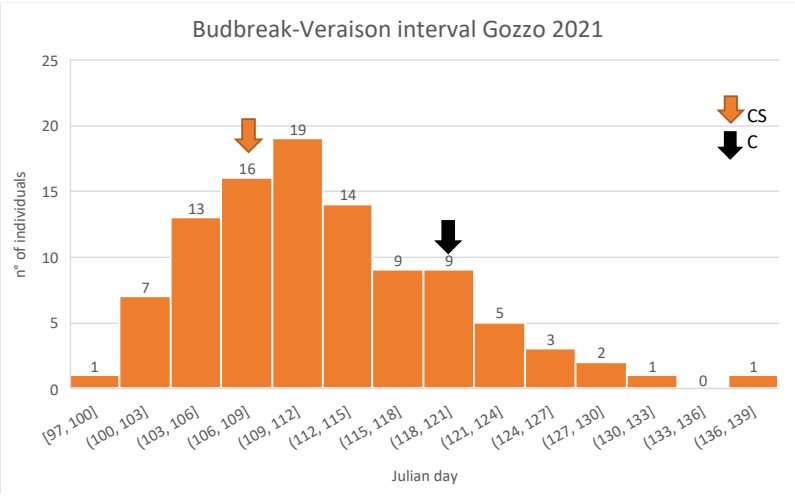
4.4.6 Budbreak-Veraison interval distribution

For budbreak-veraison interval parents show a divergent and reproducible behavior in the two years of the Gozzo's population (**graphics 6a-b**), while behavior was not reproducible in Guerrieri-Rizzardi's population in line with the observed not reproducible behavior already observed in budbreak (**graphic 6c**).

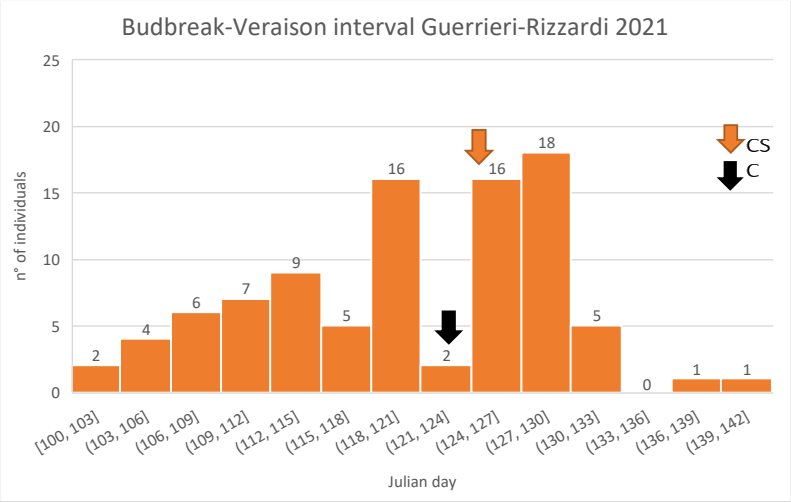
The extension of the whole interval looks similar in the three graphics but, interestingly, in Guerrieri-Rizzardi's population (**graph 6c**) most individuals have a wider interval (121-130 days) compared to Gozzo's vineyard.



Graphic 6a - Distribution of the "interval budbreak-veraison" in CsxC population in Gozzo's vineyard 2020



Graphic 6b - Distribution of the trait "interval budbreak-veraison" in CsxC population in Gozzo's vineyard 2021



Graphic 6c - Distribution of the trait interval budbreak-veraison in CsxC population in Guerrieri-Rizzardi's vineyard 2021

4.5 Technological trait measurements

4.5.1 Maturity distribution

Ripening process was followed by measuring of Brix degree according to the criteria described in *chapter 3.6.4* (page 34) on each genotype that produced grapes. Unfortunately, some genotypes did not produce grapes during the vintages, since they were still unstable due to young age.

Maturity was scored as the day the berries reached a brix degree value equal to 19; however, by testing brix 3 times per week, it could happen that we missed the exact day that brix value reaches 19°. Therefore, exact maturity dates were interpolated from sugar growth curve.

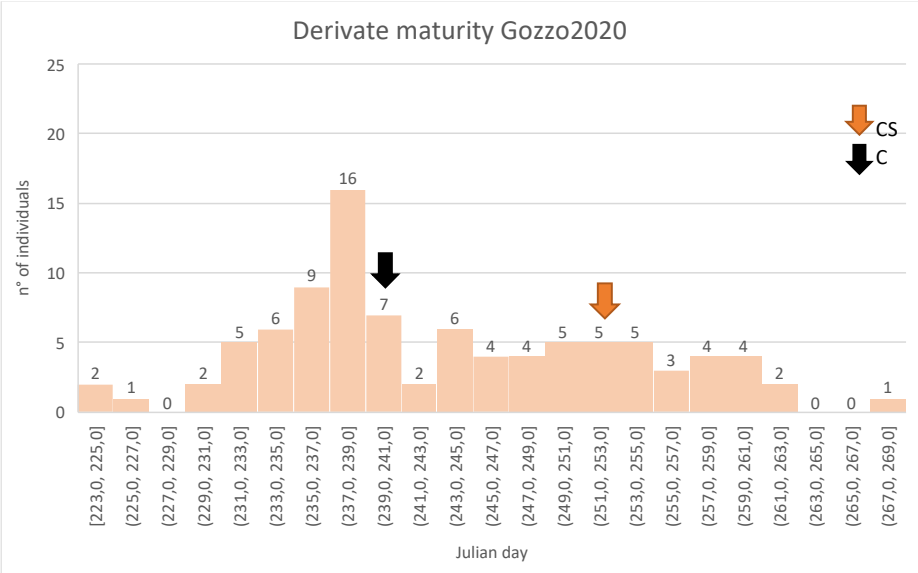
The graphs of the derived maturity mediated for the copies of Gozzo's population and those for the copy of Guerrieri-Rizzardi's vineyard are shown below.

Parents show a non-reproducible behavior between the 2020 and 2021 vintages in Gozzo's vineyard, with a marked divergence in 2020 and an inversion of the timing of maturity in 2021, that is however similar to the behavior of parents in Guerrieri-Rizzardi's vineyard in 2021 (**graphic 7a, b and c**).

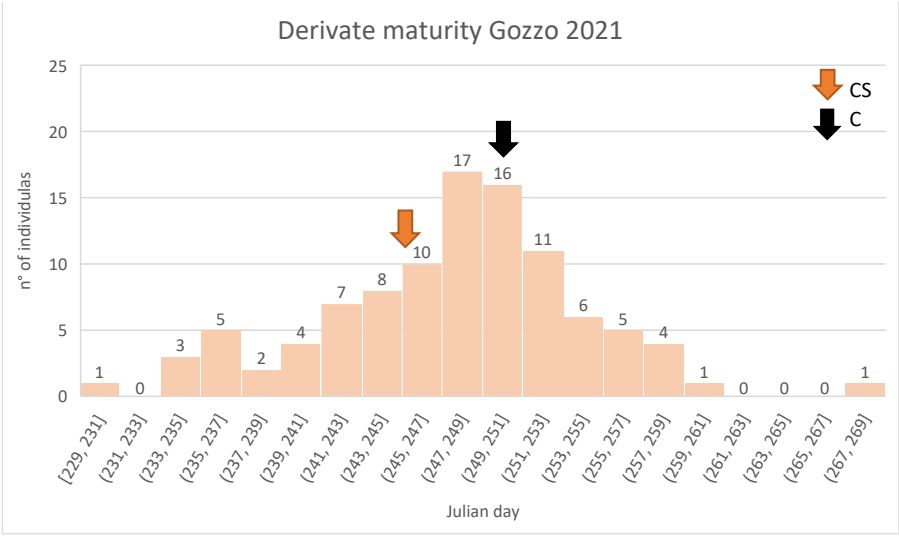
Furthermore, observing the trend of the populations in the three graphs, we saw a difference in the ripening intervals: **graphic 7c** shows a wider interval in Guerrieri-Rizzardi's population (about 66 days) compared to Gozzo population in the same

year. Moreover the 2020 vintage of Gozzo's vineyard (**graphics 7a**) was significantly earlier than in 2021 (**graphics 7b**), with the peak of maturity at day 238 in 2020 compared to the peak at day 247 in 2021.

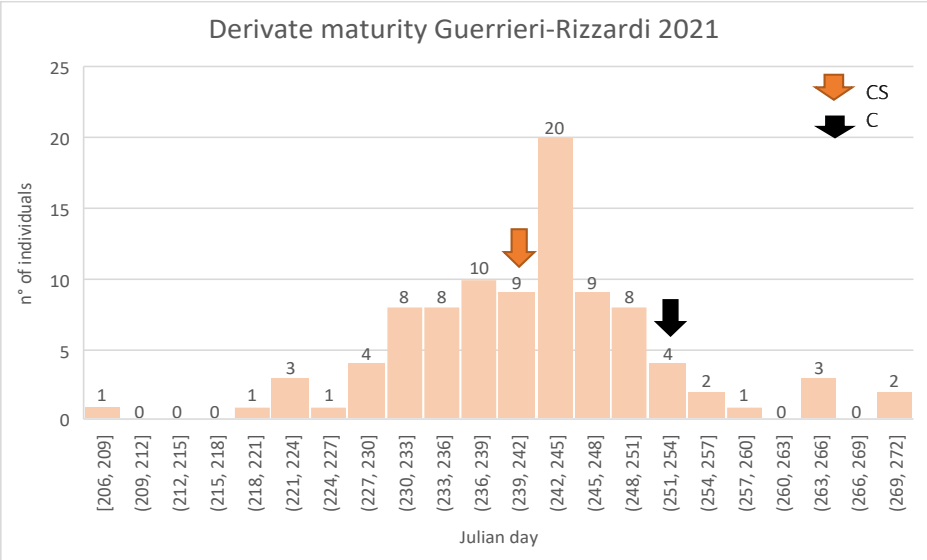
The trend of distributions has in all cases a normal and transgressive trend with respect to the parents.



Graphic 7a - Distribution of maturity time (brix 19°) in CsxC population in Gozzo's vineyard 2020



Graphic 7b - Distribution of maturity time (brix 19°) in CsxC population in Gozzo's vineyard 2021

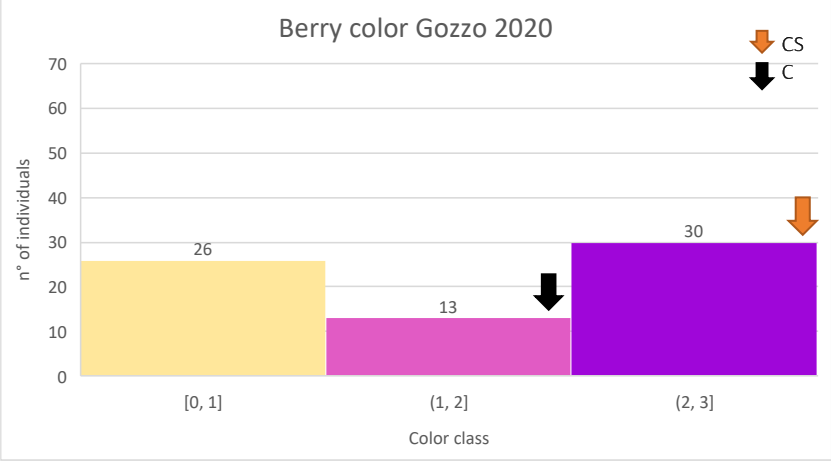


Graphic 7c - Distribution of maturity time (brix 19°) in CsxC population in Guerrieri-Rizzardi's vineyard 2021

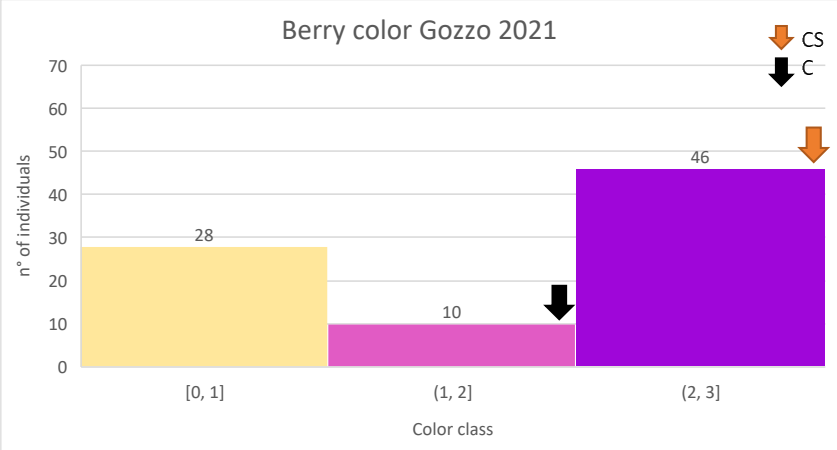
4.5.2 Berry color

The genetic control of berry color is well understood in grapevine. A cluster of four *Myb* transcription factor genes located on chromosome 2 controls the largest part of the quantitative variation in berry color (Kobayashi et al., 2004; This et al., 2007; Fournier-Level et al., 2009). However, for colored grapes, the berry color varies from the lightest pink to the darkest purple-black, and the locus on chromosome 2 doesn't fully explain these differences. Some studies already demonstrated the contribution of further minor loci and mapped additional contributions (Cavallini et al., 2015, Tirumalai et al., 2019, Sun et al., 2020). With the aim of mapping minor contributions in our segregating population, we considered therefore here the color variation as a quantitative trait by using a color scale as explained in *paragraph 3.6.5,* pages 35-36.

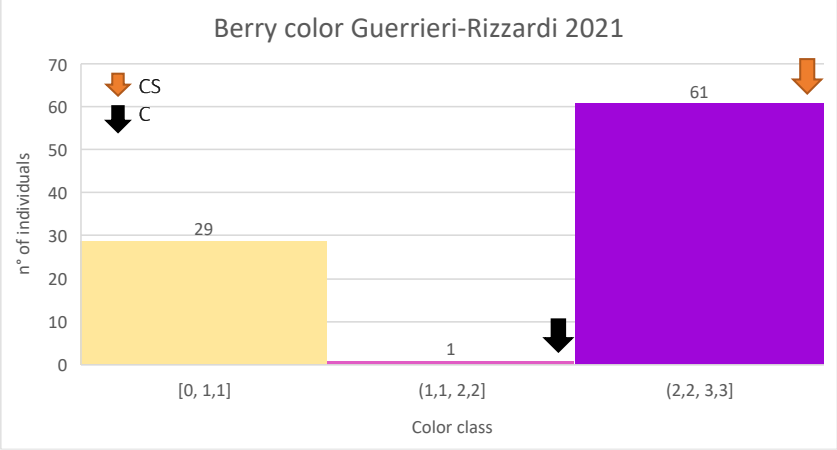
Graphics 8a-c show the color distributions of clusters produced in different vintages and for the two vineyards. As expected, a marked bimodal was confirm in line with the expected oligogenic inheritance of the trait.



Graphic 8a - Distribution of berry color in CsxC population in Gozzo's vineyard 2020



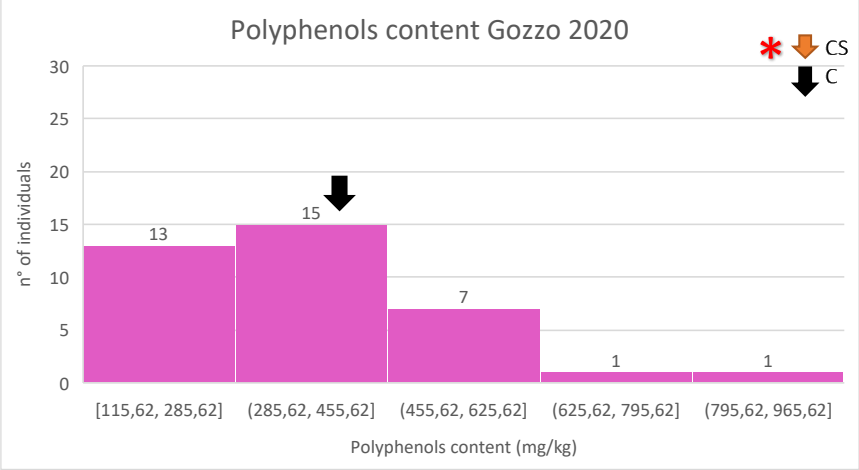
Graphic 8b - Distribution of berry color in CsxC population in Gozzo's vineyard 2021



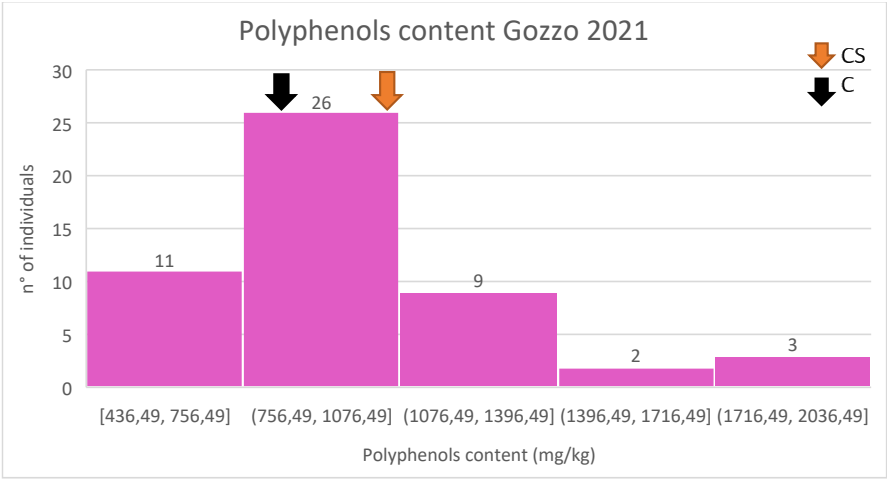
Graphic 8c - Distribution of berry color in CsxC population in Guerrieri-Rizzardi's vineyard 2021

4.5.3 Polyphenols content

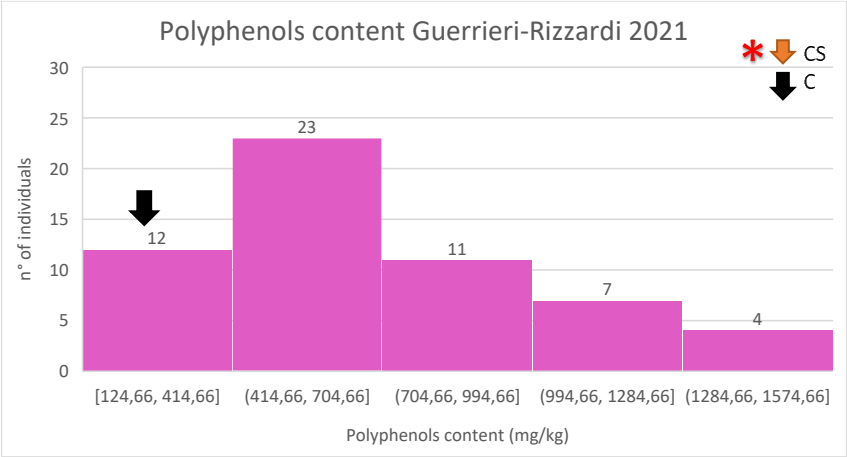
Regarding the polyphenols content in berries in all cases we observed a normal distribution of the data. Unfortunately, in Gozzo 2020 and Guerrieri-Rizzardi 2021 vintages (**graphic 9a** and **9c**) the parental Cabernet Sauvignon did not produce enough grapes for the evaluation of polyphenols content. The parental Corvina showed on average a lower content compared to the whole population average value.



Graphic 9a - Distribution of polyphenols content in CsxC population in Gozzo's vineyard 2020
*For this trait no data on Cs were available in 2020 due to low availability of berries



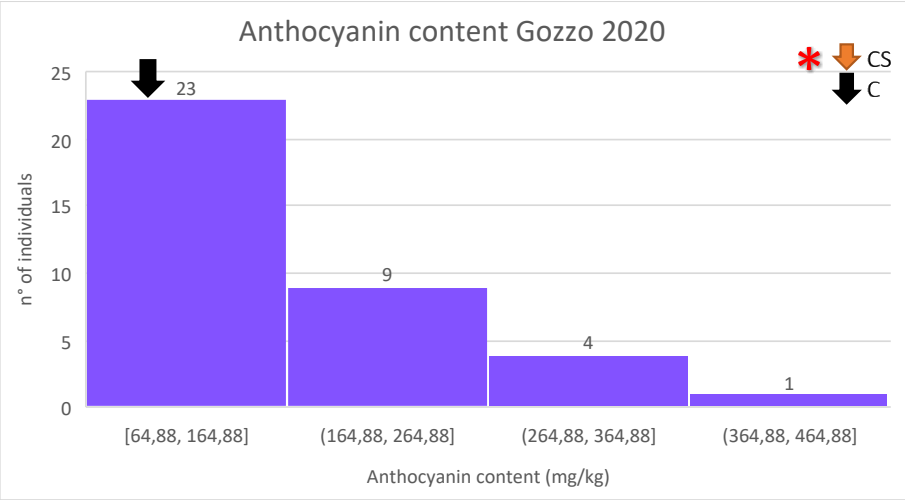
Graphic 9b - Distribution of polyphenols content in CsxC population in Gozzo's vineyard 2021.



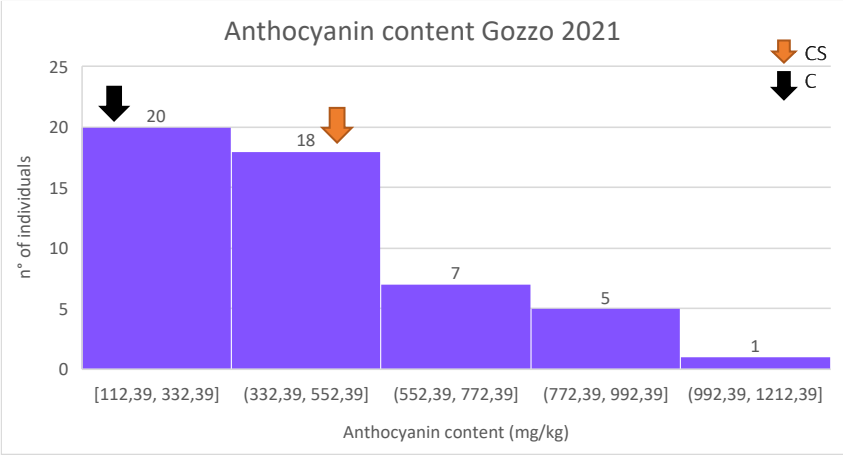
Graphic 9c - Distribution of polyphenols content in CsxC population in Guerrieri-Rizzardi's vineyard 2021.
*For this trait no data on Cs were available in 2021 due to low production of berries

4.5.4 Anthocyanin content

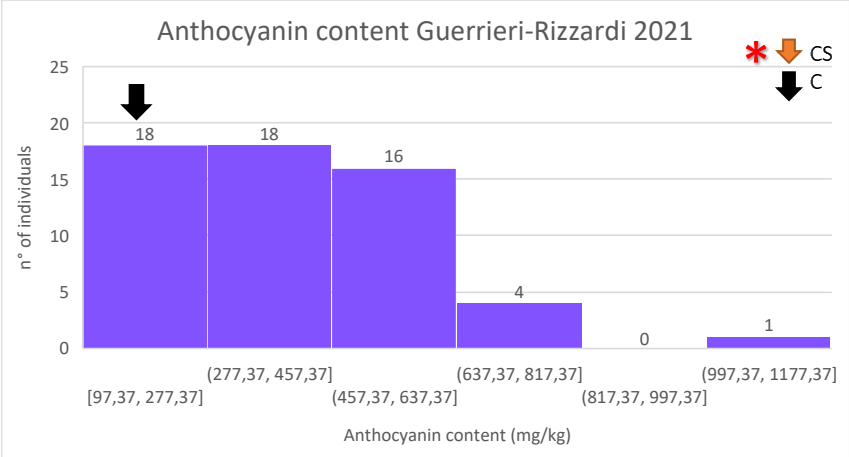
As for the whole polyphenols content, also for anthocyanin it was not possible to analyze Cs in Gozzo 2020 and Guerrieri-Rizzardi 2021 vintages. Corvina always showed a lower in anthocyanins content compared to the population average content. In 2021(**graphic 10b**) a higher accumulation was found compared to 2020.



Graphic 10a - Distribution of anthocyanin content in CsxC population in Gozzo's vineyard 2020.
*For this graphic no data on Cs were available due to low availability of berries



Graphic 10a - Distribution of anthocyanin content in CsxC population in Gozzo's vineyard 2021.



Graphic 10a - Distribution of anthocyanin content in CsxC population in Guerrieri-Rizzardi's vineyard 2021.
*For this graphic no data on Cs were available due to low production of berries

4.6 Statistical analysis

As described in *paragraph 3.7* (page 37), correlation inside trait was estimated both with Pearson's correlation and with Spearman ranks function of Excel. Indeed, Spearman ranks correlation is best suited for the evaluation of across year or across location correlations, especially for the phenology related traits.

The correlation results are presented below, with comparisons between left and right copies of the CsxC population in the Gozzo's vineyard and between the copy in Guerrieri-Rizzardi's vineyard for years 2020 and/or 2021.

4.6.1 Phenological traits

TRAIT	GOZZO Left/Right 2020	GOZZO Left/Right 2021	GOZZO Left 2020/Left 2021	GOZZO Right 2020/Right 2021	GOZZO Medium 2020/Medium 2021	GOZZO Medium 2021/GUERRIERI RIZZARDI 2021
BUDBREAK	0.371	0..450	0.341	0.331	0.430	0.394
FLOWERING	0.421	0.386	0.282	0.191	0.319	0.203
VERAISON	0.423	0.537	0.398	0.397	0.423	0.325
BUDBREAK-FLOWERING INTERVAL	0.174	0.381	0.373	0.147	0.325	0.275
FLOWERING-VERAISON INTERVAL	0.475	0.519	0.464	0.389	0.387	0.371
BUDBREAK-VERAISON INTERVAL	0.378	0.476	0.384	0.290	0.386	0.340

Table 10a. Pearson's correlation of phenological traits across years and locations

TRAIT	GOZZO Left/Right 2020	GOZZO Left/Right 2021	GOZZO Left 2020/Left 2021	GOZZO Right 2020/Right 2021	GOZZO Medium 2020/Medium 2021	GOZZO Medium 2021/GUERRIERI RIZZARDI 2021
BUDBREAK	0.365	0.326	0.383	0.268	0.434	0.381
FLOWERING	0.322	0.449	0.172	0.270	0.323	0.221
VERAISON	0.404	0.546	0.391	0.499	0.492	0.216
BUDBREAK-FLOWERING INTERVAL	0.179	0.298	0.387	0.058	0.261	0.291
FLOWERING-VERAISON INTERVAL	0.470	0.536	0.468	0.485	0.492	0.323
BUDBREAK-VERAISON INTERVAL	0.373	0.460	0.423	0.365	0.407	0.348

Table 10b. Spearman ranks correlation of phenological traits across years and locations

In general, a quite low correlation was found for the trait **flowering**, both across years and locations. A slightly better correlation was seen for **budbreak**, especially across years, while the best correlation was seen for the **veraison** trait (up to 0.499 across years and 0.492 across locations, **tables 10a - 10b**).

Concerning the intervals, **budbreak-flowering interval** showed a low correlation in all the comparisons, likely due to the low correlation of flowering. On the other hand, **flowering-veraison** showed a much better correlation across year, while the correlation was still lower by considering different locations.

4.6.2 Technological parameters

TRAIT	GOZZO Left/Right 2020	GOZZO Left/Right 2021	GOZZO Left 2020/Left 2021	GOZZO Right 2020/Right 2021	GOZZO Medium 2020/Medium 2021	GOZZO Medium 2021/GUERRIERI RIZZARDI 2021
DERIVATIVE MATURITY	0.470	0.442	0.318	0.460	0.511	0.482
COLOR	0.990	0.926	0.910	0.887	0.893	0.947
POLYPHENOLS	0.768	0.655	0.614	0.734	0.107	0.232
ANTHOCYANIN	0.870	0.324	0.582	0.898	0.705	0.353

Table 10c. Pearson's correlation of technological parameters.

TRAIT	GOZZO Left/Right 2020	GOZZO Left/Right 2021	GOZZO Left 2020/Left 2021	GOZZO Right 2020/Right 2021	GOZZO Medium 2020/Medium 2021	GOZZO Medium 2021/GUERRIERI RIZZARDI 2021
DERIVATIVE MATURITY	0.549	0.459	0.376	0.501	0.494	0.539
COLOR	0.995	0.875	0.876	0.837	0.864	0.883
POLYPHENOLS	0.870	0.381	0.774	0.683	0.625	0.303
ANTHOCYANIN	0.907	0.377	0.774	0.683	0.625	0.375

Table 10d. Spearman ranks correlation of technological parameters.

Regarding the **color**, the correlation between various comparisons was, as expected, very strong (**tables 10c** and **10d**). As previously mentioned, this trait is mainly regulated by a major trait on chromosome 2. Therefore, the environmental effect had less impact here.

The berry content in **polyphenols**, as well as in **anthocyanins**, also showed good reproducibility between vintages. The correlation between locations was instead much lower: the poor correlation may indicate a difference in the levels of production and accumulation of these molecules due to the interaction of the genotypes with the different locations, thus highlighting a relevant environmental effect on these phenotypic traits.

Finally, **maturity** also showed a good correlation, comparable to veraison, that was however, in this case, also retained across locations.

4.7 QTL mapping

The goal of this analysis was the search for significant statistical correlations between genotypic and phenotypic data, possibly collected in different years, in order to identify chromosomal regions putatively involved in phenological and qualitative traits in a segregating population of *Vitis vinifera*.

The genotypic data used are represented by three linkage maps: the P1 map, built with Corvina's genetic data, the P2 map, built with Cabernet Sauvignon's genetic data and the CP map or consensus map, generated by the integration of the two parental maps.

By using the three maps, it was possible to evaluate the contribution of the two parents for each trait and identify chromosomal regions significantly involved.

Phenotypic data in more years and on multiple replicas/locations of CsxC population allowed to score reproducible results.

The permutation test was applied to define LOD thresholds of significance for each QTL at chromosome level. These thresholds of significance were applied in order to select only the statistically significant regions to in the QTL analysis outputs.

QTL analysis was done according to different methods:

- Interval Mapping provided a first overview on the regions putatively associated with traits

- MQM analysis represents a method for searching for QTL regions in a more precise and powerful way compared to IM, using cofactors to take into account also the possibility of multiple QTLs.

4.7.1 Interval Mapping (IM)

The IM analysis was conducted on a total of 9 traits: the input files for the analysis were processed as described in *paragraph 3.8.1* (page 38). The quantitative data file (*qua.file*) was prepared with phenological and technological data collected in the different vintages and for different replicas of the population.

Furthermore, for each trait, independent runs were launched using both the CP map and the two parental maps P1 and P2 (*map.files*), each with related genetic datasets (*loc.file*). For the phenology in Gozzo's population, both the left and right replicas were run separately and the average between the two, while for the population of Guerrieri-Rizzardi data was run on the only copy available.

In *Appendix* (pages 151-169) **tables 27-36** are reported the complete results of the IM analysis.
Reproducible QTLs between the different maps and between replicates are reported and discussed instead below in **tables 11 a-b**. In details peaks of significance including map position, corresponding marker, LOD score, percentage of variance explained as well as the confidence interval are reported. To be considered reproducible, QTLs must be found in at least two years or locations and with an intersection between their confidence intervals.
For phenology related traits, starting from the **budbreak**, it is possible to note that for CP map a QTL was detected on LG7 with an overall confidence interval ranging from 27.945 to 51.334 cM, whose maximum value of LOD 5.71 is reported at 31.147 cM of chromosome 7, shared by both 2020 and 2021 datasets in both vineyards. Similarly, P2 map showed a QTL always positioned on LG7 with an interval between 4.276 and 52.26 cM and with a LOD peak of 5.37 at 34.354 cM, shared between 2020 and 2021 vintages of both vineyards.
Looking at **flowering**, CP map highlighted two regions: one on LG2 and one on LG7. The first, on chromosome 2, has a confidence interval between 19.214 and 26.787 with an LOD peak of 6.27 at 24.659 cM, shared by both vineyards but only for the vintage 2021. The second region, located on the chromosome 7, has a larger range, from 21.399 to 86.160 cM, with a LOD peak of 5.13 at 86.160 cM and detectable between the two vintages of Gozzo's vineyard but not in Guerrieri-Rizzardi's.Two distinct regions on the LG7 have been also found on P2 map: the first lies in an interval between 0 and 23.675 cM with the peak of 2.49 to 5.694 cM. The second, on the other hand, has an interval between 56.526 and 95.167 cM, with an LOD peak of 2.32 at 85.865 cM. Both regions have been identified in the two vintages of Gozzo's vineyard but not in Guerrieri-Rizzardi's.

The analysis of **veraison** highlighted from the CP map a wide region on LG5, between 1.789 and 33.194 cM with maximum LOD peaks of 3.99 at 2.328 (Gozzo 2020) and 22.838 cM (Guerrieri-Rizzardi 2021). From the P2 map, on the other hand, a region on LG17 between 56.402 and 57.821 cM was selected for veraison with a maximum LOD of 3.07 to 57.821 cM, reproducible only for 2021 vintage in both vineyards.

Observing the results for **phenological intervals**, the **budding-flowering interval** analysis highlighted again a region on LG7 from both the CP and P2 maps as already found in budbreak and flowering. For CP map, the region is in a confidence interval between 43.99 and 58.637 cM, with an LOD peak of 4.11. 54.706 cM. For P2 map, on the other hand, the region extends between 23.675 and 59.363 cM with a maximum LOD value of 3.92 at 52.26 cM. Both regions were found in 2020 and 2021 vintages of Gozzo's vineyard. Interestingly one of the associated markers for the QTL on LG7 found for budbreak was found here too.

Regarding **flowering-veraison interval**, only from CP map a reproducible region was detected on LG5 in a range of 2.868 and 33.194 cM, with a peak of 3.73 at 11.80 cM and found in Gozzo 2020/Guerrieri-Rizzardi 2021 vintages.

Finally, as regards **budding-veraison interval**, reproducible QTLs were detected only from P1 map on LG4 and LG12. The first region, located on chromosome 4, is between 0 and 13.5 76 cM with a LOD peak of 2.26 at 2.847 cM, reproducible only for Gozzo's vineyard. The second region, on chromosome 12, is instead between 15.79 and 19.346 cM with a peak of 2.4 at 19.346 cM and reproducible in the two vineyards only for 2021 vintage.

Table 11a - List of reproducible QTLs obtained from Interval Mapping on phenological traits data.

		Peak position						*Confidence Interval*	
Trait	*Dataset*	*Map*	*Linkage Group*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position (cM)*
Budbreak	*Gozzo Medium 2020*	**CP**	LG7	31.147	7_7430930	5.71	22.3	4.99 - 5.71 - 4.91	28.657 - 49.421
	Guerrieri-Rizzardi 2021			31.147	7_7579292	3.85	14.2	3.32 - 3.85. 3.46	27.945 - 36.312
	Gozzo L 2021			38.985	7_13388305	4.91	19.6	4 - 4.91 - 4.2	33.635 - 51.334
	Gozzo Medium 2021			35.058	7_10409620	4.88	18.8	4.05 - 4.88 - 4.08	31.147 - 43.99
	Gozzo Medium 2020	**P2**	LG7	34.354	7_11129686	5.37	21.2	4.55 - 5.37 - 4.51	25.104 - 50.842
	Gozzo L 2021			42.916	5_20326151	4.34	17.5	3.54 - 4.34 - 4.07	30.078 - 52.26
	Guerrieri-Rizzardi 2021			25.812	7_7348129	2.28	9	1.95 - 2.28 - 2.03	4.276 - 32.925
Flowering	*Gozzo L 2021*	**CP**	LG2	24.659	2_5557145	6.27	26	5.76 - 6.27 - 5.56	19.214 - 25.723
	Gozzo R 2021			23.58	Un_12299152	3.72	19.1	3.01 - 3.72 - 2.93	21.069 - 26.787
	Gozzo Medium 2021			23.58	Un_12299152	6.51	25.1	5.54 - 6.51 - 6.15	19.283 - 24.723
	Guerrieri-Rizzardi 2021			22.511	Un_15535432	3.16	14.5	3.03 - 3.16 - 3.16	21.069 - 22.511
	Gozzo R 2020		LG7	56.129	Un_12388826	3.67	19.7	3.34 - 3.67 - 3.22	29.369 - 58.814
	Gozzo Medium 2020			31.502	7_7830862	3.46	15.3	3.03 - 3.46 - 3.02	21.399 - 33.989
	Gozzo R 2021			82.218	7_17591750	5.13	25.3	4.48 - 5.13 - 4.22	74.015 - 86.16
	Gozzo Medium 2021			82.218	7_17591750	3.85	15.7	3.28 - 3.85 - 3.56	41.831 - 82.931
	Gozzo L 2020	**P2**	LG7	5.694	7_2780583	2.24	12	2.21 - 2.24 - 1.82	0 - 13.139
	Gozzo Medium 2020			5.694	7_2780583	2.49	11.3	2.26 - 2.49 - 2.1	3.567 - 18.638
	Right 2021			16.5	7_4810487	1.75	9.5	1.7 - 1.75 - 1.73	7.853 - 23.675
	Gozzo Medium 2020		LG7	57.944	Un_12338449	2.06	8.7	1.8 - 2.06	56.526 - 57.944
	Gozzo R 2020			76.552	7_912711	2.02	11.4	1.83 - 2.02 - 2.02	56.526 - 76.552
	Gozzo R 2021			85.865	7_17089269	2.32	12.3	1.83 - 2.32 - 1.78	55.107 - 95.167
	Gozzo Medium 2021			86.574	7_17354354	1.95	8,.3	1.95 - 1.93	86.574 - 88.733

Trait	*Dataset*	*Map*	*Linkage Group*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position (cM)*
Veraison	*Gozzo L 2020*	**CP**	LG5	22.838	5_6552770	3.99	21	3.23 - 3.99 - 3.43	14.658 - 33.194
	Guerrieri-Rizzardi 2021			2.328	5_1143888	3.99	19.1	3.4 - 3.99 - 3.41	1.789 - 19.272
	Gozzo L 2021	**P2**	LG17	57.111	17_14858602	2.16	11.2	2.05 - 2.16 - 2.16	56.402 - 57.111
	Guerrieri-Rizzardi 2021			57.821	17_16780652	3.07	14.2	2.32 - 3.07	56.402 - 57.821
Interval budbreak-flowering	*Gozzo R 2020*	**CP**	LG7	54.706	Un_14300857	4.11	22	3.66 - 4.11 - 3.39	43.99 - 55.417
	Gozzo Medium 2020			54.706	Un_14300857	3.38	14.4	3.21 - 3.38 - 3.2	37.566 - 58.637
	Gozzo L 2021			50.445	7_469791	3.95	17.9	3.04 - 3.953.25	33.635 - 52.578
	Gozzo R 2020	**P2**	LG7	57.235	Un_14300857	3.17	17.5	2.31 - 3.17 - 2.46	32.925 - 59.363
	Gozzo M 2020			34.354	7_11129686	3.07	13.2	2.28 - 3.07 - 2.69	23.675 - 52.26
	Gozzo L 2021			52.26	7_297044	3.92	17.8	2.93 - 3.92 - 2.95	30.078 - 56.526
	Gozzo Medium 2021			4.985	7_2350141	2.48	10.5	2.35 - 2.48 - 1.81	1.429 - 57.944
Interval flowering-veraison	*Gozzo Medium 2020*	**CP**	LG5	29.638	5_8455004	3.37	15.8	3.13 - 3.37 - 3	22.838 - 33.194
	Guerrieri-Rizzardi 2021			11.8	mu_5_3513954	3.73	17.5	3.1 - 3.73 - 3.44	2.868 - 23.193
Interval budbreak-veraison	*Gozzo Medium 2020*	**P1**	LG4	0	4_3181677	1.77	8.3	1,77	0
	Gozzo R 2021			2.847	4_4384093	2.26	13.2	1.8 - 2.26 - 1.72	0 - 13.579
	Gozzo R 2021		LG12	15.79	12_3084744	1.8	10.6	1,8	15.79
	Guerrieri-Rizzardi 2021			19.346	12_3914492	2.4	11.4	1.89 - 2.4	15.79 - 19.346

Examining the results related to **color**, we note the presence of a very solid QTL in all populations analyzed and on three maps. Specifically, CP map shows a range on LG2 between 39.252 and 41.742 cM with a very consistent LOD value of 32.78 at 39.252 cM. Looking at P1 map instead, the interval is shown in the positions from 38.922 - 48.203 cM, while P2 map shows it between 33.884 - 39.578 cM. Furthermore, P2 map shows reproducible QTLs also on LG6 and LG17, respectively with ranges between 31.961 - 58.378 cM and LOD peak of 3.2 at 51.964 cM and 41.426 - 57.821 cM with peak of 2.67 at 56.402 cM, also very reproducible between vintages.

The analysis on **polyphenols** content showed a reproducible region between the two vintages of Gozzo and shared by both CP map and P1 map on the LG3: in CP map the region extends from 39.108 and 49.016 cM with a maximum LOD of 3.56 at 49.016 cM, while in map P1 it is located between 32.111 and 44.729 cM with a peak of 3.87 at 36.429 cM.

Finally, as regards the **anthocyanin** content, on CP map a reproducible region between the two vintages of Gozzo on LG2 has been identified and which approximates the region identified for the color of the berry (*CP: 39.252-41.742 cM; P2: 33.884-39.578 cM*), confirming the goodness of the analysis. The region putatively involved in the anthocyanin content is located between 33.184 and 40.319 cM with a maximum LOD of 4.83 at 36.392 cM.

Table 11b - List of reproducible QTLs obtained from Interval Mapping on technological traits data.

		Peak position						*Confidence Interval*	
Trait	*Data source*	*Map*	*Linkage group*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position (cM)*
Berry color	*Gozzo Medium 2020*	**CP**	LG2	3961	2_14208969	17	67.9	17	39.61
	Gozzo Medium 2021			40.319	2_14702692	26.75	77	26.45 – 26.75 – 26.75	39.252 – 40.319
	Guerrieri-Rizzardi 2021			39.252	2_13398576	32.78	85.3	32.78 – 32.16	39.252 – 40.319
	Gozzo Medium 2020	**P1**	LG2	43.907	2_15113892	2.83	17.3	1.88 – 2.83 – 2.15	38.922 – 48.203
	Gozzo Medium 2021			43.907	2_15113892	4.47	21.8	3.91 – 4.47 – 3.76	41.769 – 48.203
	Guerrieri-Rizzardi 2021			43.907	2_15113892	4.81	21.7	3.81 – 4.81 – 3.93	40.34 – 48.203
	Gozzo Medium 2021	**P2**	LG2	37.44	2_14702692	11.15	45.7	10.29 – 11.15 – 11.15	36.022 – 37.44
	Guerrieri-Rizzardi 2021			37.44	2_14702692	7.4	31.2	7.15 – 7.4 – 6.83	36.022 – 39.578
	Gozzo Medium 2020			36.022	2_14208969	5.77	32.5	5.17 – 5.77 – 5.55	34.593 – 37.44
	Gozzo Medium 2021		LG6	53.382	6_19044643	2.13	11	1.9 – 2.13 – 1.95	31.961 – 56.24
	Guerrieri-Rizzardi 2021			51.964	6_18357854	3.2	15	2.95 – 3.2 – 2.29	51.254 – 56.949
	Gozzo Medium 2020		LG17	56.402	17_13926824	2.61	16.7	1.84 – 2.61 – 2.48	44.294 – 57.821
	Gozzo Medium 2021			43.584	17_9990564	1.88	9,8	1.78 – 1.88 – 1.78	41.426 – 57.821
	Guerrieri-Rizzardi 2021			57.821	17_16780652	2.58	12.2	2.35 – 2.58	54.264 – 57.821
Derivative maturity	-	-	-	-	-		-	-	-
Polyphenols	*Gozzo Medium 2020*	**CP**	LG3	39.108	3_9524678	3.29	33.6	3.29 - 3.13	39.108 - 45.339
	Gozzo Medium 2020	**P1**	LG3	36.429	3_9706422	3.25	33.3	2.33 - 3.25 - 2.25	32.111 - 44.729
	Gozzo Medium 2021			36.429	3_9706422	3.87	29.5	3.87 - 3.38	36.429 - 44.729
Anthocyanin	*Gozzo Medium 2020*	**CP**	LG2	35.678	2_9135820	3.83	37.9	3.56 - 3.83 - 3.43	35.321 - 39.964
	Gozzo Medium 2021			36.392	2_9223164	4.79	35.1	4.03 - 4.79 - 3.92	33.184 - 40.319

4.7.2 Multiple QTL Mapping (MQM)

QTL mapping was completed using this analysis method which has a greater resolving power and is statistically more precise than IM. During this analysis, in fact, it is possible to use cofactors selected automatically or defined by the user, to identify chromosomal regions associated with traits of interest. This allows to separate two or more QTL regions present in the same linkage group with greater precision.

The MQM analysis was conducted using only parental maps, P1 and P2 (as described in *paragraph 3.8.4*, page 40). Beside the maps, their genetic datasets and a matrix of quantitative data, that reports only means of phenotypic values between the left and right copies of Gozzo's population and phenotypic data of Guerrieri-Rizzardi's copy for each trait and vintage, were used here.

In **table 12** are reported and commented only reproducible QTLs, i.e., those present in at least two of the three datasets (Gozzo 2020, Gozzo 2021, Guerrieri-Rizzardi 2021). The relative confidence intervals are given; the upstream and downstream marker of LOD peak and immediately adjacent to the extremes of LOD +/- 1 interval was identified. Regarding flowering and budbreak-flowering interval, no overlapping and reproducible confidence intervals were found.

Once the list of reproducible QTLs was selected, the physical positions of the confidence intervals were recovered for each of these QTLs by anchoring QTLs flanking markers to the PN40024 12X.v2 reference. SNP marker expected positions, when available, were recovered from literature (Laucou et al., 2018; Vervalle et al., in press). These were indeed previously recovered by using flanking sequences of SNP markers (https://urgi.versailles.inra.fr/Species/Vitis/GrapeReSeq_Illumina_20K) as queries for BLASTN searches to identify markers theoretical position on the reference genome. Furthermore, the position was also recently further validated through mapping in three segregating population and markers with incongruent positions were discarded from the analysis. Incorrect assignment might be related to short-repeated regions and consequent BLAST issues; the problem in attributing the correct allocation by BLASTN mainly affected SNP probes originating from species other than *Vitis vinifera*.

In this way, it was possible to obtain genomic windows within which to search for genes with interesting GOs and possibly involved in phenological traits and quality. The complete results including all regions identified by the MQM analysis, including non-reproducible ones, are shown in the *Appendix* in tables 37-46, pages 170-196.

Starting from **budbreak**, peaks reproducible only from the P2 map have been identified; on LG19, two reproducible peaks were identified at the same position of 52.497 cM both for Gozzo 2020 and 2021. Physical confidence interval was calculated and covers a region from 18224260 to 20698092 bp of chromosome 19.

For **veraison**, from P1 map reproducible regions containing QTL have been identified on LG4 and LG15. LG4 contains two overlapping regions shared between Gozzo 2020 and 2021 datasets at, respectively, 41.311 (LOD max 7.52) and 42.74 cM (LOD max 5.83). The physical confidence interval covers altogether a region from 20216896 to 20839556 bp of the chromosome 4. The second region on the same chromosome is located at 56.348/58.506 cM shared between Gozzo 2020-2021 datasets covering a region of chromosome 4 from 22594795 to 23610477 bp.
On LG15, just one region shared between Gozzo 2020 and Guerrieri-Rizzardi 2021 dataset was identified, also sharing the same maker (15_11426252) at LOD max (5,65-1.97). In this case, identified QTL region covers a portion of genome from 11043258 to 11702515 bp. From map P2, QTL regions were found reproducible on LG10 at 51.07 cM (li_10_5830295 marker), shared between Gozzo's vintages from 9168618 to 9898096 bp and on LG19 at 5.006 cM (19_1792021 marker) from 1722935 to 2039032 bp, also shared from Gozzo's vintages.

MQM results from P2 map for **flowering-veraison interval** have been highlighted numerous regions on linkage groups 4, 5, 7, 9, 10 and 12. On LG4 there is a peak at 0 cM shared by datasets of year 2021 and marked by the same 4_359553 marker. This QTL region covers from 359553 to 626979 bp of chromosome 4.

Interestingly, a region shared by all dataset is located on LG5, with peak at 11.410 cM and which confirms the marker *mu_5_3513954* as associated and reproducible by both MQM analysis and previous IM analysis on veraison (table 4.7.2aIn addition, again on LG5, a region at 17.114 cM is also identified, similarly reproducible with MQM and IM results on veraison, confirming the data.
Two regions have been identified on LG7: one at 67.936 cM and one at 79.541 cM, both perfectly reproducible between populations for 2021 vintage.
On LG9 there are two reproducible regions for the 2021 vintage: one at 0 cM and one at 5.684 cM, covering overall physical confidence intervals.
Finally, four regions have been identified on LG10 and one on LG 12, all reproducible for the 2021 vintage.

For **budbreak-veraison interval** just one region was identified from P1 map on LG6 at 39.044 cM and shared by populations for 2021 vintage. Peak marker associated with this QTL is represented by 6_8154773 marker.

Reproducible QTLs for **derivative maturity** were found from both parental P1 and P2 parental maps. On P1, it was possible to identify three reproducible regions between at least two of the three vintages analyzed: one on LG 11 at 9.971 cM, one on LG 15 at 40.260 cM and one on LG18 at 22.203, which also share same marker in the first and second case. From the parental map P2, on the other hand, only one region on chromosome 17 was identified. The region is shared by two Gozzo vintages at 8.626 cM.

As for the **color of berry**, the presence of a region on LG2 and shared from both parental maps and for each map by the 3 datasets with high LOD values is in line with literature reports about the color locus of grapevine (Walker et al., 2007; Kobayashi et al., 2004; Fournier-Level et al., 2009). This confirms the goodness of the MQM approach. Furthermore, a region on LG17 was also identified for the P2 map at 27.149 cM.

The analysis on **polyphenol content** highlighted from the map P1 a region at 0.000 cM on LG 11 that shares the same marker *11_391566* in the two Gozzo datasets, covering a physical region from 391566 to 494510 bp.

To conclude, the analysis on **anthocyanin content** highlighted from map P1 an overlapping region on LG 8 at 37.616/38.326 cM shared by the Gozzo datasets.

Table 12 – List of reproducible QTL regions found by the MQM analysis. Markers of upstream (**orange**) and downstream (**green**) extremes of each confidence interval are reported. In **grey** are reported markers associated for each peak. In **red**, position of extreme markers used to search for genomic windows on reference genome PN40024 12X.v2.

			Peak position					*Confidence Interval*							
										Upstream marker		*Peak marker*		*Downstream marker*	
Trait	***Dataset***	***Map***	***LG***	***Position (cM)***	***Marker***	***LOD score***	***% var***	***Position (cM)***	***LOD***	***Marker ID***	***12X.v2 position***	***Marker ID (v0)***	***12X.v2 position***	***Marker ID***	***12X.v2 position***
Budbreak	*Gozzo 2020*	P2	19	52.497	19_18762848	2.42	4.1	49.597-52.497	2.15-2.42	19_15492086	18224260	19_18762848	19436662	19_20024278	20698092
	Gozzo 2021			52.497	19_18762848	3.98	8.8	49.597-52.497	3.76-3.98	19_15492086	18224260	19_18762848	19436662	19_20024278	20698092
Veraison	*Gozzo 2020*	P1	4	41.311	4_20165673	7.52	12.8	39.882-41.311	4.66-7.52	4_19850210	20216896	4_20165673	20532359	Un_24914324	20789611
	Gozzo 2021			42.74	Un_24914324	5.83	16.2	41.311-42.74	4.76-5.83	4_20315085	20681771	Un_24914324	20789611	Un_24964269	20839556
	Gozzo 2021			56.348	4_21965405	6.39	18.2	56.348	6.39	4_21750855	22594795	4_21965405	22809345	4_22055921	22899861
	Gozzo 2020			58.506	4_22566230	7.01	11.6	56.384-58.506	5.78-7.01	4_22273921	23117861	4_22566230	23410170	4_22766537	23610477
	Gozzo 2020		15	25.072	15_11426252	5.65	8.2	22.914-25.072	2.38-5.65	SNP1371_290	11043258	15_11426252	11426252	15_11702515	11702515
	G-R 2021			25.072	15_11426252	1.97	3.2	23.914-25.072	1.65-1.97	SNP1371_290	11043258	15_11426252	11426252	15_11702515	11702515
	Gozzo 2020	P2	10	51.07	li_10_5830295	6.49	16.1	48.924-51.07	5.05-6.49	be_10_5609983	9168618	li_10_5830295	22594795	10_6339461	9898096
	Gozzo 2021			51.07	li_10_5830295	5.67	10.5	48.924-51.07	3.5-5.67	be_10_5609983	9168618	li_10_5830295	22594795	10_6339461	9898096
	Gozzo 2020		19	5.006	19_1792021	3.68	9.2	5.006-7.006	3.68-3.13	19_1722935	1722935	19_1792021	1792021	19_2039032	2039032
	Gozzo 2021			5.006	19_1792021	2.33	4.9	5.006-6.006	2.33-2.9	19_1722935	1722935	19_1792021	1792021	19_2039032	2039032

Interval Flowering-veraison	*Gozzo 2021*	P2	4	0	4_359553	3.7	6.6	0-2	3.7-3.27	-	-	4_359553	359553	4_626979	626979
	G-R 2021			0	4_359553	5.51	11	0-2	5.51-4.79	-	-	4_359553	359553	4_626979	626979
	Gozzo 2020		5	11.41	mu_5_3513954	2.15	7.8	10.7-11.41	1.91-2.15	5_3495583	3495583	mu_5_3513954	3513964	5_3707334	3707334
	Gozzo 2021			11.41	mu_5_3513954	3.04	5.2	10.7-11.41	2.21-3.04	5_3495583	3495583	mu_5_3513954	3513964	5_3707334	3707334
	G-R 2021			11.41	mu_5_3513954	5.2	10.4	10.7-11.41	4.12-5.2	5_3495583	3495583	mu_5_3513954	3513964	5_3707334	3707334
	Gozzo 2021		7	17.114	5_4902985	1.87	3.5	17.114	1.87	5_4865007	4865007	5_4902985	4902985	5_4960783	4960783
	G-R 2021			17.114	5_4902985	4.98	9.9	17.114	4.98	5_4865007	4865007	5_4902985	4902985	5_4960783	4960783
	Gozzo 2021			67.936	Un_13943695	2.25	2.8	67.036-67.936	1.92-2.25	Un_23447761	19517011	Un_13943695	19579597	Un_22783551	19910258
	G-R 2021			67.936	Un_13943695	2.53	4.7	66.036-67.936	1.95-2.53	Un_23447761	19517011	Un_13943695	19579597	Un_22783551	19910258
	Gozzo 2021			79.451	lb_7_15393500	5.59	9.4	77.552-79.451	4.28-5.59	Un_34126916	21379229	lb_7_15393500	21701938	7_15456080	21764518
	G-R 2021			79.451	lb_7_15393500	4.59	8.9	77.552-79.451	3.3-4.59	Un_34126916	21379229	lb_7_15393500	21701938	7_15456080	21764518
	Gozzo 2021		9	0	9_137176	4.2	6.7	0	4.2	-	-	9_137176	137176	9_159487	159487
	G-R 2021			0	9_137176	4.89	9.7	0	4.89	-	-	9_137176	137176	9_159487	159487
	Gozzo 2021			5.684	9_2912792	5.59	9.1	4.975-5.684	4.24-5.59	9_2561793	2561793	9_2912792	2912792	9_2946803	2946803
	G-R 2021			5.684	9_2912792	6.24	12.8	4.975-5.684	4.95-6.24	9_2561793	2561793	9_2912792	2912792	9_2946803	2946803
	Gozzo 2021		10	15.04	Un_26075992	5.29	9.7	12.389-15.04-17.04	3.04-5.29-5.02	Un_30186356	15887454	Un_26075992	12709257	Un_33535896	4434801
	G-R 2021			15.04	Un_26075992	2.87	5.4	13.389-15.04-17.04	1.91-2.87-2.67	Un_30186356	15887454	Un_26075992	12709257	Un_33535896	4434801
	Gozzo 2021			38.2	10_2730976	5.57	9.8	37.491-38.2-39.2	5.57-5.05	10_2044634	5603269	10_2730976	6289611	10_2892377	6451012
	G-R 2021			38.2	10_2730976	3.57	6.9	37.491-38.2-39.02	3.57-3.06	10_2044634	5603269	10_2730976	6289611	10_2892377	6451012
	Gozzo 2021			44.594	10_5115610	4.07	7.2	43.884-44.594	3.17-4.07	SNP397_331	8585271	10_5115610	8674245	10_5293732	8852367
	G-R 2021			44.594	10_5115610	3.38	6.5	43.884-44.594	2.81-3.38	SNP397_331	8585271	10_5115610	8674245	10_5293732	8852367
	Gozzo 2021		12	74.443	10_10642694	4.23	7.3	70.792-74.443	3.01-4.23	10_10477108	14035743	10_10642694	14201329	10_10755483	14314118
	G-R 2021			74.443	10_10642694	3.16	6.1	70.792-74.443	2.48-3.16	10_10477108	14035743	10_10642694	14201329	10_10755483	14314118
	Gozzo 2021			16.479	12_3847351	2.2	3.5	15.77-16.479	2.2	12_3812648	3812648	12_3847351	3847351	12_3875906	3875906
	G-R 2021			16.479	12_3847351	2.99	5.6	15.77-16.479	2.99	12_3812648	3812648	12_3847351	3847351	12_3875906	3875906

Derivative maturity	*Gozzo 2021*	P1	11	9,971	11_1811149	8,27	23,1	9.971-10.971	8.27	11_1719115	1719115	11_1811149	1811149	11_1991860	1991860
	G-R 2021			9,971	11_1811149	7,69	15,1	9.971-10.971	7-69-7.68	11_1719115	1719115	11_1811149	1811149	11_1991860	1991860
	Gozzo 2020		15	40,26	15_13737462	3,56	6,1	39.832-40.26-41.26	3.25-3.56-3.05	15_13467733	13467733	15_13737462	13737462	15_13788826	13788826
	G-R 2021			40,26	15_13737462	2,68	4,7	39.832-40.26-41.26	2.56-2.68-2.24	15_13467733	13467733	15_13737462	13737462	15_13788826	13788826
	Gozzo 2020		18	22,203	18_4381193	6,5	12	22.203	6.5	SNP1001_250	4140486	18_4381193	4381193	18_4500969	4500969
	Gozzo 2021			22,203	18_4381193	2,06	4,8	20.774-22.203	1.95-2.06	SNP1001_250	4140486	18_4381193	4381193	18_4500969	4500969
	Gozzo 2020	P2	17	8,626	17_1572984	6,15	8,5	4.975-8.626	5.31-6.15	17_1343464	1475092	17_1572984	1704612	17_1842832	1974460
	Gozzo 2021			8,626	17_1572984	12,54	21,8	8.626	12.54	17_1343464	1475092	17_1572984	1704612	17_1842832	1974460
Berry color	*Gozzo 2021*	P1	2	43.907	2_15113892	4.47	21.8	43.197-43.907	4.14-4,47	2_10292647	10404646	2_15113892	15225891	2_16280570	16392569
	G-R 2021			43.907	2_15113892	4.81	21.7	43.197-43.907-	4.53-4.81	2_10292647	10404646	2_15113892	15225891	2_16280570	16392569
	Gozzo 2020			36.022	2_14208969	5.96	28.8	34.593-36.73	5.59-5.96	2_13395493	13507492	2_14208969	14320968	2_14315554	14427553
	Gozzo 2021		2	37.440	2_14702692	13.88	45.4	36.02-38.440	12.68 – 13.88	2_14208969	14320968	2_14702692	14814691	2_15113892	15225891
	G-R 2021	P2		37.440	2_14702692	8.83	29.8	36.731-37.440	8.35-8.83	2_14208969	14320968	2_14702692	14814691	2_15113892	15225891
	Gozzo 2021		17	27.149	17_6320545	2.30	4.5	26.720-27.149	2.16	17_5959988	6091616	17_6320545	6452173	17_6607716	6739344
	G-R 2021			27.149	17_6320545	2.81	4.2	25.720-27.149	2.36-2.81	17_5959988	6091616	17_6320545	6452173	17_6607716	6739344
Polyphenol content	*Gozzo 2020*	P1	11	0.000	11_391566	6.59	12.5	0.000	6.59	-	-	11_391566	391566	11_494510	494510
	Gozzo 2021			0	11_391566	6.09	6.8	0	6.09	-	-	11_391566	391566	11_494510	494510
Anthocyanin content	*Gozzo 2021*	P1	8	37.616	8_12974771	7.51	0.8	35.639-37.616	7.19-7.51	8_12046749	12046749	8_12974771	12974771	8_13485811	13485811
	Gozzo 2020			38.326	8_13093399	7.75	44.9	37.616-38.326	7.75	8_12974771	12974771	8_13093399	13093399	8_13690958	13690958

4.7.3 Gene ontologies and selection of candidate regions

Recovered regions were then mined to search for interesting candidates putatively involved in the regulation of the investigated traits. The research produced an extensive list of genes for each trait and for each interval (Supplementary table 1). A selection was made on these candidates based on the gene ontology (GO) of each gene that may have interesting and compatible functions with the processes under consideration. In details, we decided to select candidates belonging to the following GOs: “transcription factors”, “carbohydrate metabolic process”, “response to hormone stimulus”, “developmental processes”, “response to stress” and “signal transduction”.

Table 13, shown below, lists the candidates identified for each of the reproducible QTLs identified from the MQM analysis.

Table 13 - List of candidate genes with interesting functional annotation and gene ontology included in reproducible QTL regions.

Trait	*Chr*	*Gene ID*	*Start position*	*End position*	*Functional annotation*	*Gene ontology*
Budbreak	**19**	VIT_19s0140g00120	*18229921*	*18231721*	gibberellin 2-oxidase	Response to Hormone Stimulus
		VIT_19s0140g00140	*18335488*	*18337040*	gibberellin 2-oxidase	Response to Hormone Stimulus
		VIT_19s0027g00460	*19983415*	*19984468*	probable glutathione s-transferase	Secondary Metabolic Process
		VIT_19s0027g00230	*19459106*	*19463910*	nac domain protein VvNAC33	Transcription Factor Activity
Veraison	**4**	VIT_04s0023g03650	*20558693*	*20576365*	glycosyl hydrolase family 38 protein	Carbohydrate Metabolic Process
		VIT_04s0023g03320	*20207562*	*20217990*	transducin wd40 domain-containing protein	Cellular Process
		VIT_04s0023g03520	*20447472*	*20448369*	cold shock domain protein 3	Transcription Factor Activity
		VIT_04s0023g03710	*20618667*	*20619631*	myb transcription factor MYB4b	Transcription Factor Activity
		VIT_04s0044g01110	*23299884*	*23302672*	alcohol dehydrogenase 1	Carbohydrate Metabolic Process
		VIT_04s0044g01120	*23307283*	*23314411*	alcohol dehydrogenase 1	Carbohydrate Metabolic Process
		VIT_04s0044g01130	*23319484*	*23321778*	alcohol dehydrogenase 1	Carbohydrate Metabolic Process
		VIT_04s0044g01090	*23255303*	*23280622*	transducin wd40 domain-containing protein	Signal Transduction
		VIT_04s0044g01220	*23550337*	*23557330*	nac domain containing protein 82	Transcription Factor Activity
	15	VIT_15s0021g01140	*11181159*	*11185013*	nucleotide-diphospho-sugar transferase domain-containing protein	Signal Transduction
		VIT_15s0021g01150	*11203877*	*11207272*	nucleotide-diphospho-sugar transferase domain-containing protein	Signal Transduction
		VIT_15s0021g01310	*11499221*	*11503179*	wrky transcription factor 12	Transcription Factor Activity
	10	VIT_10s0003g03490	18543927	18544228	gibberellin 2-oxidase	Response to Hormone Stimulus
		VIT_10s0003g03730	19496032	19477575	f-box kelch-repeat protein	binding
		VIT_10s0003g03710	19428061	19436718	cullin 2	binding
	19	VIT_19s0014g01680	18441617	18442075	transcription factor tcp3-like	Transcription Factor Activity

		VIT_19s0014g01780	18803237	18804689	transcription factor rf2b	Transcription Factor Activity
		VIT_19s0014g01840	19005891	19013021	protein argonaute 1-like	Transcription Factor Activity
		VIT_19s0014g01850	19028850	19029563	c3hc4-type ring finger protein	Transcription Factor Activity
Interval Flowering-Veraison	**4**	VIT_04s0008g00520	*458271*	*466257*	6-phosphofructokinase 3-like	Carbohydrate Metabolic Process
		VIT_04s0008g00660	*595518*	*601744*	protein early flowering 3-like VvELF3_1	Developmental Process
		VIT_04s0008g00620	*550884*	*552122*	myb-related transcription factor lbm2-like	Transcription Factor Activity
	5	VIT_05s0020g01930	*3626919*	*3630035*	tubby-like f-box protein 3	Transcription Factor Activity
		VIT_05s0020g03140	*4868741*	*4873368*	hexose transporter	Carbohydrate Metabolic Process
	7	VIT_07s0129g00030	*21633505*	*21635150*	gras family transcription factor	Transcription Factor Activity
	9	VIT_09s0002g02980	*2679371*	*2685353*	brassinosteroid insensitive 1-associated receptor kinase 1	Response to Stress
		VIT_09s0002g03050	2799091	2799551	gdsl esterase lipase apg	Lipid Metabolic Process
	10	VIT_10s0003g00990	*5912736*	*5922763*	pyruvate decarboxylase	Carbohydrate Metabolic Process
		VIT_10s0003g01380	*6288268*	*6290668*	myb family transcription factor apl	Transcription Factor Activity
		VIT_10s0003g05740	*14314636*	*14317876*	probable wrky transcription factor 14-like	Transcription Factor Activity
Interval Budbreak-Verasion	**6**	VIT_06s0004g06920	*7629694*	*7641875*	fructose-6-phosphate 2-kinase fructose- -bisphosphatase	Carbohydrate Metabolic Process
		VIT_06s0004g07020	*7765480*	*7771216*	ethylene-overproduction protein 1-like	Carbohydrate Metabolic Process
		VIT_06s0004g06710	*7418557*	*7422349*	zinc finger ccch domain-containing protein	Carbohydrate Metabolic Process
Derivative maturity	**11**	VIT_11s0016g02200	*1800943*	*1807928*	beta-galactosidase	Response to Hormone Stimulus
		VIT_11s0016g02410	*1948504*	*1950635*	transcription factor myb1r1	Transcription Factor Activity
		VIT_11s0016g02140	*1758493*	*1761259*	ap2 erf domain-containing transcription factor	Carbohydrate Metabolic Process
	15	VIT_15s0021g02700	*13735914*	*13737731*	expansin b protein	Cell Wall Metabolism
	18	VIT_18s0001g05300	*4254678*	*4257480*	trehalose-phosphate phosphatase-like	Carbohydrate Metabolic Process
		VIT_18s0001g05570	*4363673*	*4368991*	sugar transport protein 5-like	Carbohydrate Metabolic Process

		VIT_18s0001g05250	*4220268*	*4222313*	ap2 erf domain-containing transcription factor	Response to Hormone Stimulus
	17	VIT_17s0000g01820	*1483421*	*1486575*	malate synthase	Carbohydrate Metabolic Process
		VIT_17s0000g01860	*1550198*	*1551818*	gaga-binding transcriptional activator	Transcription Factor Activity
Berry color	**2**	VIT_02s0033g00370	*14275372*	*14275752*	myb-related transcription factor	Transcription Factor Activity
		VIT_02s0033g00380	*14291438*	*14291683*	transcription factor myba1	Transcription Factor Activity
		VIT_02s0033g00390	*14291727*	*14292732*	transcription factor myba3	Transcription Factor Activity
		VIT_02s0033g00400	*14302187*	*14303282*	transcription factor myba3	Transcription Factor Activity
		VIT_02s0033g00410	*14351791*	*14352807*	transcription factor myba3	Transcription Factor Activity
		VIT_02s0033g00420	*14386896*	*14387181*	40s ribosomal protein s28-like	DNA/RNA Metabolic Process
		VIT_02s0033g00430	*14392775*	*14393842*	transcription factor myba3	Transcription Factor Activity
		VIT_02s0033g00440	*14400037*	*14401092*	transcription factor myba3	Transcription Factor Activity
		VIT_02s0033g00450	*14420525*	*14421283*	transcription factor myba3	Transcription Factor Activity
		VIT_02s0033g00460	*14425416*	*14426478*	transcription factor myba3	Transcription Factor Activity
		VIT_02s0033g00960	*15974263*	*15978140*	nuclear-interacting partner of alk	DNA/RNA Metabolic Process
		VIT_02s0033g01000	16093114	16095441	benzyl alcohol o-benzoyltransferase	Cellular Metabolic Process
		VIT_02s0033g01020	16190948	16193334	benzyl alcohol o-benzoyltransferase	Cellular Metabolic Process
		VIT_02s0033g01030	16252684	16254941	benzyl alcohol o-benzoyltransferase	Cellular Metabolic Process
		VIT_02s0033g01050	16299329	16301080	benzyl alcohol o-benzoyltransferase	Cellular Metabolic Process
		VIT_02s0033g01060	16349045	16351278	benzyl alcohol o-benzoyltransferase	Cellular Metabolic Process
		VIT_02s0033g01070	16356696	16359431	benzyl alcohol o-benzoyltransferase	Cellular Metabolic Process
	17	VIT_17s0000g05520	*6144571*	*6153251*	calcium-dependent protein kinase 30	Signal Transduction
		VIT_17s0000g05810	*6447796*	*6451662*	wrky transcription VvWRKY53	Transcription Factor Activity
Polyphenol content	**11**	VIT_11s0016g04080	*3340651*	*3341251*	ethylene-responsive transciptional coactivator-like protein	Response to Hormone Stimulus
		VIT_11s0016g05340	*4677731*	*4680660*	ethylene-responsive transcription factor win1-like	Response to Hormone Stimulus

		VIT_11s0016g03940	*3224068*	*3225265*	heat stress transcription factor c-1	Transcription Factor Activity
		VIT_11s0016g04280	*3571380*	*3572024*	zinc finger ccch domain-containing protein 14-like	Transcription Factor Activity
		VIT_11s0016g04600	*3913885*	*3916168*	myb family transcription factor	Transcription Factor Activity
Anthocyanin content	**8**	VIT_08s0040g01730	*12843518*	*12852631*	ethylene insensitive	Response to Hormone Stimulus
		VIT_08s0040g01810	*12924098*	*12929097*	auxin response factor 18-like	Response to Hormone Stimulus
		VIT_08s0040g02110	*13224257*	*13225775*	nac domain ipr003441	Transcription Factor Activity

5. Discussion 1

In this study a QTL mapping analysis was conducted on a segregating population obtained from the cross between the local cultivar Corvina and the worldwide cultivated variety Cabernet Sauvignon with the main aim of dissecting the genetic regulation of phenology and quality traits in grapevine

The selection of segregating individuals by genotyping with SSR markers successfully identified 142 genotypes confirmed as cross and on which genetic and phenotypic data on different vintages were collected.
Hybridization of genomic DNA of these individuals to *Vitis*18kSNP chip allowed to obtain the large number of markers needed to build a high-resolution genetic map to be used for QTL mapping.
The *Vitis*18KSNP chip showed to be a useful tool for the high-throughput genotyping of 18071 SNPs. It selectively genotypes SNPs distributed throughout the grapevine genome (Le Paslier et al., 2013). On the technical side successful genotyping rate, i.e., the percentage of retained SNPs after removal of low quality and no-call loci, was high in our analysis (91.2%) compared to previous work. Indeed, the observed score was close to the upper limit of the range 55.6-92.7% observed in previous studies applying similar thresholds for SNP filtering when using the *Vitis*18K SNP chip (De Lorenzis et al., 2015, 2017, 2019; Mercati et al., 2016; Ruffa et al., 2016; Sunseri et al., 2018; Bianchi et al., 2020). This can probably be attributed to the large set of genotypes that was used in our case to classify SNPs, since we used altogether data from three independent segregating population for extraction by Genome studio (Vervalle et al., in press). Moreover, ASSiST takes into account the pedigree of each mapping population, thus the progeny genotypes are expected, while the possible genotypes of germplasm accessions are not known *a priori*.
Genotyping reproducibility was also comparable to that found in other studies employing the same array (99.8% for our Cabernet Sauvignon genotype control genotype hybridized twice compared to 99-100% in Houel et al., 2015; Mercati et

al., 2016; De Lorenzis et al., 2019. This very low error rate in the positive detection of SNPs is an evident advantage of this tool when compared to highly multiplexed, shallow sequencing strategies like GBS (Genotyping-By-Sequencing) and RAD (Restriction Site-Associated DNA) sequencing that result in a high proportion of genotyping error, as well as missing data and under-calling of heterozygous sites (Mardis 2008; Barba et al., 2014; Hyma et al., 2015; Marrano et al., 2018).
Despite the successful genotyping rate and reproducibility, genotyping failure was observed for 1586 SNPs. The *Vitis*18K SNP array consists of approximately 25% loci identified from different *Vitis* species (Le Paslier et al., 2013), which can account for the low quality, or the lack of hybridization, for a number of SNP loci (De Lorenzis et al., 2015; Mercati et al., 2016; Sunseri et al., 2018; De Lorenzis et al., 2019). Accordingly, the fraction of non-*vinifera* SNPs loci among SNPs failing to genotype was abundant (832 SNPs corresponding to more than 50% of failed ones), in agreement with the well-recognized low marker transferability in *Vitis* (Zou et al., 2020).

Genotypic data were successfully used then to build a high-resolution genetic map. Our analysis demonstrate that data provided by the *Vitis*18KSNP array can be successfully used for the building of a genetic map. This SNP chip has been mostly used to assess population genetic diversity, structure, and relatedness (De Lorenzis et al., 2015; Degu et al., 2015; Mercati et al., 2016; Ruffa et al., 2016; Laucou et al., 2018; Marrano et al., 2018; Sunseri et al., 2018; De Lorenzis et al., 2019; Bianchi et al., 2020; Boccacci et al., 2020; Crespan et al., 2020; Raimondi et al., 2020; Crespan et al., 2021; D'Onofrio et al., 2021; Mercati et al., 2021), or to identify somatic variants to explain intra-varietal diversity (De Lorenzis et al., 2017).
The application of the *Vitis*18KSNP array to build high resolution genetic maps and identify marker-trait associations was much more limited so far (Houel et al., 2015; Duchêne et al., 2020; Mamani et al., 2021). In detail, Houel et al. (2015) built two framework parental genetic maps including respectively 714 SNPs and 408 SNPs starting from about 6000 genotyped SNPs. The maps lengths were 1301 and 606 cM respectively with a much shorter Picovine map due to its high homozygosity.

Duchêne et al. also built two parental maps including respectively 953 and 806 SNPs (total length of 1413 e 1164 cM) starting from 6837 markers segregating in a parental fashion. Finally, Mamani et al. also built a consensus map including 1717 total SNPs markers including also *hk* x *hk* markers similarly as done in our work. However, the map we present here includes more markers even compared to this last map, with a final number of 3467 SNPs and a length of 1284 cM. Finally, the successful mapping of QTLs conducted in our study and these previous studies confirms that the *Vitis*18KSNP chip can be successfully used for the development of biparental maps, including consensus maps and for the mappings of QTLs in both parental and consensus maps.

All progenies were phenotyped for phenology and quality related traits. Except for berry color, as expected, all traits showed a normal trend, suggesting a complex genetic architecture. While a divergence in color (R2-R3) was maintained between the two parental genotypes, phenology divergence, especially for veraison time, was less obvious. By looking at distribution of the phenology related traits a transgressive behavior was however in general observed in the population. This justified QTL mapping of phenology related traits in this population, despite the non-highly divergent behavior observed in parental lines. For phenotyping replicates of the CsxC population were positioned both in the same location or in locations with different climates and soils. Our data clearly highlighted that phenotypic variation attributable to the interaction with environment seemed to be more pronounced for some of the traits. As an example, comparing data on the content of anthocyanins and polyphenols from the population copy present in the Guerrieri-Rizzardi's vineyard and in Gozzo's vineyard (same vintage 2021) a low correlation was observed, despite a higher overall correlation across replicates in same location or across years. Interestingly the lower reproducibility of phenotypes across location was less pronounced for other traits (e.g., maturity). This phenomenon can be explained by the different climatic characteristics between the two cultivation areas and their effect on some of the phenotypic traits: in fact, Valpolicella area represents the ideal cultivation environment for Corvina cultivar, that is known to shows relevant differences if grown in a different environment. It

would be interesting to further evaluate these aspects in further years, to explore the suitability of this material to eventually dissect the genetic regulation of the interaction with the environment in the different genotypes.
Finally, we should also mention that the correlation indices were sometimes quite low with a variable behavior between vintages. This is likely due to the still young age of the plants which are therefore not yet fully stable phenotypically Indeed such low reproducibility was especially observed in 2020 vintage. The integration of these datasets with data collected in future years is expected to increase the statistical relevance of the mapped QTLs.

Regarding the identification of QTLs, reproducible genomic regions between at least two datasets and with LOD significance values above threshold have been identified for almost every trait, with the only exceptions of flowering time and budding-flowering interval. These reproducible QTLs match to genomic regions previously identified in further studies as involved in the traits, providing solidity to our study.
For budbreak a QTL on the lower part of LG 19 was indeed previously reported in Duchêne et al. (2012) explaining up to 18% of the observed variance. Similarly, we identified a reproducible QTL for veraison time on chromosome 15 between 11043258bp and 117025151bp. Interestingly in the same region a QTL for veraison time was previously reported by Grzeskowiak et al. (2013).
Despite these confirmed QTLs, this study also identified a new reproducible QTL for veraison time (on linkage group 4) and several new QTLs for Flowering – Veraison interval (on linkage groups 4, 5, 7, 9, 10, 12 respectively) and for budbreak-veraison interval (on linkage groups 6).
For QTL related to quality traits the strongest QTL was that related to berry color on linkage group 2. This was confirmed for all dataset with LOD values ranging from 4.47 to 13.88 and the interval centered on the *MybA1* gene that is well known to control the largest part of the quantitative variation in berry color. This demonstrates that our approach, including the production of a high-resolution genetic map and the QTL mapping procedure, is suitable for successfully mapping QTLs for the enquired traits. Furthermore, a reproducible QTL for berry color on

LG 17 was also found, in agreement with previous literature reports (Costantini et al., 2015 and Sun et al., 2020).
Finally reproducible QTLs for the anthocyanins and polyphenols content as well as for the sugar content at maturity were also retrieved, also in this case with some in agreement with previous findings (Costantini et al., 2015 for example for anthocyanins and polyphenols), that can be now further explored and fine mapped.

The identification step of putative candidate genes was carried out considering the Gene Ontology (GO) of the various candidates included in the regions after anchoring to the assembly as described in *paragraph 4.7.2* (page 76). Among the various GO of biological functions greater attention was paid to candidates with GO belonging to transcription factors (TF), signal transduction, signaling, development and carbohydrate metabolism classes.
For budbreak trait, several genes involved in the hormonal response have been identified on chromosome 19: VIT_19s0140g00120 and VIT_19s0140g00140 encode a gibberellin 2-oxidase involved in GA's biosynthesis, with a potential role in budbreak regulation. Several transcription factors potentially related to the budbreak were also identified; among these, the most interesting is represented by VIT_19s0027g00230 which encodes the VvNAC33 factor. This TF, as reported by Sun et al. (2012), appears to be regulated by the miR64 microRNA and participates not only in foliar development, but also in abiotic stress response. Furthermore, its interactions with the transcription factor *MybA3*, which plays a key role in berry ripening, was shown. It represents a key regulator in transcriptome remodeling during plant development, including maturation, and was recently functionally characterized in this contest (Palumbo et al., 2014; Wang et al., 2013; Fasoli et al., 2012; D'Incà et al., 2021). A putative role in budbreak was however not explored so far. Finally, and most interestingly, a gene encoding for a glutathione s-transferase (VIT_19s0027g00460) was also identified among the candidates that was previously highlighted by Duchêne et al. (2012) where a budbreak QTL in the same region was found. Indeed, an induction of this gene was previously reported in association to budbreak and with the oxidative stress occurring in the dormancy release process in other species (Ophir et al., 2009).

For veraison trait, in the region identified on chromosomes 4are located two transcription factors, the *VvMYB4b* and the NAC factor 82 (VIT_04s0044g01220). *VvMYB4b* (VIT_04s0023g03710) is orthologs to a gene which constitutive expression in petunia represses phenylpropanoid biosynthetic genes and selectively reduces the amount of small-weight phenolic compounds that are associated to the berry ripening (Cavallini et al., 2015). Furthermore, VIT_04s0023g03520, also included among candidates for the same QTL, encodes an UV up-regulated cold shock domain protein 3 that regulate the expression of several phenylpropanoid biosynthesis-related genes. (Carbonell-Bejerano et al., 2014). Two genes coding for a transducin WD40-containing motif (VIT_04s0044g01090, VIT_04s0023g03320), the main components of a regulatory complex (MBW complex) were also found in the same confidence interval. This protein is particularly interesting as the WD40 domain establishes combinatorial interaction with *Myb* factors that, among other functions, are also involved in determining the set of anthocyanin or proanthocyanidin (PA) genes to be expressed (Winkel-Shirley, 2001; Koes et al., 2005; Ramsay and Glover, 2005). As regards carbohydrate metabolism, several interesting genes are present in this region: VIT_04s0023g03650 encodes a galactose mutarotase-like protein, orthologue of At3g01260 gene in *Arabidopsis thaliana*, involved in galactose catabolic process via UDP-galactose. There are also 3 genes (VIT_04s0044g01110, VIT_04s0044g01120, VIT_04s0044g01130) encoding an alcohol dehydrogenase 1 involved in the production of volatile organic compounds (VOCs) (Ma et al., 2021) that are the main determinants of wine and organoleptic properties in grape flavor (Rambla et al., 2016) under control also by environmental factors (Black et al., 2015; Belda et al., 2017).

On chromosome 15, on the other hand, a WRKY transcription factor (VIT_15s0021g01310) has been identified. This gene was previously characterized by Wang et al., (2014) in the regulation of various physiological processes, including the response to abiotic stress. Two genes (VIT_15s0021g01140, VIT_15s0021g01150) encoding glycosyltransferases involved in the biosynthesis of disaccharides, oligosaccharides and polysaccharides were also included among candidates.

Regarding flowering-veraison interval, several QTLs have been identified on chromosomes 4, 5, 7, 9 and 10. Among interesting genes putatively involved in the trait on chromosome 4 a 6-phosphofructokinase 3-like (VIT_04s0008g00520) the *VvELF3* protein early flowering 3-like (VIT_04s0008g00660) involved in circadian clock signaling (Carbonell-Bejerano et al., 2014) and a *Myb* transcription factor (VIT_04s0008g00620) were included. A tubby-like f-box protein 3 (VIT_05s0020g01930) belonging to a family of transcription factors regulated by ABA (Lai et al., 2004) and a transporter of hexose sugars (VIT_05s0020g03140) have been identified on chromosome 5. On chromosome 7, on the other hand, *VviSHR1* (VIT_07s0129g00030) was identified, belonging to the gras transcription factor family, expressed in several reproductive and vegetative tissues. Concerning this candidate, it was already reported that during ripening, its expression appeared higher (Grimplet et al., 2016). On chromosome 9, in addition to a brassinosteroid kinase receptor (VIT_09s0002g02980), a TF belonging to *Myb* family (VIT_10s0003g01380), a pyruvate decarboxylase (VIT_10s0003g00990) and a gdsl esterase lipase apg (VIT_09s0002g03050), a gene belonging to a class of pointed also by Guo et al. (2019) as candidate for a veraison QTL they found on chromosome 16 have been identified. Finally, two TFs belonging to *Myb* family (VIT_10s0003g01380) and to the WRKY family (VIT_10s0003g05740) have been highlighted on chromosome 10.

For Budbreak-veraison interval, three interesting genes have been identified on chromosome 6: a fructose-6-phosphate-2-kinase (VIT_06s0004g06920), involved in anthocyanin content (Pastore et al., 2011), an ethylene -overproduction protein 1-like (VIT_06s0004g07020) and a Cys-3-His zinc finger TF (VIT_06s0004g06710).

Further QTL regions harboring candidate genes have been identify on chromosomes 11, 15, 17, 18 for derivative maturity. Most interesting candidates here are those involved in sugar metabolism, like a beta-galactosidase (VIT_11s0016g02200), sugar transport protein 5-like encoding tonoplast-localized transporters specific of the berry cells (Hayes et al., 2007; Afoufa-Bastien et al., 2010 - VIT_18s0001g05570), malate synthase (VIT_17s0000g01820). However, TFs including two ap2 erf domain-containing family genes (VIT_18s0001g05250,

VIT_11s0016g02140) and a *Myb1r1* (VIT_11s0016g02410) were also included in these regions.

For berry color, as expected according to literature (Guo et al., 2019; Kobayashi et al., 2004; Fournier-Level et al., 2009; This et al., 2007) we found the cluster of *Myb* genes including *MybA1* and *MybA3* in the identified major QTL on chromosome 2. We also identified seven genes reported also from Guo et al. (2019): nuclear-interacting partner of alk (VIT_02s0033g00960) and many copies of benzyl alcohol o-benzoyltransferase (VIT_02s0033g01000-VIT_02s0033g01070). On chromosome 17 were also found a calcium-dependent protein kinase 30 (VIT_17s0000g05520) and the WRKY TF *VvWRKY53* (VIT_17s0000g05810).

Concerning polyphenols content, five genes have been identified on chromosome 11 QTL confidence interval: two genes responsive to ethylene, ethylene-responsive transcriptional coactivator-like protein (VIT_11s0016g04080) and ethylene-responsive transcription factor win1-like (VIT_11s0016g05340), a heat stress transcription factor c-1 (VIT_11s0016g03940) and two TF belonging to zinc finger CCCH domain-containing family (VIT_11s0016g04280) and *Myb* family (VIT_11s0016g04600).

Finally, for anthocyanin content, three genes have been identified on chromosome 8: an ethylene-insensitive protein 2-like and (VIT_08s0040g01730), an auxin response factor 18-like (VIT_08s0040g01810) and a NAC domain protein (VIT_08s0040g02110).

CHAPTER 2

Validation of association results from a genome-wide association study to target genetic regulators of phenological events in a grapevine collection

2. Abstract

QTL mapping is an excellent tool for the genetic dissection of complex traits but, using a bi-parental population, it provides genomic information regarding only the two cultivars used to produce the segregating population. On the other hand, Genome-Wide Association Study (GWAS) represents a genetic instrument for investigating complex traits that allows an understanding of more extensive genetic architecture of grapevine species. In fact, in the GWAS a large genetic variability panel is usually used, represented by several different cultivars. This allows to obtain not only the identification of regions possibly involved with the phenotype, but also to surpass the cultivar-specific characteristics and to obtain genetic information not limited to the only two varieties used instead in the QTL mapping.

GWAS was performed in our lab during a previous study: starting from a panel of more than 600 cultivars grown in the germplasm collection of CREA-VE in Susegana (Treviso, Italy), a core collection of 132 unique genotypes was pulled out, based on data at 45 SSR markers, representing the minimum number of cultivars covering all the alleles and reaching the genetic diversity of the initial panel. Then, SNPs genetic data obtained through hybridization to the GrapeReSeq *Vitis*18KSNP Chip were retrieved for the whole core collection. Altogether, 9536 SNP markers were selected and managed with phenotypic data, i.e., flowering time, veraison time and interval flowering-veraison collected over 13 years, to perform GWAS. After bioinformatic analysis conducted with 3 different software, 15 SNPs significantly associated to veraison time and to the interval flowering-veraison were found. Among these, 3 SNPs located on chromosomes 12, 16 and 18, were detected as significantly associated to both phenotypic traits and chosen to validate the associations.

In this study, the association between the phenotype of flowering-veraison interval and SNPs found on chromosomes 12, 16 and 18 was verified through Sanger sequencing of genomic regions surrounding the three SNPs. From the initial panel

of 600 cultivars, a subset of 94 individuals showing extreme phenotypes in flowering-veraison interval was created and, for each individual, ad-hoc designed regions were Sanger-sequenced to confirm the presence or absence of each SNP. After a chi-square test on allelic distributions among 94 individuals, each of the 3 SNPs identified was confirmed as significantly associated with phenotype. Potential candidate genes in all the identified regions are discussed.

2. Introduction

In *Chapter 1* introduction, main problems for modern grapevine cultivation and especially the deleterious effects that climate change is having on it were already introduced. As mentioned, the main target of breeders is to obtain new stable cultivars more resistant to biotic and abiotic stress. However, in order to start breeding programs, it is necessary to dissect the genetic determinants that govern the main phenological and qualitative traits and get information about their position in the genome and functions.
Chapter 1 illustrated one of the most used genetic approaches for understanding and identifying the complex quantitative traits that often govern phenotypic traits in *V. vinifera*, QTL mapping. Although this type of approach has repeatedly proved to be a valid tool for the research of genetic key elements (Duchêne et al., 2012; Costantini et al., 2008; Grzeskowiak et al., 2013; Zyprian et al., 2016 as examples), it also has limitations. QTL mapping, in fact, uses a bi-parental segregating population as analysis material, which therefore allows to analyze the genetic variants present in the population but which are inherited only from two parental cultivars, thus preventing the extension of the analysis of phenological traits also to other cultivars and therefore other variants.

At the same time, Genome-Wide Association Study (GWAS) represents an alternative genetic approach in the study of complex traits that allows to know the genetic basis of a given trait using a larger portion of the genetic variability of a species (Lipka et al., 2015). In fact, this approach is applied to a large natural population, usually a collection of germplasm, which allows to analyze many more variants for a phenotypic trait and is thus not limited to causal segregating polymorphisms in parents of a bi-parental population, providing a higher mapping resolution (Huang et al., 2014). The source of variability can be also constituted by a subset of the genotypes present in germplasm collections and composed of individuals carefully selected in order to represent the entire genetic variability of the belonging collection (Emanuelli et al., 2010, 2013, Chitwood et al., 2014, Trenti

et al., 2021). The potential of GWAS has been demonstrated in recent years by various studies, for example to explore the genetic structure and the history of domestication of grapevines using a collection of the United States Department of Agriculture (USDA) (Myles et al., 2011), characterization of genetic diversity (Nicolas et al., 2016; Migicovsky et al., 2017) in germplasm available but also in the dissection of key agronomic traits in other fruit and vegetable crops, such as peach (Cao et al., 2016).

The linkage disequilibrium (LD) represents a non-random association between alleles of two or more loci within a population. The extent of this association is characteristic of each species and can vary depending on the evolutionary and domestication history of the organism. In grapevine, it has been estimated that this phenomenon affects relatively small regions (about 50 kb, Nicolas et al., 2016). Regarding GWAS approaches in *V. vinifera*, the narrowness with which the LD decays requires the use of high-resolution genotyping technologies. Several studies have shown indeed that GWAS mapping of traits involved in domestication can be also accomplished through relatively low-resolution genotyping strategies but rarely allows to resolve regions that govern complex traits (Myles et al., 2011; Migicovsky et al., 2017).

As already mentioned in *Background* paragraph (page 8), the advent of NGS technologies has also brought upgrades to GWAS; in addition to GBS as a tool for high resolution genotyping (Elshire et al., 2011), array-based technology has also been introduced in GWAS approaches. Chitwood et al. (2014) conducted a study by using Illumina SNP chip technology to characterize the genetic basis of leaf shape by genotyping 961 vine accessions through the *Vitis*9kSNP chip. More recently, Laucou et al. (2018) used the most modern version of the SNP chip, *Vitis*18KSNP chip, to analyze the structure and genetic diversity of a large collection of germplasm, also identifying an LD in a range of 28.8 to 58.2 kb related to non-recombinant haplotype blocks, thus also highlighting some limitations in GWAS of complex traits. GWAS studies aiming also to map traits related to phenology were more recently conducted in grapevine, but their success was limited as they could retrieve just few QTLs for such traits mainly located on

chromosome 16 (Guo et al., 2019, Flutre et al., 2022). Alongside many other studies have been conducted through genotyping with *Vitis*18KSNP chip, confirming it as a valid instrument for example for the analysis of population structures (D'Onofrio et al., 2021; Crespan et al., 2020, 2021; De Lorenzis et al., 2015, 2017, 2019).

A GWAS study has been previously conducted in the Laboratory of Genetic Biotechnology in the University of Verona (Pietro Delfino, PhD thesis) aimed at exploring the genetic diversity of a collection of grapevines and searching for genetic associations and identifying regulatory elements of the main phenological processes. The study was conducted starting from a large collection of germplasm located at the CREA-VE in Susegana - Conegliano (TV). The collection, composed from more than 2000 grape accessions, has been phenotyped for over 50 years for the main phenological stages such as: flowering beginning, veraison beginning and ripening.

From the collection, a total of 617 unique molecular genotypes were considered, according to previous genotyping with 45 SSR markers. Genetic diversity and population structure of these were also analyzed by using the software Genalex v.6.5.1 (Peakall and Smouse, 2021) the PCA method and Structure v.2.3.4 and the results validated with DAPC Bayesian approach.

A genome-wide association study was conducted on a subset created from the 617 individuals, a core collection, consisting of a minimum number of genotypes that represented all the variability of the entire. The constitution of the core collection involved the analysis of 530 individuals (the 617 genotypes were filtered for MAF>0.05) and their genetic data at 45 SSR markers. By using the software Core Hunter 3 (De Beukelaer et al., 2018)132 individuals able to represent the entire genetic variability of the collection, were selected for the core.

For the GWAS study, SNP chip profiles for 95 of these 132 individuals already available from a literature dataset coming from the hybridization on *Vitis*18kSNP chip of a French germplasm collection (Vassal collection) (Laucou et al., 2018) were considered. The missing SNP chip profiles for the remaining 37 individuals of CREA-VE collection were obtained by extracting the genomic DNA and

hybridizing them on the same chip at the Sequencing Platform Laboratory, Edmund Mach Foundation (San Michele all'Adige - TN). The association analysis was conducted on three different phenotypic traits: flowering beginning, veraison beginning and flowering to veraison interval. The analysis was conducted in parallel with three software: Emmax (Kang et al., 2010), Gapit (Lipka et al., 2012) and Qtcat (Klasen et al., 2016).

Overall, GWAS analysis conducted identified 15 SNPs markers significantly associated with the three phenotypic traits investigated but, among these, only three were associated with at least two of the three software used. As can be seen from **table 14**, the markers chr12_19533740_A_G (chromosome 12) chr16_348557_T_G (chromosome 16) and chr18_3381571_C_T (chromosome 18) are all associated with the flowering to veraison interval, in addition to the "veraison beginning ". These 3 SNPs for which multiple evidence were found, were therefore chosen for the validation of the association in this work through Sanger sequencing in a panel of genotypes showing extreme phenotypes for the flowering-veraison interval phenotypic trait.

	Trait	Software	Significance	Multiple evidence
Chr3_379409	VB	QTCAT	*	No
Chr4_6106210	F-V	QTCAT	**	No
Chr8_3847595	F-V	Gapit	*	No
Chr8_7340158	F-V	QTCAT	*	No
Chr8_8216855	F-V	Gapit	**	No
Chr11_963022	F-V	QTCAT	*	No
Chr11_1343142	F-V	QTCAT	*	No
Chr12_21100465	VB, F-V	QTCAT, EMMAX	**, *p < 0.1*	Yes
Chr13_26217739	VB	QTCAT	**	No
Chr14_23248717	F-V	QTCAT	**	No
Chr14_29502138	F-V	QTCAT	**	No
Chr16_348557	VB, F-V, F-V	QTCAT, Gapit, QTCAT	**, *, **	Yes
Chr16_20303373	F-V	QTCAT	**	No
Chr18_3381571	VB, F-V	QTCAT, QTCAT	**, **	Yes
Chr18_22124028	F-V	QTCAT	**	No

Table 14 - The table reports the most significantly associated SNPs for the phenological traits averaged across the years. * FDR < 0.05; ** FDR < 0.01 as obtained from the approach under which the association was found. For QTCAT, the concept of FDR does not stand, so the stars are given according to the p-value.

3. Materials and methods

3.1 CREA-VE germplasm collection

The materials used for previous GWAS analysis and for this study are both grapevine plants belonging to the CREA-VE Conegliano's collection. GWAS material is represented by 617 unique genotypes including national and international cultivars, located in the vineyards of the research center for viticulture of Susegana, Conegliano (TV, Italy). These plants have been phenotyped for 50 years for the main phenological stages (beginning of flowering, veraison and ripening). Here we used data of 11 years. In addition, the interval between flowering and veraison was calculated, expressed in Julian days, and used as further phenotypic trait for GWAS analysis. In order to validate association signals obtained from the GWAS study, we focused on a portion of the collection.

The validation was carried out by identifying the members of the collection with the most extreme phenotypes, i.e., those with wider or narrower flowering-veraison interval. The first 47 genotypes with the widest interval (named *high* group, highest number of Julian days) and the first 47 individuals with the narrowest interval (named *low* group, lower number of Julian days) were then selected from the collection for validation.

3.2 SNPs localization and annotation

The annotation of the three significant SNPs obtained from GWAS approach made it possible to identify the genomic localization of each one and to evaluate, if SNP falls into a coding sequence (CDS), the function of the gene and predictive effect of the mutation. This operation was manually conducted by examining the annotation v1 (Vitulo et *al*., 2014) on the assembly 12x.v2, by using physical positions of each SNP (see *Chapter 1*, *paragraph 4.7.2*, page 76). Once position of the SNP was verified, the predictive effect of the mutation was evaluated by PROVEAN software (Choi, 2021).

In order to create amplicons on genomic region of each of the three SNPs identified and subjected to Sanger sequencing, genomic intervals have been defined from which the reference sequence can be extrapolated and on which primer pairs for amplification can be designed. Using the physical position of each SNP, it was possible to create an interval of + 500 / -500 bp from the SNP to obtain the reference sequence from the reference genome *Vitis vinifera* 12X.v2 of Grape Genome Browse (Genescope.cns.fr).

Furthermore, the upstream and downstream regions flanking each SNP were also investigated to search for interesting genes in linkage disequilibrium with SNPs and possibly involved in phenological processes.

3.3 Primer design

The design of primers for the amplification of genomic regions of each SNP was carried out through Primer3 (Koressaar et *al*., 2018); by entering the nucleotide sequence of the genomic interval at + 500 / -500 bp from the SNP, it was possible to obtain primer pairs that meet design requirements. The requirements for each primer were:

- distance of at least + 150 / -150 bp from the SNP to ensure a good reading during Sanger sequencing
- maximum length of the PCR product of 800 bp
- GC content of 45-55%
- ideal melting temperature 60° C

Once obtained a primer pair for each region, by using Primer BLAST software (Ye J. et al., 2012) it was verified that each pair found matches only in the selected region and did not produce non-specific products.

3.4 Genomic DNA extraction from pool high/low individuals

Genomic DNA was extracted from each of 94 individuals using the DNeasy Plant Mini kit (Qiagen, Germany). Leaf samples of each cultivar composing the Conegliano's Collection were sampled in 2017 and stored individually in plastic bags at -80° C at the Department of Biotechnology (Verona). The extraction protocol included the following steps:

19) Sterile 1.5 mL Eppendorf tubes are prepared with 1.7 mm glass beads (Roth, Germany) for mechanical grinding of the tissue, marking the cap with the name of the genotype.
20) For each genotype, 300 mg of frozen leaf tissue are sampled and placed in the tube with glass beads. Tubes are then immediately placed in liquid nitrogen and proceed with mechanical grinding of the tissues through the Silamat S5 grinder, taking care not to thaw the sample. Tubes are then removed from liquid nitrogen with pliers and quickly placed in the grinder. Each sample are grinded for about 6 seconds and then frozen again in nitrogen. The operation is repeated 3 times until a powder is obtained.
21) In the meantime, prepare an equal number of 1.5 ml labeled Eppendorf tubes and aliquot 500 µL of the buffer AP1 pre-heated to 65° C in a water bath. Tubes with buffer are then placed in the water bath to maintain the temperature.
22) After grinding, all the powder is poured into the tubes containing the lysis buffer and careful mixing is carried out, so that all the powder meets the buffer.
23) Aliquot 5 µL of RNase A in each sample and mix thoroughly.
24) The samples are incubated in water bath at 65 ° C for 10 minutes, inverting the tubes every 3 or 4 minutes.
25) After incubation, 163 µL of buffer P3 are aliquoted in each sample; the solution is mixed and incubated on ice for 5 minutes.
26) The lysate is centrifuged for 5 minutes at 14.000 rpm.

27) The supernatant is then carefully removed without touching the pellet and placed in the previously marked QIAshredder column. The column is centrifuged for 2 minutes at 14.000 rpm.
28) The eluate is then transferred to a new marked 1.5 mL Eppendorf tube without touching the pellet at the bottom of the column (if present).
29) 1.5 volumes of buffer AW1 are aliquoted to the solution and mixed by pipetting.
30) 650 µL of the obtained solution are transferred into the DNeasy Mini spin column and centrifuged for 1 minute at a speed of at least 8.000 rpm.
31) The eluate is discarded and the step is repeated with the remaining solution always using the same column.
32) The column is then moved to a new 2 mL elution tube and 500 µL of the AW2 buffer is added. Column is centrifuged for 1 minute at a speed of at least 8.000 rpm.
33) Another 500 µL of AW2 buffer are added and centrifuged for 2 minutes at 14.000 rpm.
34) Transfer the column to a new sterile 1.5 ml labeled Eppendorf tube and aliquot 100 µL of sterile MilliQ water exactly on the column filter, but without touching it, to allow the elution of the DNA from the silica matrix.
35) Tubes are then incubated for 5 minutes at room temperature and centrifuged for 1 minute at a speed of at least 8.000 rpm
36) Repeat step 17. Keep the extracted DNA on ice.

Each DNA extracted was then quantified with NanoDrop One/OneC spectrophotometer (Thermo Fisher, USA) and qualitative analyzed through 0,8% agarose gel run.

3.5 Optimization of amplification conditions

Optimization of amplification conditions was carried out in order to obtain specific products. The amplification of the regions was carried out using the PCRBio Hi-

Fidelity taq polymerase (Resnova, Italy) to limit as much as possible the introduction of accidental errors during the extension phase and amplified with SimplyAmp thermocycler (Applied Biosystems, USA). Once conditions were set, each primer pair was used to amplify 94 genotypes of pool high/low in a final volume of 25 uL of PCR mix. Once the amplification program was complete, the products were run on 1.5% agarose gel, to verify the correct size of amplicons and to exclude the presence of non-specific products. **Tables 15a** and **15b-c-d** show PCR mixes and amplification conditions for the three pairs of primers.

	SNP chr12_21100465_A_G	**SNP chr16_348557_T_G**	**SNP chr18_3381571_G_A**
Reagents	Final concentration	Final concentration	Final concentration
PCRBio Buffer	1x	1x	1x
Primer F	0.4 uM	0.4 uM	0.4 uM
Primer R	0.4 uM	0.4 uM	0.4 uM
PCRBio Hi-Fi Taq	0.5 U	0.5 U	0.5 U
genomic DNA	1ng/uL	1ng/uL	1ng/uL

Table 15a - PCR mixes. PCRBio Buffer contains 1.5 mM $MgCl_2$ and dNTPs

SNP chr12_19533740_A_G			
Initial Denaturation		95°C	1 minute
35X	Denaturation	94°C	15 seconds
	Annealing	67°C	15 seconds
	Extension	72°C	30 seconds
Final Extension		72°C	5 minutes
		4°C	∞

Table 15b – Amplification conditions of primer pair for SNP chr12_19533740_A_G.

SNP chr16_348557_T_G			
Initial Denaturation		95°C	1 minute
35X	Denaturation	95°C	15 seconds
	Annealing	57°C	15 seconds
	Extension	72°C	30 seconds
Final Extension		72°C	5 minutes
		4°C	∞

Table 15c – Amplification conditions of primer pair for SNP chr16_348557_T_G.

SNP chr18_3381571_G_A			
Initial Denaturation		95°C	1 minute
35X	Denaturation	95°C	15 seconds
	Annealing	60°C	15 seconds
	Extension	72°C	30 seconds
Final Extension		72°C	5 minutes
		4°C	∞

Table 15d – Amplification conditions of primer pair for SNP chr18_3381571_G_A.

3.6 Sanger sequencing

3.6.1 Amplicon purification

The sample preparation pipeline involves first the removal of salts, primers and other reagents present in the amplification mix. For this purpose, Ampure XP magnetic beads (Beckman Coulter, USA) were used which allow to electrostatically retain all nucleotide fragments higher than 100 bp and eliminate all other reagents.

For each sample, the purification protocol was performed as follows:

- Vortex and resuspend the Ampure XP beads
- Add 1.8 uL of beads every 1 uL of each PCR product (45 uL of Ampure for 25 uL) in a 96-well plate.
- Mix by pipetting slowly 10 times and incubate at room temperature for 10 minutes.
- Move the plate to the appropriate support with magnets and incubate for 2 minutes. The solution must become clear.
- While holding the plate on magnets, discard the supernatant without disturbing the beads and leave 5 uL of solution to keep the beads from drying out.
- Still holding the plate on magnets, add 200 uL of fresh 70% ethanol solution and incubate for 30 seconds. Aspirate all the ethanol and discard it.
- Repeat the 70% ethanol wash and allow the ethanol residues to evaporate without drying the beads too much.
- Remove plate from magnets and elute with 36 uL of sterile MilliQ water. Incubate at room temperature for 5 minutes.
- After that, pipette gently to resuspend the beads and incubate for 10-15 minutes, pipetting a couple more times during this interval.
- Place the plate back on the magnets and incubate for 2 minutes to separate the beads from the DNA.
- Keeping the plate on the magnets, transfer the eluate into new tubes, taking care not to carry any bead residues.

In this way, purified PCR products were obtained to be checked on 1.5% agarose gel to ensure no material was lost during purification.
Once the successful completion of the purification has been verified, all samples were quantified through NanoDrop One/OneC spectrophotometer (Thermo Fisher, USA).

3.6.2 BigDye v3.1 Terminator labeling

The labeling of purified amplicons was carried out with BigDye v3.1 Terminator kit. This sample preparation was performed with a 2X quantity of DNA, as recommended from the guide, and in forward or reverse orientation. The following labeling mix (**table 16a**), was prepared for all 94 samples for each pair of primers:

Reagent	Volume for 1 reaction
Sequencing Buffer 5X	1.5 uL
BigDye Terminator	1 uL
Primer R or F (3.2 uM)	1 uL
Purified PCR product	*See table x*
H2O	to 10 uL

Table 16a– Composition of BigDye v3.1 terminator mix

The amount of DNA was chosen based on the length of amplicon. From **table 16b** it is therefore possible to deduce the quantitative range (minimum and maximum ng) to be used to obtain a good quality reading.

Length of amplicon (bp)	**ng for 2X quantity**	
	Minimum	**Maximum**
380	7.6	15.2
400	8	16
420	8.4	16.8
440	8.8	17.6
460	9.2	18.4
480	9.6	19.2
500	10	20
520	10.4	20.8
540	10.8	21.6
560	11.2	22.4
580	11.6	23.2

Table 16b – Quantity of DNA required for the labeling reaction with BigDye v3.1 Terminator related to the length of the amplicon.

In our case, for the amplicon on chr12, 8 to16 ng are required while for chr16 and chr18 are required 10.4 to 20.8 ng. According to concertation obtained from the reading with Nanodrop for each purified sample, the corresponding uL of DNA were therefore added to obtain the required quantity.

Once labeling mix was assembled, samples were placed in SimplyAmp thermocycler and the following program was used for labelling:

Initial Denaturation		96°C	1 minute
25X	Denaturation	96°C	10 seconds
	Annealing	50°C	5 seconds
	Extension	60°C	4 minutes
	4°C	∞	

Table 17– BigDye v3.1 terminator program on SymplyAmp thermocycler

Finally, the last step of sample preparation involved a purification for the removal of reagents before loading onto the sequencer. The purification protocol, in accordance with the kit manual, suggested the following phases:

- Prepare a new 96-well plate
- Sequentially, aliquot in each tube 2.5 uL of EDTA 125 mM pH 8 and 1 uL of NaAc 3M
- Add 10 uL of sequence product
- Add 30 uL of EtOH 100%
- Vortex briefly and incubate at RT for 15 minutes (important not to exceed with time) covering the samples with aluminum
- After incubation, place samples in a centrifuge and spin at 2250xg for 30 minutes at 4° C
- Then, remove samples. Place a sheet of paper towels on each centrifuge slot and over them the upside-down open plate with samples to remove the purification solution.
- Spin the inverted samples at 180xg for 1 minute.

- At the end, remove the samples and proceed with a wash by adding 70% EtOH, aliquoting 40 uL in each sample.
- Centrifuge the samples with 70% EtOH for 15 minutes at 1870xg
- At the end, discard the ethanol using the same system with sheets of paper towels and inverted samples. Spin at 180xg for 1 minute.
- After the centrifuge, dry the samples completely for 5-10 minutes under the fume hood and then closed with caps.

Samples were stored at -20° C until the time of sequencing. At the moment of analysis, samples were thawed and rehydrated by adding 10 uL of sterile MilliQ water to each. After resuspension and 30-minutes incubations at 4 ° C, 5 µL of each sample was placed in a new plate. 5 uL of Hi-Di formamide (Applied Biosystems, USA) were added to each one and were finally loaded into Abi 3130XL automatic sequencer (Applied Biosystems, USA) for sequence analysis.

All sequencing operations were carried out at the Laboratory of Molecular Biology at CREA-VE in Conegliano (TV) under the supervision of Prof. Manna Crespan.

3.6.3 Analysis of Sanger sequencing results

Once runs on Sanger sequencer were finished, sequencing outputs were analyzed with MEGA 7 software (Kumar et al., 2016). The software allows to view the chromatograms resulting from the sequencing, to ascertain their quality. Thus, it also allows to create alignments between a query sequence and a reference sequence, to search for polymorphisms between the two.

In this way, by aligning the reference sequence to the sequences obtained from the samples under analysis, it was possible to identify the variants carried by each of the 94 genomes tested.

3.7 Statistical analysis

The *gene count* method was used to estimate allele frequencies in the groups high and low, for each of the three SNPs.

$$p = f(\mathrm{A}) = \frac{2 \times (\text{Individuals AA}) + (\text{Individuals Aa})}{2 \times (\text{total number of individuals})}$$

$$q = f(\mathrm{a}) = \frac{2 \times (\text{Individuals aa}) + (\text{Individuals Aa})}{2 \times (\text{total number of individuals})}$$

$$p + q = 1$$

Finally, to validate significative differences of allele frequency distributions in a group respect to the other, a chi-square test was conducted.

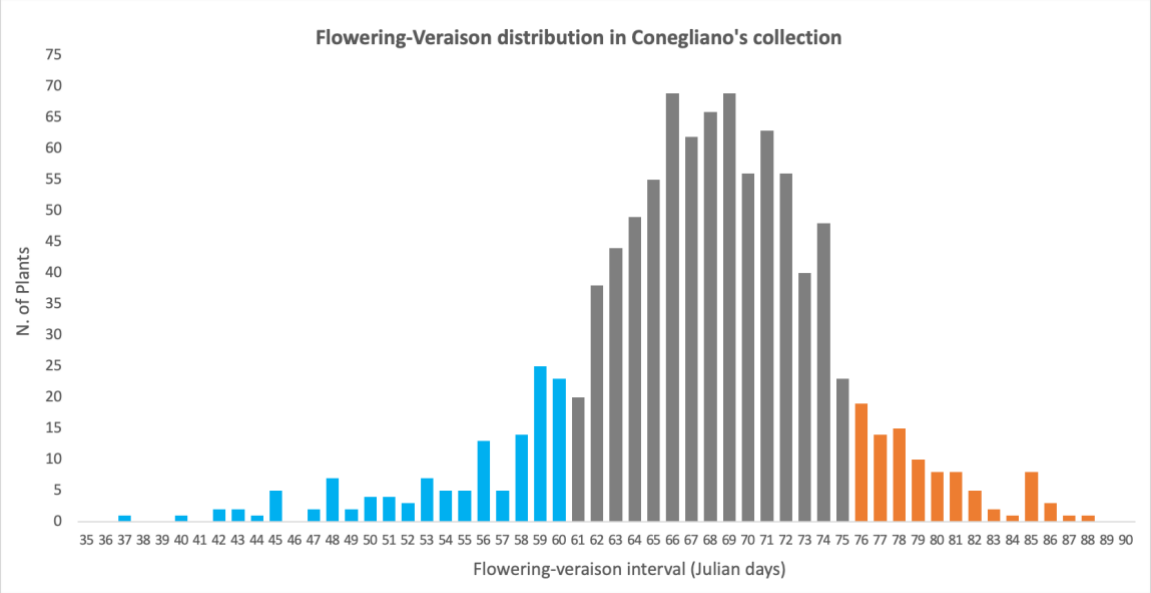
4. Results

4.1 Selection of individuals with extreme phenotype for validation of association

For the validation of association with the interval flowering-veraison phenotype for the three SNPs identified by GWAS approach conducted on the core collection, two groups of plants with extreme and divergent phenotype were selected from the whole plants originally used to build up the core.

More in detail phenotypic data collected over 11 years on Conegliano's collection were considered to select those genotypes with a very wide (group high) or very short (group low) interval. **Graphic 11** shows the distribution of the entire collection for flowering-veraison interval (reported in the abscissas in Julian days) obtained as average from raw data of 11 years.

Phenotypic tails were therefore chosen for the validation purpose, represented **in blue** by individuals with a short interval and **in orange** by individuals with a wide interval. For each tail, 47 individuals (7,6% of the collection) were selected to compare the allelic frequencies of SNPs previously identified as associated to the phenotype.



Graphic 11 – Distribution of average flowering-veraison interval of entire Conegliano's collection. In blue, individuals composing *low* phenotypic tail group of the collection with short interval. In orange, individuals composing *high* phenotypic tail group of the collection with large interval.

In **tables 18a-b** are reported cultivars included in each group used in this study, with the given name of the cultivar, its identification code in the collection and the flowering-veraison interval (average of 11 years) expressed in Julian days.

Code	Flowering-veraison interval (Julian days)	Cultivar	Color
731	88	FOGARINA	noir
1093	87	CATARRATTONE	blanc
437	86	LATTUARIO_NERO	noir
580	86	RABOSO_PIAVE	noir
701	85	BRAMBANA	noir
1068	84	CAVRARA_NERA	noir
260	83	RABOSO_VERONESE	noir
703	83	ARAMONE	NA
429	82	CIMINNITA	blanc
1074	82	CORBINONA_13_GT	noir
1084	82	ULIVELLO	noir
1109	82	OLIVETTA_VIBONESE	noir
569	82	NERA_GROSSA	NA
471	82	INZOLIA_IMPERIALE	noir
1107	81	EMPEROR	NA
1194	81	SCHIOCCOLETTO	NA
688	81	AXINA_DE_TRES_BIAS	noir
465	80	PIZZUTELLO_BIANCO	blanc
150	80	LUMASSINA	blanc
78	80	COLOMBANA_NERA	NA
135	79	INVERNENGA	blanc
142	79	LAMBRUSCO_DI_SORBARA	noir
141	79	LAMBRUSCO_A_FOGLIA_FRAST	noir
354	79	VERNACCIA_NERA	noir
1176	79	GORGOTTESCO	NA
179	79	MOLINARA	noir
120	79	GRECO_DI_TUFO	blanc
685	78	ARSILICO	noir
518	78	CORBINELLA_PADOVANA	noir
436	78	PERLONA	blanc
42	78	BOSCHERA	blanc
715	78	CICINESE_GROSSO	noir
717	78	CODELUNGHE	noir
1151	78	BOGGIONE_NERO	noir
1167	78	COLORINO_FORTE	NA
128	78	GUARDAVALLE	blanc
97	78	FERTILIA	noir
514	78	CIAVARELLA_FRANCESE	NA
1115	77	RED_OHANEZ	red
575	77	PEDICELLO_ROSSO	noir
277	77	SGIUSEPPE_NERO	noir
250	77	POLLERA_NERA	noir
269	77	RONDINELLA	noir
1191	77	ROSETTA	NA
951	77	CACCIO_DI_FERMO_1	blanc/noir
39	77	BOMBINO_NERO	noir
1110	76	INZOLIA_PARCHITANA	blanc

Table 18a – Composition of *high* group. From left, the cultivar's code of the Conegliano collection, the flowering-veraison interval (average value from 11 years), name of cultivar given in Conegliano and color of berry.

Code	Flowering-veraison interval (Julian days)	Cultivar	Color
485	40	ORA	blanc
464	42	PERLA_DI_CSABA	blanc
1113	42	DIAMANT	blanc
469	43	S_ANNA_DI_LIPSIA	blanc
483	43	CANNER_SEEDLESS	blanc
458	45	VICTORIA	blanc
493	46	DELIGHT	blanc
479	46	EARLY_MUSCAT	blanc
494	47	PERLETTE	blanc
1099	47	SULTANA	blanc
421	48	MATILDE	blanc
457	48	MADDALENA_BRUNI	blanc
1098	48	SULTANINA_BIANCA	blanc
468	49	REGINA_DEI_VIGNETI	blanc
426	49	CHASSELAS_BLANC	blanc
484	49	DANUTA	NA
433	51	CARDINAL	rouge/noir
1097	51	MONUKKA	blanc/noir
488	51	CORINTO_NERO	noir
480	51	RED_GLOBE_1	rouge
472	53	JULY_MUSCAT	blanc
35	53	BLANC_DE_MORGEX	blanc
1116	53	AURORA_2	NA
495	53	APIRENA_BRUNI	NA
112	53	GAMARET	noir
642	54	BACCHUS	NA
534	54	UNKNOWN	NA
477	55	CORINTO_BIANCO	blanc
132	55	INCROCIO_BRUNI_54	blanc
115	55	GARANOIR	NA
645	56	WURZER	NA
602	56	TERBASCI_cp2	noir
1029	56	BELLISE	NA
1183	56	MOSCATO_DI_SPAGNA	NA
283	56	SANTA_MARIA	blanc
651	57	GOLDRIESLING	NA
593	57	SAN_LORENZO	noir
74	57	CILIEGIOLO	noir
424	58	BICANE	blanc
449	58	CLOTILDE_PROSPERI	blanc
693	58	BIANCHETTA_DI_BACEDASCO	blanc
694	58	BIANCHETTA_DI_DIOLO	blanc
197	58	MULLER_THURGAU	blanc
637	58	MUSCAT_OTTONEL	NA
846	59	UVA_DI_CRIMEA	noir
6	59	ALBANA	blanc
251	59	PORTOGHESE	ukn

Table 18b – Composition of *low* group. From left, cultivar's code of the collection in Conegliano, the flowering-veraison interval (average value from 11 years), cultivar name given in Conegliano and color of berry.

4.2. SNPs annotation

The annotation of genomic regions in which the three SNPs significantly associated with the flowering-veraison interval from the GWAS analysis fall was carried out manually. By using the genomic positions on v1 annotation and assembly 12x.v2, it was possible to determine whether SNPs fall into intergenic or intragenic regions. The search was carried out using physical coordinates of markers on 12X.v2 assembly which made it possible to visualize the region of genome in which each SNP is located. In all three cases, SNPs are placed within coding sequences of genes with different functions.

Observing **table 19**, SNP chr12_19533740_A_G is located in *VIT_12s0035g00280* gene, and in particular in its first exon. The gene extends from position 21083032 to 21100612 (12x.v2) of chromosome 12 and encodes an endomembrane protein 70.

SNP chr16_348557_T_G is instead found in the fourth exon of *VIT_16s0039g0700* gene, positioned from 346728 to 353020 bp of chromosome 16 and encoding a maltose transporter (or RCP1 in *Arabidopsis thaliana*).

Finally, SNP chr18_3381571_C_T is located in the sixth exon of *VIT_18s0001g03540* gene, located between 3380455 and 3383288 bp of chromosome 18, encoding for an auxin transducer and involved in endogenous signaling.

Table 19 – Annotation of genomics regions in which significant SNPs from GWAS annotation fall.

					Gene information				
SNP ID	***Chromosome***	**Position 12X.v0 (bp)**	**Position 12X.v2 (bp)**	***Genomic region***	***Gene ID***	***Gene position 12x.v2 (bp)***	***SNP location***	***Functional annotation***	***Gene ontology***
chr12_19533740_A_G	12	19533740	21100465	Gene (CDS)	VIT_12s0035g00280	21083032-21100612	Exon 1	Endomembrane protein 70	-
chr16_348557_T_G	16	348557	348557	Gene (CDS)	VIT_16s0039g0700	346728-353020	Exon 4	RCP1	Maltose excess protein transporter
chr18_3381571_C_T	18	3381571	3381571	Gene (CDS)	VIT_18s0001g03540	3380455-3383288	Exon 6	Auxin transporter protein 4	Auxin transport, signal transduction, response to endogenous signals

Since all polymorphisms fall into CDS, the effect of the mutation on encoded protein was also evaluated for each SNP. Going back to the coding sequence of each gene, the variation of single nucleotide was predicted through PROVEAN software.

The result, in all three cases, is a synonymous mutation of the aminoacidic sequence and therefore no variation in the functions of the gene product is expected.

4.3 Primer design for SNP genotyping by Sanger sequencing

The design of primers for the amplification of the genomic regions containing SNPs was carried by using different software to ensure the design of specific primer pairs matched only the target region. In **table 20** are listed the three pairs of primer finally used in this study for the amplification of regions.

SNP	Primer Forward (5'-3')	Primer Reverse (5'-3')	Annealing Temperature	Product size (bp)
Chr12	ACCTACTTGAGATCTAAGCCCTACG	CTCTGTACCTCGCCCTCCTTCTC	67°C	390
Chr16	GGAAAGACAATAAAATCTCTCCAAT	CTCAAGCTGCAGATCAAAGGAAACT	57°C	513
Chr18	ACCTAAATTTTTGCAAACTTTAGCC	TTCAGAACTCAAAATTTCTTAAAAT	60°C	509

Table 20 – Primer pairs for amplification of genomic regions including the associated SNPs.

4.4 SNP genotyping by Sanger sequencing

The sequencing runs conducted at the Laboratory of Molecular Biology at CREA-VE in Conegliano (Prof. Manna Crespan) allowed to obtain the exact sequences of amplicons under analysis and to know whether the wild-type or mutant variant was

present, in homozygosity or heterozygosity, for each of the three SNPs and in every genome belonging to the *high* and *low* groups.

However, for some samples it was not possible to get this information. In some cases, in fact, it was not possible to obtain a clean sequence from the chromatograms, which gave ambiguous or absent signals. These errors are likely due to incomplete labeling reactions or difficulty of the sequencer capillaries in reaching the sample in the plate.

However, it was possible to obtain a sufficiently large number of sequences to calculate and compare SNP's allele frequencies in the two groups.

Below, in **tables 21a-c**, sequencing results with variants for each of the SNPs analyzed are reported.

chr12_19533740_A_G *(forward orientation)*			
Sample	*Variant*	*Sample*	*Variant*
731	G/G	**433**	ND
429	A/G	**112**	A/G
1093	A/G	**1097**	ND
1074	A/G	**642**	G/G
437	G/G	**488**	ND
1084	A/G	**534**	A/G
580	G/G	**480**	ND
1109	G/G	**477**	ND
701	A/G	**472**	A/G
569	G/G	**132**	ND
1068	G/G	**32**	G/G
471	G/G	**115**	G/G
260	G/G	**1116**	A/G
1107	G/G	**645**	A/G
703	A/G	**495**	A/A
1194	G/G	**602**	A/G
688	G/G	**1029**	A/G
1176	ND	**693**	A/G
465	A/G	**1183**	A/A
179	A/G	**694**	ND
150	ND	**283**	A/G
120	ND	**197**	A/G
78	A/G	**651**	ND
685	A/G	**637**	A/A
135	A/G	**593**	A/G
518	ND	**846**	ND
142	A/G	**74**	ND
436	G/G	**6**	ND
141	ND	**424**	ND
42	ND	**251**	A/G
354	A/G	**449**	ND
715	A/G	**485**	G/G
717	G/G	**494**	A/G
277	A/G	**464**	A/A
1151	ND	**1099**	A/G
250	ND	**1113**	G/G
1167	G/G	**421**	G/G
269	G/G	**469**	G/G
128	ND	**457**	A/A
1191	ND	**483**	A/G
97	ND	**1098**	A/A
951	A/G	**458**	A/G
514	ND	**468**	A/G
39	G/G	**493**	A/A
1115	ND	**426**	A/G
1110	ND	**479**	A/A
575	ND	**484**	A/G

Table 21a – Sequencing results on SNP chr12_19533740_A_G. Sample in orange belong to the *high* group. Sample in blue belong to the *low* group. Green cells: homozygous wild-type variants. Light red cells: heterozygous variants. Red cells: homozygous alternative variant. ND: not detectable.

Chr16_348557_T_G *(forward orientation)*			
Sample	*Variant*	*Sample*	*Variant*
731	G/T	**485**	T
437	G/T	**483**	T
701	G/T	**426**	T
260	T	**433**	T
1074	G/G	**1097**	T
465	G/G	**488**	T
150	G/G	**480**	T
78	T	**472**	T
135	ND	**35**	ND
142	G/G	**1116**	T
141	G/T	**495**	T
354	G/T	**112**	T
1176	T	**642**	G/T
179	T	**534**	T
120	T	**477**	T
685	T	**132**	ND
518	G/G	**115**	T
436	ND	**645**	G/T
42	T	**602**	T
715	G/T	**1029**	G/G
717	T	**1183**	G/T
1151	T	**283**	ND
1167	G/T	**593**	G/T
128	ND	**74**	ND
97	G/T	**424**	ND
514	G/T	**449**	ND
1115	G/G	**693**	ND
575	T	**694**	ND
277	G/T	**197**	ND
250	T	**637**	ND
269	G/G	**846**	ND
1191	ND	**6**	ND
951	G/T	**251**	ND
39	G/T	**464**	T
1110	ND	**1113**	G/G
1093	T	**469**	T
437	ND	**483**	ND
580	T	**458**	T
701	ND	**493**	T
1068	T	**479**	G/T
260	ND	**494**	T
703	G/T	**1099**	T
429	G/G	**421**	ND
1074	ND	**457**	T
1084	T	**1098**	T
1109	G/G	**468**	T
569	T	**426**	T
471	G/G	**484**	G/G
1107	G/T	**433**	ND
1194	G/T		
688	G/G		

Table 21b – Sequencing results on SNP Chr16_348557_T_G. Sample in orange belong to *high* group. Sample in blue belong to the *low* group. Green cells: homozygous wild-type variants. Light red cells: heterozygous variants. Red cells: homozygous alternative variant. ND: not detectable.

Chr18_3381571_G_A *(reverse orientation)*			
Sample	*Variant*	*Sample*	*Variant*
471	T	**494**	ND
1194	T	**1099**	ND
150	T	**421**	C/T
78	ND	**457**	ND
141	ND	**488**	C/T
1176	T	**472**	ND
685	T	**35**	T
518	T	**495**	T
436	T	**642**	C/T
42	T	**543**	C/T
715	T	**645**	C/T
717	T	**602**	C/T
1151	C	**1029**	ND
97	ND	**1183**	C/T
514	T	**283**	T
1115	T	**651**	C/T
575	ND	**593**	C/T
277	T	**74**	C/T
250	T	**424**	ND
269	C/T	**693**	T
H2O	ND	**694**	T
951	C/T	**197**	T
39	C/T	**637**	C
731	T	**846**	T
1093	C/T	**6**	T
580	T	**251**	T
179	C/T	**1116**	C/T
437	C/T	**485**	T
701	T	**464**	C
1068	T	**1113**	T
447	T	**483**	C/T
260	T	**458**	T
569	T	**493**	C
1176	ND	**495**	ND
1109	T	**494**	ND
1074	T	**1099**	ND
429	T	**421**	ND
688	T	**112**	C/T
1107	C/T	**1098**	T
150	ND	**468**	C
465	T	**426**	C
354	T	**433**	T
120	T	**1097**	T
142	T	**488**	ND
135	T	**480**	C/T
97	ND	**115**	C/T
128	C/T	**477**	C/T
1167	T	**479**	C/T
703	T	**469**	T
1084	C/T	**484**	C
1110	T		

Table 21c – Sequencing results on SNP Chr18_3381571_G_A. Sample in orange belong to *high* group. Sample in blue belong to the *low* group. Green cells: homozygous wild-type variants. Light red cells: heterozygous variants. Red cells: homozygous alternative variant. ND: not detectable.

4.5 Allele frequencies in the *high* and *low* groups and statistical analysis

Allele frequencies for each of the three SNPs in the two different pools were calculated. The method used, as reported in *paragraph 3.7* (page 113), allowed to evaluate differences in the distributions of the wild-type or alternative variants in the two pools.
Below, for each SNP, the allele frequencies calculated on the total of the available sequences are reported.

4.5.1 Allele frequencies of SNP chr12_19533740_A_G

chr12_19533740_A_G	
Group HIGH	Group LOW
Total number of genotypes **32**	*Total number of genotypes* **34**
Number of genotypes AA **0**	*Number of genotypes AA* **8**
Number of genotypes AG **16**	*Number of genotypes AG* **19**
Number of genotypes GG **16**	*Number of genotypes GG* **7**

Table 22a – Summary of sequencing results on SNP chr12_19533740_A_G

Allele frequency A	Allele frequency A
0.250	0.515
Allele frequency G	Allele frequency G
0.750	0.485

Table 22b – Allele frequencies of wild-type/alternative variant of SNP chr12_19533740_A_G among the two groups.

Table 22a shows the overall results of sequencing carried out on the 47 individuals of the high and 47 of the low groups for the analysis of variants of SNP chr12_19533740_A_G. Overall, 32 readable sequences were obtained for

individuals belonging to the *high* group and 34 for individual belonging to the *low* group. The wild-type variant "A" of the base was not found in homozygosity for any of the genomes analyzed belonging to the group *high*, while in the *low* group seems to be present an enrichment for this genotype, finding 8 individuals with the AA genotype. The heterozygous genotype AG is found in 16 individuals from the pool high and 19 individuals from the pool low. As for the alternative variant G, in the individuals with *high* phenotype there is an enrichment for the homozygous genotype GG, with 16 individuals presenting it, while only 7 individuals in the group *low* show the alternative variant in homozygosity.

These results are evident from the allele frequencies shown in **table 22b**, in which a higher frequency of the wild-type A allele in the low group is evident, while a higher frequency of the mutated G allele in the high group is observed.

Overall, it is possible to state that most individuals with a large flowering-veraison interval share the alternative variant of the SNP more often than individuals with a short interval.

4.5.2 Allele frequencies of SNP chr16_348557_T_G

Chr16_348557_T_G	
Group HIGH	Group LOW
Total number of genotypes **42**	*Total number of genotypes* **33**
Number of genotypes TT **16**	*Number of genotypes TT* **25**
Number of genotypes TG **15**	*Number of genotypes TG* **5**
Number of genotypes GG **11**	*Number of genotypes GG* **3**

Table 23a – Summary of sequencing results on SNP chr16_348557_T_G

Allele frequency T	Allele frequency T
0.560	0.833
Allele frequency G	Allele frequency G
0.440	0.167

Table 23b – Allele frequencies of wild-type/alternative variant of SNP chr16_348557_T_G among the two groups.

As for the analysis of SNP chr16_348557_T_G, from **table 23a** it is possible to note that 42 readable sequences were obtained for the individual belonging to the group high and 33 for the group low. The wild-type variant "T" of the base is shared in homozygosity by 16 individuals in the group high, while in low it is found in 25 individuals. 15 individuals in the group high have a heterozygous genotype, while only 5 have it in the low. Finally, the mutated variant GG is found in homozygosity in 11 genomes of the high, compared to only 3 in the low group.

It is therefore evident, also from **table 23b**, that there is an enrichment for the wild-type base in low phenotype, while an enrichment of the mutated base G of the high phenotype is even more marked compared to the low.

Also in this case, therefore, it seems that individuals with a wide flowering-veraison interval more often share the alternative variant than the wild-type base.

4.5.3 Allele frequencies of SNP chr18_3381571_C_T

Chr18_3381571_C_T	
Group HIGH	Group LOW
Total number of genotypes **43**	*Total number of genotypes* **49**
Number of genotypes CC **1**	*Number of genotypes CC* **6**
Number of genotypes CT **9**	*Number of genotypes CT* **26**
Number of genotypes TT **33**	*Number of genotypes TT* **17**

Table 24a – Summary of sequencing results on SNP Chr18_3381571_C_T.

Allele frequency C	Allele frequency C
0.128	0.388
Allele frequency T	Allele frequency T
0.872	0.612

Table 24b – Allele frequencies of wild-type/variant of SNP Chr18_3381571_C_T among the two groups.

A total of 43 readable sequences for the individual belonging to the high group and 40 for the low were obtained by sequencing the samples for the SNP chr18_3381571_C_T. As can be seen from **table 24a**, in this case, the wild-type base "C" was found in homozygosity in only one individual of the high group and in 6 individuals of the low. Continuing, 9 individuals in the high group exhibit a heterozygous genotype versus 26 in the low.

Finally, 33 individuals in the pool high have the mutated variant "T" in homozygosity, while in the pool low only 17 have a TT genotype.

As can also be seen from **table 24b**, the wild-type variant appears significantly more frequent in genomes of low phenotype than in those with high phenotype. The alternative variant, on the other hand, seems to have a slight enrichment among the individuals of the group high, although it would appear to be fairly represented also among individuals with a short interval as it is often present in heterozygosity.

4.5.4 Statistical analysis

To confirm the significance of allele frequencies differences in groups *high* and *low* calculated on each variant and for each SNP, a chi-square test was conducted. This confirmed a significant enrichment of the alternative variants in all three cases (**table 25**) in the *high* group.

Chr18_3381571_C_T	HIGH	LOW	***p-value***
C	11	38	0.0001382
T	75	60	

Chr16_348557_T_G	HIGH	LOW	***p-value***
T	47	55	0.0006934
G	37	11	

Chr12_19533740_A_G	HIGH	LOW	***p-value***
A	16	35	0.003254
G	48	33	

Table 25 – Summary of chi-square test conducted on allele frequencies of SNPs variants. From left, nucleotide variant of SNP, number of allele representation (homozygous and heterozygous) in group *high* and in group *low* and the p-value of the chi-square statistical test.

4.6 Investigation of flanking regions in linkage disequilibrium

The upstream and downstream regions flanking the three SNPs were also investigated. As reported in *Introduction* (page 100), considering an average linkage disequilibrium in grapevine of +/- 50.000 bp, flanking regions were investigated in order to search for genes that could be in linkage disequilibrium with the SNPs and involved in the regulation of the phenotype. **Table 47** reported in *Appendix* on page 197 lists interesting genes in these regions. Possible candidates are discussed below.

5. Discussion 2

In this second chapter we validate the results of a genome-wide association study performed on a grapevine germplasm collection. The GWAS study was previously conducted analyzing the genetic diversity of the collection, in order to identify the most genetically diverse and representative panel of individuals to be used for genetic association. From the collection, a final number of 617 unique genotypes were considered according to analysis at 45 SSR markers. Among these, a diversity panels of 132 genotypes to be used for association analysis was created. Out of 132 genotypes, SNP genetic profiles of 95 individuals were previously published from Laucou et al. (2018) and this genetic information were integrated with missing profiles of 37 individuals generated through the use of the GrapeReseq *Vitis*18KSNP chip.

To find the most interesting regions linked to veraison time traits, three approaches represented by different software that apply different statistical procedures were integrated. Overall, GWAS analysis identified 15 SNPs markers putatively associated with the three phenotypic traits investigated but, among these, for three in particular multiple evidence were found (either identified as associated to more traits or by more software) (Pietro Delfino, PhD thesis).

The three markers selected by the analysis are positioned on chromosomes 12 (chr12_19533740_A_G), 16 (chr16_348557_T_G) and 18 (chr18_3381571_C_T) and their association was validated in this study by using a subset of the collection created with individuals with the most extreme phenotype. The individuals of the subset were selected on the basis of phenotypic data collected for 11 years on the CREA-VE collection and on which the average interval from flowering to veraison was calculated. From the phenotypic tails of the collection, individuals belonging to the so called *high* and *low* groups on which to test the association with the three candidate SNPs were then selected. The allelic variants of each SNP were verified

by Sanger sequencing of the regions containing the SNPs on the 94 individuals of the subsets.
The calculation of the allelic frequencies of each variant, wild-type or alternative, showed in all three loci a tendency of the *high* group to present enrichments for the alternative variant. In the case of the SNPs chr12_19533740_A_G and chr18_3381571_C_T it was in fact evident an enrichment of the alternative variants in homozygosity in the *high* group compared to the *low* which tends to have enrichments for the wild-type base. As for the SNP chr16_348557_T_G, the *high* group shows an enrichment of the alternative base in heterozygosity and a notable enrichment of the wild-type base in the *low* group.

For each SNPs, a verification of the position on genome and a study on the effect of the mutation were carried out. In all three cases, SNPs fall into gene coding portions: SNP chr12_19533740_A_G is positioned on exon 1 of the gene VIT_12s0035g00280, which encodes for an endomembrane protein 70. SNP chr16_348557_T_G is positioned on exon 4 of the gene VIT_16s0039g0700 called RCP1 protein, a maltose transporter responsible for sugar processing.
Finally, SNP chr18_3381571_C_T is positioned on exon 6 of the VIT_18s0001g03540 gene and which encodes an auxin transducer involved in signal transduction in response to endogenous stimuli. This SNP is particularly interesting because it seems to co-localize with a meta-QTL region for phenology and resulting from the intersection of the results of Carreno Ruiz et al. (2012) and Zyprian et al. (2016).
For each case, the effect that mutations produce on protein sequences was evaluated; all three SNPs produce synonymous effects on proteins.

From the studies carried out for these three polymorphisms it is possible to state that they are not causative SNPs of the phenotype. However, they may be part of a haplotype that encompasses the true causative SNP. As discussed in the introduction, linkage disequilibrium in grapevine tends to be restricted to about 50.000 bp (Nicolas et al., 2016). For this reason, regions flanking the 3 SNPs in the neighborhood of + 50.000 / -50.000 bp were also considered. In this way it was

possible to obtain a list of genes with interesting ontology genes that could be involved in the regulation of veraison. Table 47 reported in *Appendix* (page 197) shows a list of genes in LD for each of the 3 chromosomes.
Similarly, to what was done for QTL mapping of chapter 1, also in this case gene ontology including transcription factors, hormonal regulation and signaling and sugar metabolism were considered. On chromosome 12, two genes possibly involved in the transport of sugars have been identified (VIT_12s0035g00290 and VIT_12s0035g00300), encoding respectively for ERD-like 6 and SFP1 carbohydrate transporters. In addition, a gene encoding an SPK-3 ASK1 kinase (VIT_12s0035g00310) has also been identified.

In the vicinity of VIT_16s0039g0700 there are two kinase receptors (VIT_16s0039g00610 and VIT_16s0050g01880) and a sigma factor not essential for transcription (VIT_16s0039g00660).
On chromosome 18 instead, in addition to the aforementioned VIT_18s0001g03540 for the transport of auxins there are, among the most interesting, a transcription factor of the CCCH-zinc finger family (VIT_18s0001g03670), the gene VIT_18s0001g03610 involved in the biosynthesis of auxins, the gene CRK1 involved in signal transduction (VIT_18s0072g00140), three genes for the biosynthesis of flavonoids (VIT_18s0001g03430, VIT_18s0001g03470, VIT_18s0001g03490) and the BRT1 gene (VIT_18s0072g00170), a receptor with activity of signal development and signal kinase stress-regulated genes, cell elongation and flowering.

Further approaches to understand how much LD actually extends into the genome at these regions and to identify causative SNP for the three validated signals of association can be now undertaken. In particular, the associated SNP located on chromosome 18, now validated, seems very promising, also given by its co-localization with regions previously identified as involved in the phenology regulation according to QTL studies (Zyprian et al., 2016, Carreno et al., 2012).

Conclusions

Climate change, which is becoming more and more influential, represent a modern landscape with which humans have to deal to maintain and preserve crops all over the world.

In this work, environmental effects on grapevine phenology driving winegrowers to seek new solutions to preserve quality of vineyards and their wines were discussed. Global warming is affecting the onset and duration of phenological events, with an acceleration in their timing in many areas (Jones et al., 2000, 2005; Duchêne et al., 2005; Tomasi et al., 2011; Ramos et al., 2019) and leading to a more rapid ripening phase with earlier harvesting period and missed accumulation of flavor components that require more time to store. High temperatures impact the formation of typical compounds of berry skin, like flavonoids and anthocyanins (Mori et al., 2007; Pastore et al., 2017; Yamane et al., 2006) and tannins. Moreover, when grapes are harvested with higher sugar concentrations, produced wines exhibit undesirable higher alcohol content.

One of the main breeding purposes in viticulture is to preserve quality peculiarities that are threatened by climate change, but the understanding of the genetic mechanisms underlying the main phenological processes in grapevine remains still poor known.

The genetic approaches used to understand complex traits of phenology and other qualitative traits that contribute to the quality of grapes are based on the research of new interesting natural alleles within large cultivars collections. By exploiting the genetic diversity within the species, it is possible to identify variants to understand phenological mechanisms and to select new traits of interest to adapt commercial cultivars through classical or newest breeding technologies.

Two different genetic approaches were used in this study based on the same genotyping tool, the Illumina bead chip *Vitis*18KSNP.

From the first approach, the QTL mapping approach, a large number of SNP markers were obtained by SNP chip hybridization which allowed the construction of a high-resolution genetic map and the identification of multiple QTL regions by using phenotypic data collected over two years. The high number of candidates found confirmed the suitability of SNP chip-based genotyping to identify genetic variants within the analyzed population, providing data in agreement with literature, beside new data (Duchêne et al., 2012; Guo et al., 2019). However, the use of this technology on a bi-parental population is limited, as it allows to analyze variants inherited only from the two parents and does not allow a wider overview on the QTLome of the entire species.

The second approach, a genome-wide association study, originally used the same genetic tool, the *Vitis*18K SNP chip, to generate markers dataset. In this case, the genetic diversity resource was represented by a core collection built from a large collection of unique genotypes of the *V. vinifera* species. By exploring such wide diversity association signals in genomic regions on chromosomes already mentioned in previous studies were identified (Zyprian et al., 2016; Duchêne et al., 2012; Costantini et al., 2008).

According to our data the association signals found in each of the three regions are not the causative ones. Further studies will explore the extension of these unique association signals and eventually confirm whether they are in LD with flanking genes involved in the phenotype. Our GWAS findings suggests that SNP chip might not be the more appropriate tool for GWAS in grapevine to identify all the small contributors of loci for complex traits like veraison and flowering time, because of its relatively low resolution. Indeed, it provided in our previous study just sparse association signals (see Pietro Delfino thesis and discussion therein), in agreement with the expected LD in grapevine and with previous findings (Flutre et al., 2022). Alternative genotyping tools with higher resolution are recommended. Despite of that the punctual association signal previously found were confirmed in this study by exploring a panel of genotypes with extreme phenotypes.

On the other hand, the use of germplasm collection makes possible to analyze and dissect at the same time more genetic regulators of the phenological processes

providing a wider overview on the genetic regulation of the trait in the species not limited to the characteristics of few parental cultivars.

Finally, we should also acknowledge that not shared QTLs were identified by the two approaches, even though for each of the two approaches some of the identified QTLs were in agreement with previous literature findings (on chromosome 19, 15, and 7 for the QTL mapping approach and on chromosome 18 for the GWAS approach). Overall, it can be state that the two genetic approaches used in this study represent therefore excellent complementary tools for the genetic research of highly complex traits like phenology and the dissection of its genetic regulations in grapevine.

Bibliography

Afoufa-Bastien D, Medici A, Jeauffre J, Coutos-Thévenot P, Lemoine R, Atanassova R, Laloi M. (2010) The *Vitis vinifera* sugar transporter gene family: phylogenetic overview and macroarray expression profiling. BMC Plant Biol. 2010 Nov 12;10:245. doi: 10.1186/1471-2229-10-245. PMID: 21073695; PMCID: PMC3095327.

Bacilieri R, Lacombe T, Le Cunff L, Di Vecchi-Staraz M, Laucou V, Genna B, Péros JP, This P, Boursiquot JM. (2013). Genetic structure in cultivated grapevines is linked to geography and human selection. BMC Plant Biol. 13. DOI: 10.1186/1471-2229-13-25

Barba P, Cadle-Davidson L, Harriman J, Glaubitz JC, Brooks S, Hyma K, Reisch B. (2014). Grapevine powdery mildew resistance and susceptibility loci identified on a high-resolution SNP map. Theor Appl Genet. 2014 Jan;127(1):73-84. doi: 10.1007/s00122-013-2202-x. Epub 2013 Sep 27. PMID: 24072208.

Belda I, Ruiz J, Esteban-Fernández A, Navascués E, Marquina D, Santos A, Moreno-Arribas MV. (2017) Microbial Contribution to Wine Aroma and Its Intended Use for Wine Quality Improvement. Molecules. 2017 Jan 24;22(2):189. doi: 10.3390/molecules22020189. PMID: 28125039; PMCID: PMC6155689.

Bianchi D, Caramanico L, Grossi D, Brancadoro L, Lorenzis GD. (2020) How Do Novel M-Rootstock (*Vitis* Spp.) Genotypes Cope with Drought? Plants. 2020; 9(10):1385. https://doi.org/10.3390/plants9101385

Black CA, Parker M, Siebert TE, Capone DL, Francis IL. (2015) Terpe- noids and their role in wine flavour: recent advances. Aust J Grape Wine Res 21:582–600

Boccacci P, Chitarra W, Schneider A, Rolle L, Gambino G. (2020) Single-nucleotide polymorphism (SNP) genotyping assays for the varietal authentication of 'Nebbiolo' musts and wines. Food Chem. 2020 May 15;312:126100. doi: 10.1016/j.foodchem.2019.126100. Epub 2019 Dec 26. PMID: 31901826.

Bowers J, Dangl GS, Vignani R, Meredith CP. (1996) Isolation and characterization of new polymorphic simple sequence repeat loci in grape (*Vitis vinifera* L.). Genome.;39(4):628-33. DOI: 10.1139/g96-080

Bowers J.; Dangel, G.; Meredith, C. (1999) Development and characterization of additional microsatellite DNA markers for grape. Am. J. Enol. Vitic. 50, 243-246.

Brault C., Segura V., This P., Le Cunff L., Flutre T., François P, Pons T., Jean, Péros JP., Doligez A. (2022) Across-population genomic prediction in grapevine opens up promising prospects for breeding, Horticulture Research, Volume 9, 2022, uhac041, https://doi.org/10.1093/hr/uhac041

Buttrose MS. (1974) Climatic factors and fruitfulness in grapevines. Horticultural Abstracts 44, 319–326.

Canaguier A, Grimplet J, Di Gaspero G, Scalabrin S, Duchêne E, Choisne N, Mohellibi N, Guichard C, Rombauts S, Le Clainche I, Bérard A, Chauveau A, Bounon R, Rustenholz C, Morgante M, Le Paslier MC, Brunel D, Adam-Blondon AF. (2017) A new version of the grapevine reference genome assembly (12X.v2) and of its annotation (VCost.v3). Genome Data. 2017 Sep 18;14:56-62. DOI: 10.1016/j.gdata.2017.09.002

Cao K, Zhou Z, Wang Q, Guo J, Zhao P, Zhu G, Fang W, Chen C, Wang X, Wang X, Tian Z, Wang L. (2016) Genome-wide association study of 12 agronomic traits in peach. Nat Commun. 2016 Nov 8;7:13246. DOI: 10.1038/ncomms13246.

Carbonell-Bejerano, P., Diago, MP., Martínez-Abaigar, J. et al. (2014) Solar ultraviolet radiation is necessary to enhance grapevine fruit ripening transcriptional and phenolic responses.BMC Plant Biol 14, 183. https://doi.org/10.1186/1471-2229-14-183

Carbonell-Bejerano, P., Rodríguez, V., Royo, C. et al. (2014) Circadian oscillatory transcriptional programs in grapevine ripening fruits. BMC Plant Biol 14, 78. https://doi.org/10.1186/1471-2229-14-78

Carreno I., Cabezas J.A., Martinez-Mora C., Arroyo-Garcia R., Cenis J.L., Martinez-Zapater L.M., Carreno J., Ruiz-Garcia L. (2015). "Quantitative genetic analysis of berry firmness in table grape (*Vitis vinifera* L.)". Tree Genetics & genomes 11, 818. DOI: 10.1007/s11295-014-0818-x

Carreño Ruiz I (2012) Identificación de regiones cromosómicas implicadas en el control genético de caracteres de interés para la mejora genética de la uva de mesa. Proyecto de investigación, PhD Thesis. Universidad de Murcia.

Cavallini E, Matus JT, 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. 2015;167(4):1448-1470. doi:10.1104/pp.114.256172

Chitwood DH., Ranjan A., Martinez C.C., Headland LR., Thiem T., Kumar R., Covington MF., Hatcher T., Naylor D.T., Zimmerman S., Downs N., Raymundo N., Buckler E.S., Maloof J.N., Aradhya M, Prins B., Li L., Myles S., Sinha N.R. (2014) "A modern ampelography: a genetic basis for leaf shape and venation patterning in grape". Plant Physiol., 164 (1), 259-272. DOI: 10.1104/pp.113.229708

Choi Y (2012) A Fast Computation of Pairwise Sequence Alignment Scores Between a Protein and a Set of Single-Locus Variants of Another Protein. In Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine (BCB '12). ACM, New York, NY, USA, 414-417. DOI: 10.1145/2382936.2382989

Cipriani G, Marrazzo MT, 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 Biol. 8: 1–13. DOI: 10.1186/1471-2229-8-127

Cipriani G, Spadotto A, Jurman I, Gaspero G Di, Crespan M, Meneghetti S, Frare E, Vignani R, Cresti M, Morgante M, Pezzotti M, Pe E, Policriti A, Testolin R. (2010). The SSR-based molecular profile of 1005 grapevine (*Vitis vinifera* L.) accessions uncovers new synonymy and parentages and reveals a large admixture amongst varieties of different geographic origin. Theor. Appl. Genet. 121: 1569–1585. DOI: 10.1007/s00122-010-1411-9

Coombe, B.G. (1995) Growth Stages of the Grapevine: Adoption of a System for Identifying Grapevine Growth Stages. Australian Journal of Grape and Wine Research, 1, 104-110. DOI: 10.1111/j.1755-0238.1995.tb00086.x

Crespan M, Migliaro D, Larger S, et al. (2020) Unraveling the genetic origin of 'Glera', 'RibollaGialla' and other autochthonous grapevine varieties from Friuli Venezia Giulia (northeastern Italy). Sci Rep 10:1–11. DOI: 10.1038/s41598-020-64061-w

Crespan M, Migliaro D, Larger S, et al. (2021) Grapevine (*Vitis vinifera* L.) varietal assortment and evolution in the Marche region (central Italy). OENO One 55:17–37. DOI: 10.20870/oeno-one.2021.55.3.4628

D'Incà E, Cazzaniga S, Foresti C, Vitulo N, Bertini E, Galli M, Gallavotti A, Pezzotti M, Battista Tornielli G, Zenoni S. (2021) VviNAC33 promotes organ de-greening and represses vegetative growth during the vegetative-to-mature phase transition in grapevine. New Phytol. 2021 Jul;231(2):726-746. doi: 10.1111/nph.17263. Epub 2021 Mar 16. PMID: 33567124; PMCID: PMC8251598.

D'Onofrio C, Tumino G, Gardiman M, et al. (2021) Parentage atlas of Italian grapevine varieties as inferred from SNP genotyping. Front Plant Sci 11:1–16. DOI: 10.3389/fpls.2020.605934

De Beukelaer, H., Davenport, G. F., and Fack, V. (2018). Core Hunter 3: flexible core subset selection. BMC Bioinformatics 19, 203. DOI: 10.1186/s12859-018-2209-z

De Lorenzis G, Chipashvili R, Failla O, Maghradze D (2015) Study of genetic variability in *Vitis vinifera* L. germplasm by high-throughput Vitis18kSNP array: the case of Georgian genetic resources. BMC Plant Biol 15:154. DOI: 10.1186/s12870-015-0510-9

De Lorenzis G, Mercati F, Bergamini C, et al. (2019) SNP genotyping elucidates the genetic diversity of *Magna Graecia* grapevine germplasm and its historical origin and dissemination. BMC Plant Biol 19:1–15. DOI: 10.1186/s12870-018-1576-y

De Lorenzis G, Squadrito M, Rossoni M, et al. (2017) Study of intra-varietal diversity in biotypes of Aglianico and Muscat of Alexandria (*Vitis vinifera* L.) cultivars. Aust J Grape Wine Res 23:132–142. DOI: 10.1111/ajgw.12247

Di Gaspero, G., Cipriani, G., Marrazzo, M.T. *et al.* (2005) Isolation of (AC)*n*-microsatellites in *Vitis vinifera* L. and analysis of genetic background in grapevines under marker assisted selection. *Mol Breeding* **15,** 11–20. DOI: 10.1007/s11032-004-1362-4

Di Guardo M, Micheletti D, Bianco L, Koehorst-van Putten HJ, Longhi S, Costa F, Aranzana MJ, Velasco R, Arús P, Troggio M, van de Weg EW. (2015) ASSIsT: an automatic SNP scoring tool for in- and outbreeding species. Bioinformatics. 2015 Dec 1;31(23):3873-4. doi: 10.1093/bioinformatics/btv446. Epub 2015 Aug 6. PMID: 26249809; PMCID: PMC4653386.

Doligez, A., Bertrand, Y., Farnos, M. et al. (2013) New stable QTLs for berry weight do not colocalize with QTLs for seed traits in cultivated grapevine (*Vitis vinifera* L.). BMC Plant Biol 13, 217. DOI: 10.1186/1471-2229-13-217

Duchêne E., Schneider C., (2005). Grapevine and climatic changes: a glance at the situation in Alsace. Agronomy for Sustainable Development, Springer Verlag/EDP Sciences/INRA, 25 (1), 93-99 DOI: 10.1051/agro:2004057.

Duchêne, E., Butterlin, G., Claudel, P., Dumas, V., Jaegli, N., and Merdinoglu, D. (2009). A grapevine (*Vitis vinifera* L.) deoxy-d-xylulose synthase gene colocates with a major quantitative trait loci for terpenol content. Theor. Appl. Genet. 118, 541–552. DOI:10.1007/s00122-008-0919-8.

Duchêne, É., Dumas, V., Butterlin, G. et al. (2020) Genetic variations of acidity in grape berries are controlled by the interplay between organic acids and potassium. Theor Appl Genet133, 993–1008. https://doi.org/10.1007/s00122-019-03524-9

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, et al. (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLOS ONE 6(5): e19379. DOI:10.1371/journal.pone.0019379

Emanuelli, **F**., Battilana, J., Costantini, L. et al. (2011) A candidate gene association study on muscat flavor in grapevine (*Vitis vinifera* L.). BMC Plant Biol 10, 241. DOI: 10.1186/1471-2229-10-241

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 Biol 13, 39. DOI: 10.1186/1471-2229-13-39

Fasoli M, Dal Santo S, Zenoni S, et al. (2012) The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell. 2012;24(9):3489-3505. doi:10.1105/tpc.112.100230

Flutre T, Le Cunff L, Fodor A, Launay A, Romieu C, Berger G, Bertrand Y, Terrier N, Beccavin I, Bouckenooghe V, Roques M, Pinasseau L, Verbaere A, Sommerer N, Cheynier V, Bacilieri R, Boursiquot JM, Lacombe T, Laucou V, This P, Péros JP, Doligez A. (2022) A genome-wide association and prediction study in grapevine deciphers the genetic architecture of multiple traits and identifies genes under many new QTLs. G3 (Bethesda). 2022 Apr 29:jkac103. doi: 10.1093/g3journal/jkac103

Fournier-Level, A., Le Cunff, L., Gomez, C., Doligez, A., Ageorges, A., Roux, C., et al. (2009). Quantitative genetic bases of anthocyanin variation in grape (*Vitis vinifera* L. ssp. sativa) berry: a quantitative trait locus to quantitative trait nucleotide integrated study. Genetics 183, 1127–39. DOI:10.1534/genetics.109.103929.

Ganal MW, Altmann T, Röder MS. (2009). SNP identification in crop plants. Curr. Opin. Plant Biol. 12: 211–217. DOI: 10.1016/j.pbi.2008.12.009

Garris, A., Clark, L., Owens, C., McKay, S., Luby, J., Mathiason, K., et al. (2009). Mapping of Photoperiod-induced Growth Cessation in the Wild Grape *Vitis riparia*. J. Am. Soc. Hortic. Sci. 134, 261–272. Available at: http://journal.ashspublications.org/content/134/2/261.abstract.

Grimplet J, Agudelo-Romero P, Teixeira RT, Martinez-Zapater JM, Fortes AM. (2016) Structural and Functional Analysis of the GRAS Gene Family in Grapevine Indicates a Role of GRAS Proteins in the Control of Development and Stress Responses. Front Plant Sci. 2016 Mar 30;7:353. doi: 10.3389/fpls.2016.00353. PMID: 27065316; PMCID: PMC4811876.

Grzeskowiak, L., Costantini, L., Lorenzi, S. et al. (2013) Candidate loci for phenology and fruitfulness contributing to the phenotypic variability observed in grapevine. Theor Appl Genet 126, 2763–2776. DOI: 10.1007/s00122-013-2170-1

Hayes MA, Davies C, Dry IB. (2007) Isolation, functional characterization, and expression analysis of grapevine (*Vitis vinifera* L.) hexose transporters: differential roles in sink and source tissues. J Exp Bot. 2007;58(8):1985-97. doi: 10.1093/jxb/erm061. Epub 2007 Apr 23. PMID: 17452752.

Houel C, Chatbanyong R, Doligez A, et al. (2015) Identification of stable QTLs for vegetative and reproductive traits in the microvine (*Vitis vinifera* L.) using the 18 K Infinium chip. BMC Plant Biol 15:205. DOI: 10.1186/s12870-015-0588-0

Huang XH, Han B. (2014) Natural variations and genome-wide association studies in crop plants. Annu Rev Plant Biol. 2014;65:531–51. DOI: 10.1146/annurev-arplant-050213-035715

Hyma KE, Barba P, Wang M, Londo JP, Acharya CB, Mitchell SE, 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. 2015 Aug 5;10(8):e0134880. doi: 10.1371/journal.pone.0134880. PMID: 26244767; PMCID: PMC4526651.

Imazio S, Maghradze D, de Lorenzis G, Bacilieri R, Laucou V, This P, Scienza A, Failla O. (2013) From the cradle of grapevine domestication: Molecular overview and description of Georgian grapevine (*Vitis vinifera* L.) germplasm. Tree Genet. Genomes 9: 641–658. DOI: 10.1007%2Fs11295-013-0597-9

International Grape Genome Browser Genescope https://www.genoscope.cns.fr/externe/GenomeBrowser/Vitis/

Jansen RC. (1993). Interval mapping of multiple quantitative trait loci. Genetics DOI: 135: 205-211.

Jansen RC., Stam P. (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics. 1994 Apr;136(4):1447-55. DOI: 10.1093/genetics/136.4.1447

Jones, G. V., White, M. A., Cooper, O. R., and Storchmann, K. (2005). Climate Change and Global Wine Quality. Clim. Change 73, 319–343. doi:10.1007/s10584-005- 4704-2.

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. 42, 348–354. DOI: 10.1038/ng.548.

Keller, M., Tarara, J. M., and Mills, L. J. (2010) Spring temperatures alter reproductive development in grapevines. Aust. J. Grape Wine Res. 16, 445–454. DOI: 10.1111/j.1755-0238.2010.00105.x

Klasen, J. R., Barbez, E., Meier, L., Meinshausen, N., Bühlmann, P., Koornneef, M., et al. (2016). A multi-marker association method for genome-wide association studies without the need for population structure correction. Nat. Commun. 7. DOI: 10.1038/ncomms13299.

Kobayashi, S., Goto-Yamamoto, N., and Hirochika, H. (2004). Retrotransposon-induced mutations in grape skin color. Science 304, 982. DOI:10.1126/science.1095011.

Koes R, Verweij W, Quattrocchio F. (2005) Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends Plant Sci. 2005 May;10(5):236-42. doi: 10.1016/j.tplants.2005.03.002. PMID: 15882656.

Kõressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M. (2018) Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics. 2018 Jun 1;34(11):1937-1938. DOI: 10.1093/bioinformatics/bty036

Kumar S, Stecher G, Tamura K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016 Jul;33(7):1870-4. DOI: 10.1093/molbev/msw054

Lai CP, Lee CL, Chen PH, Wu SH, Yang CC, Shaw JF. (2004) Molecular analyses of the Arabidopsis TUBBY-like protein gene family. Plant Physiol. 2004 Apr;134(4):1586-97. doi: 10.1104/pp.103.037820. Epub 2004 Apr 2. PMID: 15064372; PMCID: PMC419833.

Lander & Botstein, (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics DOI: 121: 185-199.

Laucou V, Launay A, Bacilieri R, Lacombe T, Adam-Blondon AF, Bérard A, Chauveau A, De Andrés MT, Hausmann L, Ibáñez J, Le Paslier MC, Maghradze D, Martinez-Zapater J, Maul E, Ponnaiah M, Töpfer R, Péros JP, Boursiquot JM. (2018). Extended diversity analysis of cultivated grapevine *Vitis vinifera* with 10K genome-wide SNPs. PLoS One 13: 1–27. DOI: 10.1371/journal.pone.0192540

Le Paslier M-C, Choisne N, Bacilieri R, Bounon R, Boursiquot J-MB, Brunel D, Di Gaspero G, Hausmann L, Lacombe T, Laucou V LA, Martinez-Zapater J, Morgante M, Raj P, PonnaiahM Q, Scalabrin S, Torres-Perez R. (2013). The GrapeReSeq 18k Vitis genotyping chip. In: IX International Symposium on Grapevine Physiology and Biotechnology. International Society for Horticultural Science, Abstr. Book p. 123.

Lipka A. E, Kandianis CB, Hudson ME, Yu J, Drnevich J, Bradbury PJ, Gore MA. (2015) From association to prediction: statistical methods for the dissection and selection of complex traits in plants. Curr Opin Plant Biol. 2015 Apr;24:110-8. DOI: 10.1016/j.pbi.2015.02.010.

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. Bioinformatics 28, 2397–2399. DOI: 10.1093/bioinformatics/bts444.

Ma, Z., Yang, S., Mao, J. et al. (2021) Effects of Shading on the Synthesis of Volatile Organic Compounds in 'Marselan' Grape Berries (*Vitis vinifera* L.). J Plant Growth Regul 40, 679–693. https://doi.org/10.1007/s00344-020-10123-2

Mackay T., Stone E., Ayroles J., (2009). "The genetic of quantitative traits: challenges and prospects". Nat Rev Genet 10, 565-577. DOI: 10.1038/nrg2612

Mamani, M., LÓpez, M., Correa, J., Ravest, G. and Hinrichsen, P. (2021), "Identification of stable quantitative trait loci and candidate genes for sweetness and acidity in tablegrape using a highly saturated single-nucleotide polymorphism-based linkage map". Australian Journal of Grape and Wine Research. https://doi.org/10.1111/ajgw.12497

Mardis ER (2008) The impact of next-generation sequencing technology on genetics. Trends Genet 24:133–141. DOI: 10.1016/j.tig.2007.12.007

Marrano A, Micheletti D, Lorenzi S, Neale D, Grando MS (2018). Genomic signatures of different adaptations to environmental stimuli between wild and cultivated Vitis vinifera L. Hortic Res. 2018 Jul 1;5:34. doi: 10.1038/s41438-018-0041-2. PMID: 29977570; PMCID: PMC6026492.

Mauricio R., (2001). "Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology". Nat Rev genet 2, 370-381. DOI: 10.1038/35072085

Mercati F, De Lorenzis G, Brancadoro L, et al. (2016) High-throughput 18K SNP array to assess genetic variability of the main grapevine cultivars from Sicily. Tree Genet Genomes 12:59. DOI: 10.1007/s11295-016-1021-z

Migicovsky, Z., Sawler, J., Gardner, K. et al. (2017) Patterns of genomic and phenomic diversity in wine and table grapes. Hortic Res 4, 17035. DOI: 10.1038/hortres.2017.35

Migliaro D, De Nardi B, Vezzulli S, Crespan M. (2017). An upgraded core set of 11 SSR markers for grapevine cultivar identification: the case of berry color mutants. Am. J. Enol. Vitic. 68: 496–498. DOI: 10.5344/ajev.2017.17048

Mira de Orduña Heidinger R., (2010). Climate change associated effects on grape and wine quality and production. Food Research International. 43. 1844-1855. 10.1016/j.foodres.2010.05.001. DOI: 10.1016/j.foodres.2010.05.001

Moreira FM, Madini A, Marino R, et al. (2011) Genetic linkage maps of two interspecific grape crosses (*Vitis* spp.) used to localize quantitative trait loci for downy mildew resistance. Tree Genet Genomes 7:153–167. DOI: 10.1007/s11295-010-0322-x

Mori, K., Goto-Yamamoto, N., Kitayama, M., and Hashizume, K. (2007). Loss of anthocyanins in red-wine grape under high temperature. J. Exp. Bot. 58, 1935–1945. DOI :10.1093/jxb/erm055.

Mosedale, J. R., Wilson, R. J., and Maclean, I. M. D. (2015). Climate Change and Crop Exposure to Adverse Weather: Changes to Frost Risk and Grapevine Flowering Conditions. PLoS One 10, e0141218. DOI: 10.1371/journal.pone.0141218

Mullins MG., Bouquet A., Williams LE. (1992) Biology of the grapevine - Cambridge: Cambridge University Press.

Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK, Prins B, Reynolds A, Chia JM, Ware D, Bustamante CD, Buckler ES. (2011) Genetic structure and domestication history of the grape. Proc Natl Acad Sci U S A. 2011 Mar 1;108(9):3530-5. DOI: 10.1073/pnas.1009363108.

Myles S, Chia JM, Hurwitz B, Simon C, Zhong GY, Buckler E, Ware D. (2010). Rapid genomic characterization of the genus *Vitis*. PLoS One 5. DOI: 10.1371/journal.pone.0008219

Nicolas P, Lecourieux D, Kappel C, Cluzet S, Cramer G, Delrot S, Lecourieux F. (2013) The basic leucine zipper transcription factor ABSCISIC ACID RESPONSE ELEMENT-BINDING FACTOR2 is an important transcriptional regulator of abscisic acid-dependent grape berry ripening processes. Plant Physiol. 2014 Jan;164(1):365-83. doi: 10.1104/pp.113.231977. Epub 2013 Nov 25. PMID: 24276949; PMCID: PMC3875815.

Nicolas, S.D., Péros, JP., Lacombe, T. et al. (2016) Genetic diversity, linkage disequilibrium and power of a large grapevine (*Vitis vinifera* L.) diversity panel newly designed for association studies. BMC Plant Biol 16, 74. DOI: 10.1186/s12870-016-

Ophir R, Pang X, Halaly T, Venkateswari J, Lavee S, Galbraith D, Or E. (2009) Gene-expression profiling of grape bud response to two alternative dormancy-release stimuli expose possible links between impaired mitochondrial activity, hypoxia, ethylene-ABA interplay and cell enlargement. Plant Mol Biol. 2009 Nov;71(4-5):403-23. doi: 10.1007/s11103-009-9531-9. Epub 2009 Aug 4. PMID: 19653104.

Palumbo MC, Zenoni S, Fasoli M, Massonnet M, Farina L, Castiglione F, Pezzotti M, Paci P. (2014) 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. 2014 Dec;26(12):4617-35. doi: 10.1105/tpc.114.133710. Epub 2014 Dec 9. PMID: 25490918; PMCID: PMC4311215.

Pastore, C., Zenoni, S., Tornielli, G.B. et al. (2011) Increasing the source/sink ratio in Vitis vinifera (cv Sangiovese) induces extensive transcriptome reprogramming and modifies berry ripening. BMC Genomics 12, 631. https://doi.org/10.1186/1471-2164-12-631

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics 28, 2537–2539. DOI: 10.1093/bioinformatics/bts460

Petrie, P. R., and Clingeleffer, P. R. (2005) Effects of temperature and light (before and after budburst) on inflorescence morphology and flower number of Chardonnay grapevines (*Vitis vinifera* L.). Aust. J. Grape Wine Res. 11, 59–65. DOI: 10.1111/j.1755-0238.2005.tb00279.x

Pilati S, Bagagli G, Sonego P, et al. (2017) Abscisic Acid Is a Major Regulator of Grape Berry Ripening Onset: New Insights into ABA Signaling Network. Front Plant Sci. 2017;8:1093. Published 2017 Jun 21. doi:10.3389/fpls.2017.01093

Raimondi S, Tumino G, Ruffa P, Boccacci P, Gambino G, Schneider A. (2020) DNA-based genealogy reconstruction of Nebbiolo, Barbera and other ancient grapevine cultivars from northwestern Italy. Sci Rep. 2020 Sep 25;10(1):15782. doi: 10.1038/s41598-020-72799-6. PMID: 32978486; PMCID: PMC7519648.

Rambla JL, Trapero-Mozos A, Diretto G, Rubio-Moraga A, Granell A, Gómez-Gómez L, Ahrazem O. (2016) Gene-metabolite net- works of volatile metabolism in Airen and tempranillo grape cultivars revealed a distinct mechanism of aroma bouquet pro- duction. Front Plant Sci 7:1619

Ramsay NA, Glover BJ. (2005) MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci. 2005 Feb;10(2):63-70. doi: 10.1016/j.tplants.2004.12.011. PMID: 15708343.

Rao, M.J., Wang, L. (2021) CRISPR/Cas9 technology for improving agronomic traits and future prospective in agriculture. Planta 254, 68. https://doi.org/10.1007/s00425-021-03716-y

Ruffa, P., Raimondi, S., Boccacci, P. et al. (2016) The key role of "Moscato bianco" and "Malvasia aromatica di Parma" in the parentage of traditional aromatic grape varieties. Tree Genetics & Genomes 12, 50. https://doi.org/10.1007/s11295-016-1006-y

Sadras, V. O., and Petrie, P. R. (2011) Climate shifts in south-eastern Australia: Early maturity of Chardonnay, Shiraz and Cabernet Sauvignon is associated with early onset rather than faster ripening. Aust. J. Grape Wine Res. 17, 199–205. DOI: 10.1186/1471-2229-14-99

Salmaso M, Malacarne G, Troggio M, et al. (2008) A grapevine (*Vitis vinifera* L.) genetic map integrating the position of 139 expressed genes. Theor Appl Genet 116:1129–1143. DOI: 10.1007/s00122-008-0741-3

Sefc K.; Regner, F.; Turetschek, E.; Glossl, J.; Steinkellner, H.; 1999: Identification of microsatellite sequences in *Vitis riparia* and their applicability for genotyping of different *Vitis* species. Genome, 42, 367 373.

Sun, L., Li, S., Jiang, J. et al. (2020) New quantitative trait locus (QTLs) and candidate genes associated with the grape berry color trait identified based on a high-density genetic map. BMC Plant Biol 20, 302 (2020). https://doi.org/10.1186/s12870-020-02517-x

Sun, R.-Z., Cheng, G., Li, Q., He, Y.-N., Wang, Y., Lan, Y.-B., et al. (2017). Light-induced Variation in Phenolic Compounds in Cabernet Sauvignon Grapes (*Vitis vinifera* L.) Involves Extensive Transcriptome Reprogramming of Biosynthetic Enzymes, Transcription Factors, and Phytohormonal Regulators. Front. Plant Sci. 8, 547. doi: 10.3389/fpls.2017.00547

Sunseri F, Lupini A, Mauceri A, De Lorenzis G, Araniti F, Brancadoro L, Dattola A, Gullo G, Zappia R, Mercati F. (2018). Single nucleotide polymorphism profiles reveal an admixture genetic structure of grapevine germplasm from Calabria, Italy, uncovering its key role for the diversification of cultivars in the Mediterranean Basin. Aust. J. Grape Wine Res. 24: 345–35 DOI: 10.1111/ajgw.12339

Tautz D, Renz M. (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res. 1984 May 25;12(10):4127-38. DOI: 10.1093/nar/12.10.4127

This P, Jung, A, Boccacci, P, Borrego, J, Botta, R, Costantini, L, Crespan, M, Dangl GS, Eisenheld, C, Ferreira-Monteiro, F, Grando, S, Ibanez, J, Lacombe, T, Laucou, V, Magalhaes, R, Meredith XP, Milani, N, Peterlunger, E, Regner, F, Zulini, L, Maul E (2004) Development of a standard set of microsatellite reference alleles for identification of grape cultivars. Theor Appl Genet 109:1148-1158 DOI: 10.1007/s00122-004-1760-3.

Thomas MR, Scott NS. (1993) Microsatellite repeats in grapevine reveal DNA polymorphisms when analysed as sequence-tagged sites (STSs). Theor Appl Genet. 1993 Sep;86(8):985-90. DOI: 10.1007/bf00211051

Tirumalai V, Swetha C, Nair A, Pandit A, Shivaprasad PV. (1992) miR828 and miR858 regulate VvMYB114 to promote anthocyanin and flavonol accumulation in grapes. J Exp Bot. 2019 Sep 24;70(18):4775-4792. doi: 10.1093/jxb/erz264. PMID: 31145783; PMCID: PMC6760283.

Tomasi, D., Jones, G. V., Giust, M., Lovat, L., and Gaiotti, F. (2011). Grapevine Phenology and Climate Change: Relationships and Trends in the Veneto Region of Italy for 1964-2009. Am. J. Enol. Vitic. 62, 329–339. DOI:10.5344/ajev.2011.10108.

Trenti, M., Lorenzi, S., Bianchedi, P.L. et al. (2021) Candidate genes and SNPs associated with stomatal conductance under drought stress in *Vitis*. BMC Plant Biol 21, 7 DOI: 10.1186/s12870-020-02739-z

Troggio M, Malacarne G, Coppola G, et al. (2007) A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (*Vitis vinifera* L.) anchoring Pinot Noir bacterial artificial chromosome contigs. Genetics 176:2637–2650. DOI: 10.1534/genetics.106.067462

Van Ooijen J.W. (1999), LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity, 83: 613-624. DOI: 10.1046/j.1365-2540.1999.00623.x

Van Ooijen (1992). Accuracy of mapping quantitative trait loci in autogamous species. Theor. Appl. Genet. DOI: 84: 803-811.

Varshney RK, Nayak SN, May GD, Jackson SA. (2009). Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends Biotechnol. 27: 522–530. DOI: 10.1016/j.tibtech.2009.05.006

Vervalle J. A., Costantini L., Lorenzi S.; Pindo M.; Mora R.; Bolognesi G.; Marini M.; Lashbrooke J. G.; Tobutt K. R.; Vivier M.; Roodt-Wilding R.; Grando M. S.; Bellin D. (2022) - A high- density consensus map for grapevine based on three mapping populations genotyped by the Vitis18K SNP chip - Theoretical and Applied Genetics. *In press*.

Vitis International Variety Catalogue (VIVC) - https://www.vivc.de

Vitulo N., Forcato C., Carpinelli EC., (2014) A deep survey of alternative splicing in grape reveals changes in the splicing machinery related to tissue, stress condition and genotype. *BMC Plant Biol*. 2014;14:99. Published 2014 Apr 17. DOI: 10.1186/1471-2229-14-99

Voorrips RE. (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. 2002 J Hered. Jan-Feb;93(1):77-8. DOI: 10.1093/jhered/93.1.77. PMID: 12011185.

Wang M, Vannozzi A, Wang G, et al. (2014) Genome and transcriptome analysis of the grapevine (Vitis vinifera L.) WRKY gene family. Hortic Res. 2014;1:14016. Published 2014 Mar 26. doi:10.1038/hortres.2014.16

Wang N, Zheng, Y, Xin H, Fang L, Li S (2013). Comprehensive analysis of NAC domain transcription factor gene family in *Vitis vinifera*. Plant Cell Rep. 32, 61-75.

Winkel-Shirley B. (2001) Flavonoid Biosynthesis. A Colorful Model for Genetics, Biochemistry, Cell Biology, and Biotechnology, Plant Physiology, Volume 126, Issue 2, June 2001, Pages 485–493, https://doi.org/10.1104/pp.126.2.485

Xu Y. (2010). Molecular breeding tools: markers and maps. Molecular plant breeding. Wallingford: CABI, p 21–58. DOI: 10.1079/9781845933920.0021

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 2012 Jun 18;13:134. DOI: 10.1186/1471-2105-13-134

Zyprian E, Ochßner I, Schwander F, Šimon S, Hausmann L, Bonow-Rex M, Moreno-Sanz P, Grando MS, Wiedemann-Merdinoglu S, Merdinoglu D, Eibach R, Töpfer R. (2016) Quantitative trait loci affecting pathogen resistance and ripening of grapevines. Mol Genet Genomics. 2016 Aug;291(4):1573-94. DOI: 10.1007/s00438-016-1200-5.

Appendix

Table 26. Complete list of 167 individuals genotyped with VVMD25 and VVMD28 SSR markers to confirm CsxC cross. In yellow, cases of self-pollination and discarded plants. In red, cases of unreadable sequencing data.

	VVMD25		VVMD28		
Sample	A1 (bp)	A2 (bp)	A1 (bp)	A2 (bp)	
Corvina	241	263	258	258	
Cabernet Sauvignon	239	249	234	236	
CsxC 1	239	241	234	258	*Cross*
CsxC 5	239	241	236	258	*Cross*
CsxC 6	239	263	234	258	*Cross*
CsxC 7	249	263	234	258	*Cross*
CsxC 8	241	249	236	258	*Cross*
CsxC 9	241	249	234	258	*Cross*
CsxC 10	239	263	236	258	*Cross*
CsxC 11	241	249	234	258	*Cross*
CsxC 12	239	263	234	258	*Cross*
CsxC 13	239	241	234	258	*Cross*
CsxC 14	239	241	234	258	*Cross*
CsxC 18	239	263	236	258	*Cross*
CsxC 19	239	263	236	258	*Cross*
CsxC 20	239	239	236	236	*Self*
CsxC 21	241	249	234	258	*Cross*
CsxC 22	249	263	236	258	*Cross*
CsxC 31	239	263	236	258	*Cross*
CsxC 32	239	241	234	258	*Cross*
CsxC 33	239	263	236	258	*Cross*
CsxC 34	239	263	236	258	*Cross*
CsxC 35	249	263	236	258	*Cross*
CsxC 36	249	263	236	258	*Cross*
CsxC 37	239	263	236	258	*Cross*
CsxC 38	239	241	236	258	*Cross*
CsxC 39	249	263	236	258	*Cross*
CsxC 40	239	263	236	258	*Cross*
CsxC 41	239	263	236	258	*Cross*
CsxC 42	239	263	236	258	*Cross*
CsxC 43	241	249	234	258	*Cross*
CsxC 44	241	249	236	258	*Cross*
CsxC 46	239	263	234	258	*Cross*
CsxC 47	239	263	236	258	*Cross*
CsxC 48	249	261	234	258	*Cross*

CsxC 49	249	263	234	258	*Cross*
CsxC 50	239	263	236	258	*Cross*
CsxC 51	241	249	234	258	*Cross*
CsxC 52	239	263	234	258	*Cross*
CsxC 53	239	263	236	258	*Cross*
CsxC 54	241	249	234	258	*Cross*
CsxC 55	241	249	234	258	*Cross*
CsxC 56	239	241	234	258	*Cross*
CsxC 57	239	263	234	258	*Cross*
CsxC 58	241	249	234	258	*Cross*
CsxC 59	241	249	234	258	*Cross*
CsxC 60	241	249	234	258	*Cross*
CsxC 61	239	263	234	258	*Cross*
CsxC 62	239	241	236	258	*Cross*
CsxC 63	239	263	234	258	*Cross*
CsxC 64	241	249	236	258	*Cross*
CsxC 65	249	263	234	258	*Cross*
CsxC 66	249	263			*Cross*
CsxC 67	241	249	236	258	*Cross*
CsxC 68	241	249	234	258	*Cross*
CsxC 72	239	241	236	258	*Cross*
CsxC 74	239	263	234	258	*Cross*
CsxC 75	249	263	236	258	*Cross*
CsxC 76	241	249	236	258	*Cross*
CsxC 77	241	249	234	258	*Cross*
CsxC 80	249	263	236	258	*Cross*
CsxC 81	249	263	234	258	*Cross*
CsxC 82	239	263	234	258	*Cross*
CsxC 83	249	263	234	258	*Cross*
CsxC 84	241	249	234	258	*Cross*
CsxC 85	241	249	234	258	*Cross*
CsxC 86	249	263	236	258	*Cross*
CsxC 87	239	263	236	258	*Cross*
CsxC 88	239	263	236	258	*Cross*
CsxC 89	241	249	234	258	*Cross*
CsxC 90	239	241	236	258	*Cross*
CsxC 91	239	263	234	258	*Cross*
CsxC 92	241	249	236	258	*Cross*
CsxC 93	241	249	236	258	*Cross*
CsxC 94	239	239	258	258	*Self*
CsxC 95	249	263	236	258	*Cross*
CsxC 96	239	241	236	258	*Cross*
CsxC 97	239	241	234	258	*Cross*
CsxC 98	239	241	236	258	*Cross*
CsxC 99	239	241	236	258	*Cross*

CsxC 100	239	263			*Cross*
CsxC 101	241	249			*Cross*
CsxC 102	239	263	236	258	*Cross*
CsxC 104	249	263	234	258	*Cross*
CsxC 105	241	249	234	258	*Cross*
CsxC 106	239	263	236	258	*Cross*
CsxC 107	239	241	234	258	*Cross*
CsxC 108	241	249	234	258	*Cross*
CsxC 109	239	263	236	258	*Cross*
CsxC 110	241	249	236	258	*Cross*
CsxC 111	239	263	234	258	*Cross*
CsxC 112	241	249	234	258	*Cross*
CsxC 113	249	263	234	258	*Cross*
CsxC 115	241	249	236	258	*Cross*
CsxC 117	239	241	234	258	*Cross*
CsxC 118	241	249	236	258	*Cross*
CsxC 119	249	249			*Self*
CsxC 120	249	263	236	258	*Cross*
CsxC 121	249	263	234	258	*Cross*
CsxC 122	239	241	236	258	*Cross*
CsxC 123	239	263	234	258	*Cross*
CsxC 124	249	263	234	258	*Cross*
CsxC 125	241	249	234	258	*Cross*
CsxC 126	239	263	234	258	*Cross*
CsxC 127	249	263	234	258	*Cross*
CsxC 128	241	249	234	258	*Cross*
CsxC 130	239	241	236	258	*Cross*
CsxC 132	239	263	234	258	*Cross*
CsxC 134	241	249	234	258	*Cross*
CsxC 135					*No data*
CsxC 137	241	249	234	258	*Cross*
CsxC 138	239	241	234	258	*Cross*
CsxC 139	239	241	234	258	*Cross*
CsxC 140	239	241	236	258	*Cross*
CsxC 141	239	263	236	258	*Cross*
CsxC 142	249	263	234	258	*Cross*
CsxC 143					*No data*
CsxC 144	239	263			*Cross*
CsxC 145	239	241	234	258	*Cross*
CsxC 146	239	241			*Cross*
CsxC 147	239	241	234	258	*Cross*
CsxC 148					*No data*
CsxC 149	239	241	236	258	*Cross*
CsxC 150	239	241	236	258	*Cross*
CsxC 155	241	241	234	234	*Self*

CsxC 156	241	249	236	258	*Cross*
CsxC 158					*No data*
CsxC 159	239	241	234	258	*Cross*
CsxC 161	241	249	236	258	*Cross*
CsxC 162					*No data*
CsxC 163	239	263	236	258	*Cross*
CsxC 164	241	241	258	258	*Self*
CsxC 165	239	263	236	258	*Cross*
CsxC 166					*No data*
CsxC 169	239	263	234	258	*Cross*
CsxC 171	241	249			*Cross*
CsxC 172	239	263			*Cross*
CsxC 173	241	249			*Cross*
CsxC 177	241	249			*Cross*
CsxC 179	249	263	236	258	*Cross*
CsxC 181	249	263	236	258	*Cross*
CsxC 184	241	249	236	258	*Cross*
CsxC 192	241	263	236	258	*Cross*
CsxC 194	239	241	234	258	*Cross*
CsxC 197	239	241	234	258	*Cross*
CsxC 198	239	263	236	258	*Cross*
CsxC 199	239	241	236	258	*Cross*
CsxC 200	241	249	236	258	*Cross*
CsxC 201	239	241	234	258	*Cross*
CsxC 202	239	241	236	258	*Cross*
CsxC 203	239	241	234	258	*Cross*
CsxC 205	239	263	236	258	*Cross*
CsxC 206	239	241	234	258	*Cross*
CsxC 207	239	241	236	258	*Cross*
CsxC 208	249	263	236	258	*Cross*
CsxC 210	249	263	236	258	*Cross*
CsxC 211	249	263	234	258	*Cross*
CsxC 212	249	263	236	258	*Cross*
CsxC 213	239	241	236	258	*Cross*
CsxC 214	239	249	234	258	*Cross*
CsxC 228	239	241	234	258	*Cross*
CsxC 229	241	249	234	258	*Cross*
CsxC 230	249	263	236	258	*Cross*
CsxC 232	239	241	236	258	*Cross*
CsxC 234	239	263	236	258	*Cross*
CsxC 235	241	249	236	258	*Cross*
CsxC 236	241	249	236	258	*Cross*
CsxC 237	239	241	236	258	*Cross*
CsxC 257	239	241	234	258	*Cross*

Table 27. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Budbreak.**

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Budbreak	*Gozzo Left 2020*	CP	14	61.368	14_26256344	3.96	17	3.13 - 3.96 - 3.18	53.004 - 76.763
			16	44.685	16_21090055	3.58	15.5	3.11 - 3.38- 3.25	43.266 - 50.531
		P1	14	88.435	14_29533991	3.18	13.9	2.21 - 3.18 - 2.89	81.312 - 89.854
		P2	7	0	7_888150	2.53	11.2	2,53	0
			14	43.552	14_25130458	2.18	9.7	1.71 - 2.18 - 2.06	42.843 - 47.848
			16	43.573	16_21090055	3.17	13.8	2.23 - 3.17 - 2.49	31.413 - 54.201
	Gozzo Right 2020	CP	7	54.706	Un_14300857	6.33	30.2	5.9 - 6.33 - 5.39	53.287 - 59.346
		P1	4	52.804	4_20532455	1.73	9.4	1,73	52,804
		P2	7	57.235	Un_14300857	6.16	29.5	5.8 - 6.16 - 5.31	56.526 - 59.363
	Gozzo Medium 2020	CP	5	0	5_1634822	3.33	13.7	3.33 - 3.31	0 - 0.894
			7	31.147	7_7430930	5.71	22.3	4.99 - 5.71 - 4.91	28.657 - 49.421
		P1	14	88.435	14_29533991	1.9	8.1	1.9 - 1.9 - 1.9	87.726 - 89.145
		P2	7	34.354	7_11129686	5.37	21.2	4.55 - 5.37 - 4.51	25.104 - 50.842
			8	5.684	8_1699376	2.15	9.1	1.78 - 2.15 - 2.1	4.975 - 7.843
			16	43.573	16_21090055	2.83	11.8	2.13 - 2.83 - 2.55	31.413 - 54.201
	Gozzo Left 2021	CP	7	38.985	7_13388305	4.91	19.6	4 - 4.91 - 4.2	33.635 - 51.334
		P1	12	60.772	12_22537667	2.27	9.6	1.76 - 2.27	50.696 - 60.772
			17	9.261	17_2236982	1.98	8.4	1.66 - 1.98 - 1.75	7.843 - 22.818
		P2	7	42.916	5_20326151	4.34	17.5	3.54 - 4.34 - 4.07	30.078 - 52.26
	Gozzo Right 2021	CP	7	27.945	mu_7_6484762	3.33	16.9	3.29 - 3.33 - 3.33	16.807 - 27.946
		P1	7	31.507	7_6966667	2.11	11	2.11 - 2.11 - 1.97	30.085 - 36.492
		P2	7	4.985	7_2350141	2.17	11.3	2.15 - 2.17 - 1.9	3.567 - 12.139
	Gozzo Medium 2021	CP	7	35.058	7_10409620	4.88	18.8	4.05 - 4.88 - 4.08	31.147 - 43.99
		P1	12	20.775	12_4064700	1.83	7.5	1.83 - 1.83 - 1.64	19.346 - 21.484
		P2	7	4.985	7_2350141	3.48	13.8	2.84 - 3.48 - 2.77	1.429 - 17.929
	Guerrieri-Rizzardi 2021	CP	7	31.147	7_7579292	3.85	14.2	3.32 - 3.85. 3.46	27.945 - 36.312
		P1	14	34.625	14_12696610	1.92	7.3	1,92	34.625
		P2	1	47.158	1_19975589	1.89	7.2	1,89	47.158
			7	25.812	7_7348129	2.28	9	1.95 - 2.28 - 2.03	4.276 - 32.925

Table 28. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Flowering.**

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Flowering	Gozzo Left 2020	CP							
		P1	16	21.433	16_14105808	2.17	11.6	1.87 - 2.17 - 1.69	20.015 - 32.884
		P2	7	5.694	7_2780583	2.24	12	2.21 - 2.24 - 1.82	0 - 13.139
	Gozzo Right 2020	CP	7	56.129	Un_12388826	3.67	19.7	3.34 - 3.67 - 3.22	29.369 - 58.814
		P1							
		P2	7	76.552	7_912711	2.02	11.4	1.83 - 2.02 - 2.02	56.526 - 76.552
			8	1.418	8_237949	1.94	11	1.89 - 1.94 - 1.92	0 - 9.271
	Gozzo Medium 2020	CP	7	31.502	7_7830862	3.46	15.3	3.03 - 3.46 - 3.02	21.399 - 33.989
		P1							
		P2	7	5.694	7_2780583	2.49	11.3	2.26 - 2.49 - 2.1	3.567 - 18.638
			8	5.684	8_1699376	2.21	10	2.03 - 2.21 - 1.89	0 - 7.843
	Gozzo Left 2021	CP	2	24.659	2_5557145	6.27	26	5.76 - 6.27 - 5.56	19.214 - 25.723
		P1	2	27.553	2_5557145	1.64	7.6	1,64	27.553
			4	47.777	Un_24135523	2.19	10	1.79 - 2.19 - 2.1	34.907 - 49.195
		P2							
	Gozzo Right 2021	CP	2	23.58	Un_12299152	3.72	19.1	3.01 - 3.72 - 2.93	21.069 - 26.787
			7	82.218	7_17591750	5.13	25.3	4.48 - 5.13 - 4.22	74.015 - 86.16
		P1	1	52.415	1_10279151	2.39	12.7	1.89 - 2.39 - 1.71	42.403 - 62.375
			7	60.05	Un_23415638	1.92	10.3	1.92 - 1.92 - 1.77	60.05 - 81.43
		P2	7	16.5	7_4810487	1.75	9.5	1.7 - 1.75 - 1.73	7.853 - 23.675
			7	85.865	7_17089269	2.32	12.3	1.83 - 2.32 - 1.78	55.107 - 95.167

	Peak position	Confidence interval

Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Flowering	Gozzo Medium 2021	CP	2	23.58	Un_12299152	6.51	25.1	5.54 - 6.51 - 6.15	19.283 - 24.723
			7	82.218	7_17591750	3.85	15.7	3.28 - 3.85 - 3.56	41.831 - 82.931
		P1	1	51.706	1_10215474	2.16	9.1	1.98 - 2.16 - 2.06	48.15 - 77.351
			4	48.486	Un_19063505	1.76	7.5	1.73 - 1.76 - 1.73	43.449 - 49.195
		P2	7	32.925	7_10943535	2.1	8.9	1.91 - 2.1 - 1.71	26.521 - 47.942
			7	57.944	Un_12338449	2.06	8.7	1.8 - 2.06	56.526 - 57.944
			7	86.574	7_17354354	1.95	8.3	1.95 - 1.93	86.574 - 88.733
	Guerrieri-Rizzardi 2021	CP	2	22.511	Un_15535432	3.16	14.5	3.03 - 3.16 - 3.16	21.069 - 22.511
			6	29.632	6_6613610	3.08	14.1	3.08	29.632
			10	29.773	10_1505500	6.39	27.1	5.79 - 6.39 - 5.67	27.646 - 33.327
			11	48.168	11_13989387	3.08	14.1	3.08 - 3.08	48.168 - 49.231
		P1	2	22.537	2_5279601	1.82	8.6	1.82 - 1.61	22.537 - 25.395
			3	54.932	3_17820622	2.11	9.9	1.83 - 2.11 - 2.11	54.223 - 58.462
			19	57.575	19_22424512	2.04	9.6	1.72 - 2.04 - 1.79	56.157 - 59.004
		P2	10	35.363	10_1358731	4.92	21.6	3.98 - 4.92 - 4.25	31.077 - 41.047
			17	15.019	17_4386786	2.03	9.6	1.73 - 2.03 - 1.81	8.626 - 16.448

Table 29. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Veraison.**

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Veraison	*Gozzo Left 2020*	CP	4	36.416	4_11091087	3.57	19	3.25 - 3.57 - 3.51	34.636 - 37.314
			5	22.838	5_6552770	3.99	21	3.23 - 3.99 - 3.43	14.658 - 33.194
			11	17.061	11_2709117	3.61	19.4	3.1 - 3.61 - 3.16	11.798 - 21.545
		P1	13	73.486	13_22875050	1.63	9.2	1.63 - 1.63 - 1.63	73.486 - 74.914
		P2	5	12.119	5_3707334	2.15	11.9	2.1 - 2.15 - 1.88	11.41 - 21.388
	Gozzo Right 2020	CP	1	38.226	1_9305946	3.87	22.2	3.46 - 3.87 - 3.35	37.421 - 58.333
		P1							
		P2	1	29.71	1_9242867	2.83	16.8	2.83 - 2.11	29.71 - 46.449
			7	67.936	Un_13943695	2.32	14	2.32 - 1.9	67.936 - 88.733
	Gozzo Medium 2020	CP	4	36.416	4_11091087	5.21	22.3	4.23 - 5.21 - 4.45	29.284 - 33.744
			5	22.838	5_6552770	4.74	20.5	3.75 - 4.74 - 3.97	20.627 - 35.549
		P1	11	35.72	11_181957	1.93	8.9	1,93	35.72 - 39.328
		P2	5	11.41	mu_5_3513954	2.35	10.8	1.85 - 2.35 - 1.71	4.996 - 23.547
	Gozzo Left 2021	CP	12	55.544	12_19007153	3.04	15.4	3.04 - 3.04	55.544 - 56.132
			18	61.156	18_18851921	3.23	16.2	3.22 - 3.23	59.466 - 61.156
			19	12.859	19_2797560	3.46	17.3	3.23 - 3.46 - 3.46	10.012 - 12.859
		P1							
		P2	17	57.111	17_14858602	2.16	11.2	2.05 - 2.16 - 2.16	56.402 - 57.111
			19	10.752	19_2534560	2.47	12.9	1.75 - 2.47 - 1.85	3.577 - 13.62

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Veraison	*Gozzo Right 2021*	CP	4	21.437	4_4718845	3.62	19.5	3.13 - 3.62 - 3.31	12.848 - 24.289
		P1	3	48.38	cn_3_15425566	2.52	14	1.64 - 2.52 - 1.9	32.111 - 54.932
			4	5.747	4_4538185	1.99	11.2	1.99 - 1.99 - 1.88	0.709 - 8.584
		P1	4	0	4_359553	1.97	11.1	1.97 - 1.83	0 - 21.421
	Gozzo Medium 2021	CP	19	10.721	19_2328994	3.92	16.4	3.5 - 3.92 - 3.07	8.223 - 14.293
		P1							
		P2	4	0	4_359553	2.2	9.6	2,2	0
	Guerrieri-Rizzardi 2021	CP	2	39.964	2_14315554	4.78	21.3	3.82 - 4.78 - 4.55	39.61 - 41.742
			5	2.328	5_1143888	3.99	19.1	3.4 - 3.99 - 3.41	1.789 - 19.272
			8	68.289	8_21030434	3.6	16.5	3.6 - 3.48	68.289 - 71.486
			17	37.462	17_8738687	4.51	20.2	3.69 - 4.51 - 4.5	31.231 - 38.351
		P1	2	47.494	2_17694758	2.44	11.5	1.75 - 2.44 - 1.87	41.34 - 48.203
			5	9.323	5_2758798	1.84	9.4	1.8 - 1.84	5.026 - 9.323
			8	65.4	8_22003454	2.85	13.3	2.62 - 2.85 - 2.85	63.272 - 66.109
			13	27.105	13_7469141	2.4	11.3	1.99 - 2.4 - 1.75	21.4 - 49.069
			17	32.82	17_5541199	2.92	13.6	1.98 - 2.92 - 1.94	26.385 - 39.213
		P2	2	37.44	2_14702692	2.19	10.4	1.75 - 2.19 - 2.06	25.311 - 39.578
			17	57.821	17_16780652	3.07	14.2	2.32 - 3.07	56.402 - 57.821

Table 30. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Interval Budbreak-flowering**

Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Interval Budbreak-flowering	*Gozzo Left 2020*	CP	14	61.368	14_26256344	3.02	16.3	3,02	61,368
		P1							
		P2	8	65.171	8_18467891	1.97	11.2	1.97 - 1.65	65.171 - 71.148
			14	42.843	14_25025034	1.67	9.4	1.67 - 1.64	42.843 - 44.261
	Gozzo Right 2020	CP	7	54.706	Un_14300857	4.11	22	3.66 - 4.11 - 3.39	43.99 - 55.417
			9	22.124	9_6263865	3.5	19.1	2.99 - 3.5 - 2.91	19.229 - 28.484
		P1	9	15.414	9_3074683	1.62	9.4	1.47 - 1.62 - 1.3	11.762 - 17.551
			19	1.428	19_556023	1.73	10	1.73 - 1.73 - 1.73	1.399 - 2.137
		P2	7	57.235	Un_14300857	3.17	17.5	2.31 - 3.17 - 2.46	32.925 - 59.363
	Gozzo Medium 2020	CP	7	54.706	Un_14300857	3.38	14.4	3.21 - 3.38 - 3.2	37.566 - 58.637
			8	35.242	8_12293738	3.41	15.2	3.41	35.242
			17	13.574	17_3349188	3.1	13.3	3.1	13.574
		P1	18	80.878	18_3406716	1.82	8	1.82	80.878 - 83.036
		P2	7	34.354	7_11129686	3.07	13.2	2.28 - 3.07 - 2.69	23.675 - 52.26
			14	43.552	14_25130458	2.02	8.9	1.83 - 2.02 - 2.02	42.843 - 44.261
	Gozzo Left 2021	CP	2	45.742		3,1	14.8	2.8 - 3.09 - 3.09	35.321 - 47.03
			7	50.445	7_469791	3.95	17.9	3.04 - 3.953.25	33.635 - 52.578
		P1	2	38.922	2_8538748	2.39	11.3	2.39 - 1.65	38.922 - 47.494
			3	16.404	3_3334282	1.65	7.9	1.65	16,404
			8	15.695	8_7736827	2.15	10.2	1.69 - 2.15 - 2.15	12.848 - 16.404
			12	60.772	12_22537667	2.88	13.4	2.21 - 2.88	50.696 - 60.772
		P2	7	52.26	7_297044	3.92	17.8	2.93 - 3.92 - 2.95	30.078 - 56.526
			10	36.073	10_1454013	1.7	8.2	1.7 - 1.68	36.073 - 38.2

Gozzo Right 2021	CP							
	P1							
	P2							
Gozzo Medium 2021	CP	2	47.03	2_17308518	2.91	12.2	2,91	47,03
		12	59.461	12_21015408	3.42	14.3	3.13 - 3.43 - 3.2	55.544 - 64.184
	P1	2	43.907	2_15113892	1.71	7.4	1.66 - 1.71 - 1.54	38.922 - 47.494
		12	60.772	12_22537667	1.99	8.5	1.79 - 1.99	19.346 - 60.772
	P2	7	4.985	7_2350141	2.48	10.5	2.35 - 2.48 - 1.81	1.429 - 57.944
Guerrieri-Rizzardi 2021	CP	14	30.552	14_16792373	3.86	17.6	3.86 - 3.11	30.552 - 35.219
		16	46.108	16_21175234	3.17	14.7	3.02 - 3.17 - 3.17	44.685 - 46.108
	P1	14	38.234	14_17080159	3.5	16.1	2.67 - 3.5 - 2.51	34.625 - 43.98
		18	61.522	18_2721473	2.18	10.4	1.91 - 2.18 - 2	57.236 - 83.036
	P2	2	6.048	2_1424334	2.09	9.9	1.7 - 2.09 - 2.09	1.772 - 6.048
		4	46.45	Vvi_5025	2.04	9.7	1.74 - 2.04 - 1.78	44.322 - 48.588
		16	7.864	lb_16_5137682	2.13	10.1	1.79 - 2.13 - 2.13	7.155 - 8.573

Table 31. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Interval Flowering-veraison.**

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Interval flowering-veraison	*Gozzo Left 2020*	CP							
		P1	3	57.779	3_18408810	1.78	10.3	1.73 - 1.78 - 1.77	57.07 - 58.462
			14	16.509	14_6301987	1.95	11.3	1.95 - 1.83	16.509 - 18.668
		P2	4	21.421	4_4817733	2.14	12.3	2.14 - 1.9	21.421 - 22.13
			5	12.119	5_3707334	2.07	11.9	1.92 - 2.07 - 2.07	11.41 - 12.119
			11	48.414	11_13503544	1.9	11		
	Gozzo Right 2020	CP	1	38.226	1_9305946	4.02	23.6	4.02 - 3.35	38.226 - 59.757
		P1							
		P2	1	29.71	1_9242867	2.99	18.1	2.99 - 2.22	29.71 - 47.158
	Gozzo Medium 2020	CP	4	36.416	4_11091087	3.98	18.4	3.45 - 3.98 - 3.42	29.284 - 38.916
			5	29.638	5_8455004	3.37	15.8	3.13 - 3.37 - 3	22.838 - 33.194
		P1							
		P2	5	11.41	mu_5_3513954	2.11	10.2	1.81 - 2.11 - 2	10.7 - 12.119
	Gozzo Left 2021	CP	11	40.263	11_7911752	3.47	17.7	3.05 - 3.47 - 3.08	35.95 - 43.352
			18	61.866	18_19399517	3.41	17.4	3.24 - 3.41 - 3.41	60.175 - 61.866
			19	11.79	19_2532513	3.54	17.8	2.99 - 3.54 - 3.47	6.059 - 12.859
		P1							
		P2	19	7.905	19_1975735	2,68	13.8	1.62 - 2.68 - 2.06	0 - 13.62

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Interval flowering-veraison	*Gozzo Right 2021*	CP	4	16.775	4_3898614	3.13	17.5	3,13	16.775
		P1	3	48.38	cn_3_15425566	2.58	14.6	2.45 - 2.58 - 1.79	44.019 - 54.932
		P2	4	0	4_359553	2.84	16	2.84 - 2.32	0 - 12.848
	Gozzo Medium 2021	CP	11	40.263	11_7911752	3.87	17.2	3.42 - 3.87 - 2,98	39.19 - 43.352
			17	58.54	17_16780652	3.46	15.4	3.34 - 3.46	57.831 - 58.54
			19	10.721	19_2328994	3.95	17.1	2.96 - 3.95 - 3.2	5.525 - 14.293
		P1	3	48.38	cn_3_15425566	2.59	11.6	1.89 - 2.59 - 1.8	44.019 - 54.223
			17	58.55	17_16420307	2.38	11.3	1.62 - 2.38	52.825 - 58.55
		P2	4	0	4_359553	2.01	9.1	2.01	0
			19	7.905	19_1975735	2.07	9.4	1.7 - 2.07 - 1.9	3.577 - 11.461
	Guerrieri-Rizzardi 2021	CP	5	11.8	mu_5_3513954	3.73	17.5	3.1 - 3.73 - 3.44	2.868 - 23.193
			8	68.289	8_21030434	3.79	17.8	3.79 - 3.31	68.289 - 72.195
			13	37.888	13_9338784	3.07	14.7	2.92 - 3.07	34.086 - 37.888
			17	37.462	17_8738687	5.81	26	5.32 - 5.81 - 5.66	36.382 - 38.531
		P1	5	9.323	5_2758798	2.68	14.1	2.15 - 2.68 - 2.09	4.317 - 14.329
			8	65.4	8_22003454	3.31	15.8	3.22 - 3.31 - 3.31	63.272 - 66.109
			13	27.105	13_7469141	2.22	10.9	1.87 - 2.22 - 2	21.4 - 37.075
			17	32.82	17_5541199	2.57	12.5	1.78 - 2.57 - 1.93	21.39 - 39.213
		P2	2	36.731	2_14315554	1.65	8.2	1.65 - 1.61	36.731 - 39.578
			17	36.42	17_8738687	2.52	12,2	1.88 - 2.52 - 2.07	34.261 - 57.821

Table 32. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Budbreak-veraison.**

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Interval budbreak-veraison	*Gozzo Left 2020*	CP							
		P1	3	57.779	3_18346272	1.8	10.3	1.8 - 1.8	57.779 - 58.462
			13	75.623	13_24351099	2.15	12.2	2.13 - 2.15	73.486 - 75.623
		P2	9	10.002	9_3642521	1.79	10.3	1.61 - 1.79 - 1.7	9.293 - 18.039
			11	18.791	11_3016530	2.08	11.8	2.08 - 1.85	18.791 - 21.69
	Gozzo2020	CP	1	38.226	1_9305946	4.28	24.5	4.28 - 3.47	38.226 - 47.654
		P1	12	28.587	12_6361641	1.91	11.8	1.91	28.587
		P2	1	29.71	1_9242867	3.4	20.1	3.4 - 2.69	29.71 - 31.129
	Gozzo Medium 2020	CP	4	36.416	4_11091087	3.27	14.8	3.01 - 3.27 - 3.22	34.636 - 37.314
			5	22.838	5_6552770	3.63	16.3	3.26 - 3.63 - 3.04	21.769 - 25.173
		P1	4	0	4_3181677	1.77	8.3	1,77	0
		P2	1	30.419	1_9532553	1.68	7.9	1.67 - 1.68	29.71 - 30.419
			5	11.41	mu_5_3513954	1.77	8.3	1.77 - 1.71	11.41 - 12.119
	Gozzo Left 2021	CP							
		P1							
		P2	10	15.04	Un_26075992	1.97	10.4	1.97 - 1.8	15.04 - 35.363

Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Interval budbreak-veraison	*Gozzo Right 2021*	CP	4	16.775	4_3898614	3.66	20.3	3.51 - 3.66 - 3.35	10.011 - 21.791
		P1	3	48.38	cn_3_15425566	2.13	12.4	1.92 - 2.13 - 2.13	44.019 - 52.794
			4	2.847	4_4384093	2.26	13.2	1.8 - 2.26 - 1.72	0 - 13.579
			12	15.79	12_3084744	1.8	10.6	1,8	15,79
		P2	4	4.297	4_785056	2.76	15.8	2.44 - 2.76 - 2	0 - 9.302
	Gozzo Medium 2021	CP	10	16.796	Un_35708800	3.12	13.4	3.12	16.796
			11	40.263	11_7911752	3.11	13.6	3.08 - 3-11	39.724 - 40.263
			17	58.54	17_16780652	3.02	13.2	3.01 - 3.02	57.831 - 58.54
		P1	17	58.55	17_16420307	2.05	9.7	1.63 - 2.05	56.391 - 58.55
		P2	10	15.04	Un_26075992	1.87	8.3	1.87 - 1.87	15.04 - 30.368
	Guerrieri-Rizzardi 2021	CP							
		P1	5	9.323	5_2758798	2.26	11.1	2.02 - 2.26 - 1.84	7.894 - 34.28
			12	19.346	12_3914492	2.4	11.4	1.89 - 2.4	15.79 - 19.346
			13	22.819	13_6439782	2.68	12.7	1.85 - 2.68 - 1.71	14.976 - 54.228
			14	38.234	14_17080159	2.65	12.6	1.86 - 2.65 - 2.1	34.625 - 45.409
		P2	18	30.101	18_7254936	2.12	10.2	1.9 - 2.12 - 1.97	28.68 - 33.71

Table 33. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Derivative maturity.**

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Derivative maturity	*Gozzo Left 2020*	CP	8	66.141	8_20127914	3.03	18.1	3.03 - 3.03	66.141 - 67.22
		P1	1	56.681	1_12284201	2.06	12.7	2.06 - 2.04	56.681 - 66.651
			8	61.134	8_20236545	2.94	17.6	2.29 - 2.94 - 2.2	58.287 - 66.109
		P2	9	38.467	ae_9_12948624	1.67	10.4	1.67	38.467
	Gozzo Right 2020	CP	18	57.323	18_1128395	3.89	23.5	3.2 - 3.89 - 3.65	56.259 - 67.855
		P1	18	61.522	18_2721473	3.04	18.9	2.6 - 3.04 - 2.15	51.521 - 66.549
		P2							
	Gozzo Medium 2020	CP	2	50.956	2_18360408	3.34	15.9	3.07 - 3.34 - 3.34	50.417 - 50.956
			18	56.791	18_18300437	3.44	16.3	3.44 - 3.42	56.791 - 61.866
		P1	1	66.651	1_16864011	2.29	11.2	2.05 - 2.29 - 2.25	65.233 - 70.228
			8	61.134	8_20236545	2.46	11.9	1.87 - 2.46 - 1.94	51.174 - 66.109
			18	64.39	18_19399517	3.21	15.3	2.21 - 3.21 - 2.57	51.521 - 66.549
		P2							

Trait	Dataset	Peak position						Confidence interval	
		Map	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Derivative maturity	*Gozzo Left 2021*	P2							
		CP	5	45.828	5_16883659	4.33	21.6	3.42 - 4.33 - 3.92	43.266 - 48.167
			11	50.65	11_17166783	3.28	16.8	3.25 - 3.28 - 2.8	47.813 - 53.653
			17	23.21	17_5098307	4	20.1	3.49 - 4 - 3.19	17.85 - 35.319
		P1	11	57.193	11_19123627	1.6	8.6	1.6 - 1.6 - 1.6	55.764 - 57.902
		P2	9	38.467	ae_9_12948624	1.88	10	1.88	38.467
			17	20.036	17_5098307	3.8	19.2	3.31 - 3.8 - 2.99	14.31 - 25.72
	Gozzo Right 2021	CP							
		P1	7	83.588	7_20233925	2.49	14.5	2.09 - 2.49 - 2.26	81.43 - 84.297
		P2	4	4.297	4_785056	2.04	12.1	1.94 - 2.04 - 1.8	0 - 17.844
	Gozzo Medium 2021	CP	4	39.27	4_13910890	3.51	15.2	3.22 - 3.51 - 3.5	34.636 - 39.27
			17	26.777	17_5423545	4.3	18.3	3.52 - 4.3 - 3.41	17.85 - 30.339
		P1							
		P2	4	29.284	4_6641297	2.24	10	2.01 - 2.24 - 1.82	27.866 - 42.163
			17	20.745	17_5136536	3.7	15.9	2.97 - 3.7 - 2.91	14.31 - 25.72
	Guerrieri-Rizzardi 2021	CP							
		P1							
		P2	17	36.42	17_8738687	2.63	12.2	1.68 - 2.63 - 1.86	25.72 - 44.294

Table 34. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Berry color.**

		Peak position						Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Berry color	*Gozzo Left 2020*	CP	2	39.61	2_14208969	9.69	68.1	9.39 - 9.69 - 9.69	39.252 - 41.742
		P1	8	15.695	8_7450052	1.96	20.7	1.96 - 1.96 - 1.96	15.695 - 16.404
		P2	2	37.44	2_14733683	3.2	31.5	2.46 - 3.2 - 3.2	33.884 - 39.578
			2	53.71	2_18360408	3.71	35.5	2.71 - 3.71 - 3.71	48.704 - 53.71
			7	0	Un_7469566	2.1	21.9	2.1 - 2.01	0 - 5.694
	Gozzo Right 2020	CP	2	39.61	2_14208969	14.4	68.7	14.4	39.61
		P1	2	43.907	2_15113892	2.19	16.2	1.65 - 2.19 - 1.83	41.769 - 48.203
			4	62.083	4_23154308	1.81	13.6	1.81	62.083
		P2	2	36.022	2_14208969	5.79	38.2	5.06 - 5.79 - 5.51	34.593 - 37.44
			17	56.402	17_13926824	2.67	20.5	1.89 - 2.67 - 2.5	44.294 - 57.821
	Gozzo Medium 2020	CP	2	39.61	2_14208969	17	67.9	17	39.61
			19	4.044	19_1490246	3.47	20.7	3.47 - 3.07	4.044 - 4.99
		P1	2	43.907	2_15113892	2.83	17.3	1.88 - 2.83 - 2.15	38.922 - 48.203
		P2	2	36.022	2_14208969	5.77	32.5	5.17 - 5.77 - 5.55	34.593 - 37.44
			17	56.402	17_13926824	2.61	16.7	1.84 - 2.61 - 2.48	44.294 - 57.821
	Gozzo Left 2021	CP	2	40.319	2_14702692	17.56	71.2	17.11 - 17.56 - 17.56	39.252 - 40.319
		P1	2	43.907	2_15113892	1.89	12.5	1.89 - 1.89	43.907 - 48.203
			12	8.625	12_1648693	2.17	14.3	2.17 - 1.9	8.625 - 10.763
		P2	2	37.44	2_14702692	9.06	47.4	8.13 - 9.06 - 9.06	36.022 - 37.44
			6	54.811	6_19075219	2.18	14.3	1.89 - 2.18 - 1.93	51.964 - 58.378
	GozzoRight 2021	CP	2	39.61	2_14208969	17.83	74	17.54 - 17.83 - 17.77	39.252 - 40.319

	P1	2	43.907	2_15113892	4.09	26.7	3.45 - 4.09 - 3.26	41.769 - 48.203
		4	13.579	4_6579179	2.17	15.1	1.88 - 2.17 - 2.01	8.584 - 30.652
		19	7.112	19_1671883	2.01	14.1	1.82 - 2.01 - 1.82	1.399 - 15.675
	P2	2	37.44	2_14702692	5.01	31.5	4.49 - 5.01 - 4.64	36.022 - 39.578
		6	31.961	6_11065703	1.91	13.4	1.91	31.961
		9	55.01	Vvi_10329	2.12	14.8	2.12	55.01
Gozzo Medium 2021	CP	2	40.319	2_14702692	26.75	77	26.45 - 26.75 - 26.75	39.252 - 40.319
	P1	2	43.907	2_15113892	4.47	21.8	3.91 - 4.47 - 3.76	41.769 - 48.203
		4	13.579	4_6579179	1.78	9.3	1.78 - 1.78	13.579 - 15.707
	P2	2	37.44	2_14702692	11.15	45.7	10.29 - 11.15 - 11.15	36.022 - 37.44
		6	53.382	6_19044643	2.13	11	1.9 - 2.13 - 1.95	31.961 - 56.24
		17	43.584	17_9990564	1.88	9.8	1.78 - 1.88 - 1.78	41.426 - 57.821
Guerrieri-Rizzardi 2021	CP	2	39.252	2_13398576	32.78	85.3	32.78 - 32.16	39.252 - 40.319
	P1	2	43.907	2_15113892	4.81	21.7	3.81 - 4.81 - 3.93	40.34 - 48.203
	P2	2	37.44	2_14702692	7.4	31.2	7.15 - 7.4 - 6.83	36.022 - 39.578
		6	51.964	6_18357854	3.2	15	2.95 - 3.2 - 2.29	51.254 - 56.949
		17	57.821	17_16780652	2.58	12.2	2.35 - 2.58	54.264 - 57.821

Table 35. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Polyphenol content.**

			Peak position				Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Polyphenol content	*Gozzo Left 2020*	CP							
		P1	3	36.429	3_9706422	2.15	42.4	2,15	36,429
			4	22.09	4_13910890	1.74	36	1.74 - 1.74	22.09 - 28.514
		P2	2	28.159	2_8329307	1.91	38.7	1.91	28.159
	Gozzo Right 2020	CP	7	69.369	Un_24730183	3.82	46.6	3.82 - 3.82	69.369 - 69.724
			9	7.524	9_2561793	3.37	44.7	3.37	7.524
		P1	3	52.794	3_17217402	2.17	30	1.69 - 2.17 - 2.16	48.38 - 58.462
			16	0	16_141454	2.25	30.9	2.25 - 1.85	0 - 21.433
		P2							
	Gozzo Media 2020	CP	3	39.108	3_9524678	3.29	33.6	3.29 - 3.13	39.108 - 45.339
		P1	3	36.429	3_9706422	3.25	33.3	2.33 - 3.25 - 2.25	32.111 - 44.729
			4	24.248	4_14366795	1.94	21.5	1.94 - 1.82	24.248 - 28.514
			13	39.234	lb_13_13395387	2.21	24.1	2.21 - 1.82	39.234 - 51.228
			14	41.081	14_18944223	2.47	26.4	2.47 - 2.12	41.081 - 52.553
		P2							
	Gozzo Left 2021	CP	3	49.016	3_360091	3.56	34.3	3.33 - 3.56 - 3.56	39.108 - 49.016
			8	61.44	8_18647348	4.25	44	3.64 - 4.25	56.44 - 61.44
		P1	3	44.729	3_360091	3.23	31.7	2.99 - 3.23 - 3.23	36.429 - 44.729
			4	19.962	4_10913217	2.22	23	2.22 - 1.99	19.962 - 22.09
		P2							
	Gozzo Right 2021	CP							

	P1	4	39.882	4_19850210	2.23	26.1	2.23 - 2.23	39.882 - 59.215
		8	37.616	8_12974771	1.9	22.6	1.82 - 1.9 - 1.9	30.93 - 40.464
		16	10.034	16_3142697	2.06	24.3	2.06 - 2.06	10.034 - 4.3
	P2							
	CP							
Gozzo Media 2021	P1	3	36.429	3_9706422	3.87	29.5	3.87 - 3.38	36.429 - 44.729
		4	19.962	4_10913217	2.51	20.3	2.51 - 2.33	19.962 - 22.09
		8	50.465	8_16841839	2.27	18.6	2.13 - 2.27 - 2.27	49.756 - 55.44
	P2	2	34.593	2_13395493	2.21	18.1	1.78 - 2.21 - 1.84	0 - 37.44

Table 36. Complete list of QTL regions targeted from **Interval Mapping** on the trait **Anthocyanin content.**

			Peak position				Confidence interval	
Trait	*Dataset*	*Map*	*LG*	*Position (cM)*	*Marker*	*LOD score*	*% var*	*LOD*	*Position*
Anthocyanin content	*Gozzo Left 2020*	CP	2	33.184	2_8329307	5.19	73.8	5.19 - 4.88	33.184 - 33.898
			8	3.,59	8_12974771	4.54	68.7	4.24 - 4.54 - 4.19	35.242 - 40.434
			10	28.355	10_1353542	3.67	60.9	3.67 - 3.64	28.355 - 32.615
		P1	4	29.223	4_16652502	1.81	37.1	1.81 - 1.81	29.223 - 32.07
			8	37.616	8_12974771	2.72	50.1	2.72 - 2.72	37.616 - 46.908
		P2	2	28.159	2_8329307	4.59	69.1	4,59	28.159
	Gozzo Right 2020	CP	2	23.225	Un_33081919	4.14	49.4	3.6 - 4.14 - 4.14	13.578 - 24.659
			5	11.8	mu_5_3513954	6.39	65.1	6.11 - 6.39 - 5.82	5.278 - 16.081
		P1							
		P2	5	4.996	5_298150	3.39	42.7	3.39 - 3.15	4.996 - 7.832
	Gozzo media 2020	CP	2	35.678	2_9135820	3.83	37.9	3.56 - 3.83 - 3.43	35.321 - 39.964
			5	5.278	5_298150	3.84	38	3.84 - 3.3	5.278 - 18.563
			19	3.217	19_1013861	3.43	34.8	3.18 - 3.43 - 3.27	0 - 6.059
		P1	14	41.081	14_18944223	2.27	24.6	2.27 - 2.27	41.081 - 45.409
			19	14.965	19_3090435	1.89	21	1.89 - 1.8	14.965 - 18.574
		P2	5	4.996	5_298150	2.35	25.4	2.35 - 1.95	4.996 - 7.832
	Gozzo Left 2021	CP	7	1.418	7_159555	5.71	50	5.71 - 5.71	1.418 - 2.128
			16	0	16_141454	3.6	36.4	3.6 - 3.36	0 - 7.186
		P1	7	1.418	7_159555	5.51	48.7	5.51 - 5.51	1.418 - 2.128
			16	0	16_141454	2.85	29.2	2.85 - 2.41	0 - 21.433
		P2	2	0	2_266133	1.94	21.1	1.94 - 1.92	0 - 9.605

		4	28.575	4_6571539	2.05	22	1.96 - 2.05 - 2.05	27.866 - 29.993
Gozzo Right 2021	CP	2	36.392	2_9223164	4.83	49	4.05 - 4.83 - 4.65	33.184 - 39.964
	P1							
	P2	2	34.593	2_13395493	3.89	41.9	2.79 - 3.89 - 2.56	27.449 - 39.578
Gozzo media 2021	CP	2	36.392	2_9223164	4.79	35.1	4.03 - 4.79 - 3.92	33.184 - 40.319
		11	44.699	11_9855734	3.9	30	3.18 - 3.9 - 3.45	39.724 - 46.734
		16	13.929	16_8004504	3.38	26.3	3.38 - 3.01	13.929 - 21.433
	P1	1	84.164	1_22498675	2.27	19.6	2.27 - 2.24	84.164 - 7.418
		5	37.931	5_11363177	1.83	15.2	1.83 - 1.83	37.931 - 65.973
		16	15.009	16_8004504	3.12	24.5	3.12 - 2.62	15.009 - 21.433
	P2	2	34.593	2_13395493	4.5	33.4	3.71 - 4.5 - 3.83	29.587 - 37.44

Table 37. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Budbreak.**

							Confidence interval	
	Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Position	LOD
Budbreak	*Gozzo 2020*	P1	4	41.311	4_20165673	3.62	11.2	39.882-41.311	3.15-3.62
			4	52.804	4_20532455	4.51	12.2	52.804	4.51
			6	39.753	6_8334007	3.44	13.7	39.044-36.753	2.56-3.44
			6	49.046	6_13942066	3.24	18.3	48.336-49.046	2.49-3.24
			6	66.036	6_20283151	3.17	19.1	65.327-66.036	3.08-3.17
			7	2.128	Un_7428334	2.23	3.7	2.128	2.23
			7	11.43	7_1403201	2.47	5.8	11.43	2.47
		P2	5	2.128	5_440666	4.3	7.7	2.128	4.3
			5	5.705	5_2217503	3.89	6.9	5.705	3.89
			5	24.975	5_8389529	2.27	3.9	23.547-24.975	1.99-2.27
			7	27.231	7_7830862	4.3	7.6	26.521-27.231	4.3
			8	26.427	8_11100449	6.19	11.5	25.718-26.427	4.82-6.19
			14	45.69	14_25632777	3.81	6.7	45.69	3.81
			16	51.688	16_21396785	8.93	17.7	46.711-51.688	8.56-8.93
			19	52.497	19_18762848	2.42	4.1	49.597-52.497	2.15-2.42
	Gozzo 2021	P1	2	33.238	2_7300456	5.85	10.4	33.238	5.85
			2	38.922	2_8538748	6.08	10.6	38.922	6.08
			12	10.053	12_1792973	2.24	5.1	10.053	2.24
			12	35.751	12_7196960	4.6	0.2	35.042-35.751	3.67-4.6
			12	60.772	12_22537667	3.4	8.6	57.121-60.772	2.93-3.4
			18	51.521	18_13442809	2.55	6.4	51.521	2.55
			18	80.878	18_3406716	2.13	5.3	80.878	2.13
		P2	7	11.43	7_3841674	5.49	12.6	9.272-11.43	5.24-5.49

		10	17.939	Un_33535896	5.24	14.2	16.04-17.939-19.939	5.02-5.24-5.16
		10	30.368	Un_26047676	5.49	14.5	30.368	5.49
		16	17.876	16_13062724	3.46	7.5	14.977-17.876	2.94-3.46
		17	20.036	17_5098307	1.87	3.9	17.877-20.036	1.62-1.87
		19	41.04	19_8394255	3.57	7.8	41.04	3.57
		19	52.497	19_18762848	3.98	8.8	49.597-52.497	3.76-3.98
	P1	14	34.625	14_12696610	1.92	7.3	34.625	1.92
Guerrieri-Rizzardi 2021	P2	5	23.547	5_7191560	7.82	14.6	22.388-23.547	6.38-7.37
		5	29.251	5_9640285	7.75	15.3	28.542-29.251	7.03-7.75
		7	32.925	7_10943535	7.6	14.2	31.496-32.925-33.925	5.15-6.93-6.81
		7	62.909	7_142105	2.62	4.6	62.909	2.62
		9	9.293	9_3405399	2.62	4.7	6.393-9.293	2.12-2.62
		10	0	10_208152	6.53	12.5	0	6.53
		13	44.553	13_15159986	2.88	4.9	44.553	2.88
		13	49.569	13_17985484	4.81	8.5	48.411-49.569	4-4.66
		13	58.184	13_2611864	1.97	3.3	58.184	1.97
		16	29.995	16_17636753	3.28	5.9	28.566-29.995	2.73-3.28
		18	17.094	Vvi_221	2.39	4.4	16.385-17.094	2.34-2.39
		18	72.375	18_25839298	4.04	7.4	72.375	4.04
		18	75.243	18_26702868	4.59	8.5	75.243	4.59

Table 38. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Flowering.**

Flowering

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	15	71.721	15_17344332	2.25	15.2	71.721	2.25
		15	75.278	15_17651852	2.15	14.8	75.278	2.15
	P2	7	5.694	7_2780583	1.89	6.1	4.985-5.694	1.73-1.89
		8	5.684	8_1699376	1.9	6	4.975-5.684	1.58-1.9
		10	4.266	Un_10557449	3.95	13.1	3.49-3.95	3.49-3.95
		10	46.012	10_5482314	2.57	8.2	46.012	2.57
		11	34.073	11_5754297	2.09	6.6	33.364-34.073	1.77-2.09
		11	41.186	11_8115351	1.93	6	40.476-41.186	1.93
Gozzo 2021	P1	1	75.922	1_20641842	3.77	9.4	75.213-75.922-76.922	3.29-3.77-3.71
		2	27.553	2_5557145	4.11	10.3	3.22-4.11	25.395-27.553
		2	38.922	2_8538748	2.44	5.9	38.922	2.44
		4	47.777	Un_24135523	2.92	7.1	47.777	2.92
		5	11.482	5_2862877	2.15	5.1	11.482	2.15
		7	56.494	Un_17857429	2.32	5.6	55.065-56.494	2.22-2.32
		15	30.196	15_12175417	5.1	13.1	29.487-30.196	5.1
		15	40.260	15_13737462	2.57	6.2	38.832-40.260-41.260	2.55-2.57-2.26
	P2	7	57.944	Un_12338449	2.87	5.1	57.944	2.87
		7	58.653	Un_12454323	2.07	3.7	58.653	2.07
		13	49.569	13_17985484	3.45	6.3	48.411-49.596	2.52-3.45
		13	53.178	13_18547627	3.5	6.4	53.178-55.178	3.5-2.71
		15	0	15_150741	2.93	5.3	0	2.93
		15	22.186	15_13445497	3.17	5.8	21.757-22.186	3.08-3.17
		16	0	16_834573	2.32	4.1	0	2.32
		16	10.002	cn_16_6630270	4.64	8.7	10.002	4.64
		17	0	Un_21136778	4.24	7.9	0	4.24

		17	14.310	17_4315095	2.39	4.3	14.310	2.39
		17	30.705	17_7469345	3.47	6.3	29.996-30.705	3.47
		19	3.577	19_1722935	2.52	4.5	2.868-3.577-4.577	2.24-2.52-2.24
		19	41.749	19_8744620	3.38	6.1	41.040-41.749	2.2-3.38
		19	52.497	19_18762848	3.83	7.1	50.597-52.496	2.89-3.83
Guerrieri-Rizzardi 2021	P1	2	20.410	2_5105885	2.05	3	2.05	20.410
		6	0	6_164244	7.41	12.7	0	7.41
		6	2.127	6_819567	9.71	17.6	1.418-2.127	9.71
		11	4.996	11_1145276	1.87	2.7	4.286-4.996	1.41-1.87
		11	30.662	11_5361364	4.47	7	30.662	4.47
		11	52.187	11_18226816	2.77	4.1	52.187	2.77
		12	12.922	12_2348601	5.27	8.4	12.922	5.27
		12	20.775	12_4064700	5.2	8.3	20.346-20.775	4.91-5.2
		18	24.341	18_4702061	2.12	3.1	23.912-24.341	1.97-2.12
		18	30.014	18_6409561	8.48	14.7	30.014	8.48
		18	31.433	18_7002635	6.82	11.3	31.396-31.433	6.80-6.82
		19	59.004	19_22531421	3.09	4.7	59.004	3.09
	P2	2	16.029	2_4166541	4.2	8.4	14.870-16.029	3.86-4.2
		2	17.447	2_4431200	4.47	9.1	17.447	4.47
		10	0	10_208152	4.79	8.5	4.79	0
		10	38.200	10_2730976	10.61	22	38.200	10.61
		11	0	11_116731	3.13	5.3	0	3.13
		11	49.123	11_13989387	3.14	5.4	48.414-49.123	2.52-3.14
		15	35.796	15_15813590	4.64	8.2	33.637-35.796	4.43-4.64
		17	2.128	17_806079	3.31	5.7	1.418-2.128	2.43-3.31
		17	8.626	17_1572984	4.1	7.2	4.975-8.626	2.09-4.10
		17	46.431	17_10796830	3.09	5.3	46.431	3.09
		17	57.111	17_16347597	6.85	12.9	57.111	6.85

Table 39. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Veraison.**

Veraison

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	2	1.772	2_1579733	2.37	4	1.772	2.37
		3	58.462	3_19209269	2.08	3.5	57.779-58.462	2.07-2.08
		4	5.747	4_4538185	7.63	14.5	2.847-5.747	5.91-7.63
		4	41.311	4_20165673	7.52	12.8	39.882-41.311	4.66-7.52
		4	58.506	4_22566230	7.01	11.6	56.384-58.506	5.78-7.01
		6	34.727	6_7306913	1.66	3	34.018-34.727-36.727	1.22-1.66-1.32
		6	66.036	6_20283151	5.23	11.4	65.327-66.036	5.13-5.23
		12	15.79	12_3038128	5.99	8.5	15.79	5.99
		12	21.484	12_4425553	6.2	8.2	21.484-21.484	6.13-6.2
		14	41.081	14_18999008	3.22	3.7	41.081-43.081	3.22-3.07
		14	73.47	14_25700297	3.1	3.3	72.041-73.47	2.84-3.10
		15	0	15_150741	2.86	3.9	0	2.86
		15	25.072	15_11426252	5.65	8.2	22.914-25.072	2.38-5.65
		15	45.266	15_14544654	5.2	7.9	44.557-45.226	5.2
		15	51.69	15_15015427	3.9	5.9	50.261-51.69	3.36-3.9
		19	25.351	19_7326716	2.73	4.6	21.574-24.351-26.351	2.37-2.73-2.56
		19	31.108	19_8394255	1.64	2.6	30.679-31.108	1.57-1.64
	P2	2	27.449	2_7970900	3.26	7.5	26.021-27.449	3.04-3.26
		4	21.421	4_4817733	4.8	11.6	21.421	4.8

		5	11.41	mu_5_3513954	4.17	10.1	11.41	4.17
		10	40.338	10_3594848	5.86	14.4	39.629-40.338	4.32-5.86
		10	51.07	li_10_5830295	6.49	16.1	48.924-51.07	5.05-6.49
		11	4.318	11_1073005	2.19	4.9	4.318	2.19
		17	0	Un_21136778	4.51	10.8	0	4.51
		17	8.626	17_1572984	4.46	10.6	5.975-8.626	3.46-4.46
		19	5.006	19_1792021	3.68	9.2	5.006-7.006	3.68-3.13
Gozzo 2021	P1	3	52.794	3_17217402	3.78	9.7	48.38-52.79-53.794	3.33-3.78-3.46
		4	6.456	4_4718845	2.73	6.7	6.456	2.73
		4	42.74	Un_24914324	5.83	16.2	41.311-42.74	4.76-5.83
		4	56.348	4_21965405	6.39	18.2	56.348	6.39
		15	35.223	15_13131764	2.37	6.1	35.223	2.37
		15	38.832	15_13467733	2.42	6.2	38.832	2.42
		15	67.393	15_17122573	2.03	5	66.684-67.393-69.393	1.94-2.03-1.88
		17	0	Un_21136778	2.14	5.5	0	2.14
		17	6.414	17_1307099	3.52	9.8	5.705-6.414	3.22-3.52
		17	20.681	17_4167305	2.56	7	19.252-20.681	2.37-2.56
		17	58.55	17_16347597	3.01	8.2	56.391-58.55	2.35-3.01
	P2	4	0	4_359553	3.29	5.2	0-2	3.29-3.06
		7	67.936	Un_13943695	3.7	6.5	65.036-67.936	2.4-3.7
		7	79.451	lb_7_15393500	6.35	12.2	78.552-79.451	5.73-6.35
		7	94.458	7_18466450	2.25	3.4	93.029-94.458	1.94-2.25
		7	97.305	7_19894211	5.44	9.4	95.876-97.305	4.58-5.44
		9	0	9_137176	5.28	10.2	0	5.28

		9	6.393	9_3142465	4.94	9.4	6.393-7.393	4.94-4.33
		10	15.04	Un_26075992	2.29	4	13.389-15.04	1.74-2.29
		10	38.2	10_2730976	4.47	8	37.491-38.3-39.2	4.47-3.92
		10	51.07	li_10_5830295	5.67	10.5	48.924-51.07	3.5-5.67
		10	65.785	10_7615145	3.69	6.7	65.785	3.69
		11	45.546	be_1_18883369	4.1	7.4	42.895-45.546	3.24-4.1
		11	51.251	11_16932395	4.2	7.6	51.251	4.2
		16	5.027	16_2545201	2.34	3.8	4.318-5.027	2.04-2.34
		16	10.711	16_8004504	2.1	3.4	10.002-10.711	1.89-2.1
		16	51.688	16_21396785	1.88	3.2	51.688	1.88
		19	0	19_75475	2.48	5.1	0	2.48
		19	5.006	19_1792021	2.33	4.9	5.006-6.006	2.33-2.9
Guerrieri-Rizzardi 2021	P1	2	46.065	2_17340911	10.14	21	43.907-46.065	9.44-10.14
		8	66.109	Un_17625284	2.92	4.8	66.109	2.92
		11	23.56	11_4101095	3.72	6.3	22.4-23.56	3.08-3.72
		11	29.952	11_5323270	4.1	7.1	28.524-29.952	3.21-4.1
		12	32.915	SNP569_266	7.11	13.4	32.915	7.11
		12	52.855	12_17505870	6.2	11.3	52.855	6.2
		13	16.395	13_5480574	2.57	4.3	16.395	2.57
		14	56.967	14_22387615	4.25	7.2	56.967	4.25
		14	70.622	14_25536444	2.93	4.8	70.622	2.93
		15	25.072	15_11426252	1.97	3.2	23.914-25.072	1.65-1.97
		15	30.196	15_12175417	1.67	2.7	30.196	1.67
		17	11.389	17_3117901	3.33	5.9	11.389	3.33

P2	5	17.114	5_4902985	4.16	8.5	16.405-17.114	3.79-4.16
	7	50.842	7_469791	4.38	8	50.842	4.38
	7	61.49	Un_25358685	5.1	9.5	60.781-61.49	4.03-5.1
	11	21.69	11_3453730	5.49	11.8	18.791-21.69	4.74-5.49
	11	33.364	11_5406647	4.97	10.7	32.654-33.364	4.97
	17	20.036	17_5098307	3.65	6.9	18.877-20.036	3.38-3.65
	17	31.414	17_7502059	3.22	5.6	30.705-31.414	3.22
	17	52.126	17_11348276	4.4	8.2	52.126	4.4
	17	57.821	17_16780652	7.56	16	57.111-57.821	7.56-5.18
	18	26.549	18_6354010	1.99	4	26.549	1.99
	18	80.27	Un_16231057	2.21	3.8	80.27	2.21
	18	96.457	18_29102170	2.9	5.3	95.028-96.547	2.86-2.9

Table 40. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Budbreak-Veraison Interval.**

Budbreak-Veraison Interval

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	1	17.382	1_3120443	1,75	4.9	17.382	1.75
		3	54.932	3_17820622	2,02	5.5	54.932	2.02
		3	57.779	3_18408810	3,69	10.5	57.779-58.462	3.69-3.61
		4	0	4_3181677	1,7	4.5	0	1.7
		6	39.753	6_8334007	2,19	6	39.044-39.753	1.74-2.19
		17	5.705	Vvi_6987	4,13	13.9	5.705	4.13
	P2	16	10.711	16_8004504	3,23	11.7	10.711	3.23
		16	17.876	16_13062724	2,88	10.4	14.977-17.876	2.88
		17	0	Un_21136778	3,37	12.3	0	3.37
		17	8.626	17_1572984	4,1	15.2	5.975-8.626	3.63-4.1
Gozzo 2021	P1	3	18.532	3_5762693	4,53	10.2	18.532-20.532	4.53-4.38
		3	22.85	lb_3_5078416	5,21	11.3	22.85	5.21
		6	6.383	6_2314404	5,76	12.8	5.674-6.383	5.76
		6	39.044	6_8154773	2,92	5.9	36.886-39.044	2.68-2.92
		12	13.631	12_2885465	4,87	10.7	13.631-15.631	4.87-4.48
		13	27.814	13_7595444	3,08	6.2	27.105-27.814	3.08
		13	32.08	13_9206375	1,7	3.4	32.08	1.7
		17	58.55	17_16347597	4,83	11	57.391-58.55	4.47-4.83
		18	12.15	18_2618673	5,99	13.3	9.991-12.15-13.15	5.99-5.26
		18	24.341	18_4702061	4,47	9.5	22.912-24.341	4.47

		5	11.41	5_3540064	2,52	8.2	11.41	2.52
		11	45.546	be_1_18883369	3,69	12.4	45.546	3.69
	P2	11	51.251	11_16932395	2,1	6.8	51.251	2.1
		18	50.307	18_11985539	4,47	15.3	49.598-50.307	3.81-4.47
		18	60.966	18_19574196	4,49	15.4	60.966-61.966	4.49-3.92
		2	38.922	2_8538748	8,47	14.4	38.922	8.47
		4	36.326	4_18735077	5,14	7.5	35.617-36.326	4.82-5.14
		4	52.804	4_20532455	4,46	6.1	52.095-52.804	3.99-4.46
		5	34.28	5_10313048	3,42	4	34.28	3.42
		6	39.044	lb_6_8047951	2,69	4.2	39.044	2.69
	P1	6	42.621	6_9091912	6,1	10.4	42.621	6.1
		6	63.188	6_18029643	1,86	3.5	63.188	1.86
		12	30.015	12_6485036	8,66	14.8	28.587-30.015-32.015	7.96-8.66-8.08
		12	55.702	12_20528808	5,3	8.5	55.702	5.3
Guerrieri-Rizzardi 2021		13	22.11	13_6610364	6,27	10	21.4-22.11	6.27
		14	39.652	14_18139541	7,88	13.2	39.652	7.88
		2	39.578	2_16280570	4,78	6.6	38.869-39.578	4.78
		2	47.995	2_17308518	5,59	8	47.578-47.995	5.49-5.59
		4	48.588	4_17007118	9,21	10.3	48.588-49.588	9.21-8.16
		4	50.016	4_17357790	5,39	6	50.016	5.39
	P2	4	68.193	cn_4_21956835	3,26	4.1	68.193	3.26
		8	0	8_130628	2,81	4.3	0	2.81
		8	5.684	8_1699376	3,14	4.7	5.684	3.14
		10	35.363	10_1358731	2,58	3.1	35.363	2.58
		10	67.945	10_8013843	6,05	8.5	67.945	6.05

12	45.181	12_14294532	5,92	8.1	43.022-45.181	5.6-5.92
12	63.217	12_21809346	3,3	4.5	61.058-63.217	3.04-3.3
15	0	15_150741	6,68	6.7	0	6.68
15	5.07	15_5494292	6,31	6.5	4.36-5.07	6.31
16	20.004	16_14544942	4,14	5.5	20.004	4.14
16	35.7	16_19488863	7,17	10.4	35.7	7.17
16	54.201	16_21961919	3,67	3.8	54.201	3.67
18	38.125	18_8812753	4,81	6.2	38.125	4.81

Table 41. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Flowering-Veraison Interval.**

Flowering-Veraison Interval

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	3	57.779	3_18408810	1.65	6.9	57.779-58.462	1.65-1.64
		14	18.668	14_6505467	1.45	6.2	16.509-18.668	1.14-1.45
		18	67.258	18_19509312	1.76	9.1	67.258	1.76
		18	72.952	18_23143805	1.97	10	72.243-72.952	1.97
	P2	5	11.41	mu_5_3513954	2.15	7.8	10.7-11.41	1.91-2.15
		11	18.791	11_3016530	1.81	6.3	18.139-18.791	1.77-1.81
		16	11.42	16_10063467	2.43	8.7	10.711-11.42	2.43
		16	17.876	16_13062724	2.09	7.4	13.839-17.876	1.71-2.09
		17	36.42	17_8738687	1.84	6.5	35.261-36.42-37.42	1.84-1.75
Gozzo 2021	P1	3	17.823	3_5887310	6.7	14.9	17.114-17.823	3.41-6.7
		3	22.85	lb_3_5078416	6.54	14	22.141-22.85	4.16-6.54
		3	52.794	3_17217402	4.44	8.9	49.38-52.794-53.794	3.64-4.44-3.99
		3	55.642	3_17935812	3.89	7.8	55.642	3.89
		6	18.958	6_3589490	2.97	6.4	18.958-19.958	2.97-2.58
		6	39.044	lb_6_8047951	4.08	9	39.044	4.08
		7	0	7_601012	2.66	4.8	0	2.66
		7	23.633	Vvi_3400	4.98	10	23.633	4.98
		14	1.418	14_1391555	3.95	7.9	1.418-2.418	3.95-3.88
		14	10.711	14_4397003	3.5	6.7	10.001-10.711	2.15-3.5

	14	57.677	14_23323355	3.3	7.6	57.677-60.677	3.3-2.82
	14	64.949	14_23631907	4.53	10.6	64.949	4.53
	17	18.543	17_4064095	3.18	7.1	17.384-18.543	2.35-3.18
	17	24.247	17_4794234	3.1	6.7	23.818-24.247	2.79-3.1
	17	58.55	17_16347597	3.2	6.6	56.391-58.55	2.56-3.2
P2	4	0	4_359553	3.7	6.6	0-2	3.7-3.27
	5	11.41	mu_5_3513954	3.04	5.2	10.7-11.41	2.21-3.04
	5	17.114	5_4902985	1.87	3.5	17.114	1.87
	7	67.936	Un_13943695	2.25	2.8	67.036-67.936	1.92-2.25
	7	79.451	lb_7_15393500	5.59	9.4	77.552-79.451	4.28-5.59
	7	97.305	7_19894211	2.97	5	95.876-97.305	2.59-2.97
	9	0	9_137176	4.2	6.7	0	4.2
	9	5.684	9_2912792	5.59	9.1	4.975-5.684	4.24-5.59
	10	15.04	Un_26075992	5.29	9.7	12.389-15.04	3.04-5.29
	10	25.288	Un_26456876	2.24	4.7	21.098-25.288	1.82-4.7
	10	38.2	10_2730976	5.57	9.8	37.491-38.2	5.57
	10	44.594	10_5115610	4.07	7.2	43.884-44.594	3.17-4.07
	10	74.443	10_10642694	4.23	7.3	70.792-74.443	3.01-4.23
	11	45.546	be_1_18883369	3.56	5.9	45.546	3.56
	11	51.251	11_16932395	4.2	6.3	50.542-51.251	2.63-4.2
	12	16.479	12_3847351	2.2	3.5	15.77-16.479	2.2
	18	50.307	18_11985539	2.79	5	49.598-50.307	2.39-2.79
	18	75.243	18_26702868	4.08	6.9	73.085-75.243	3.52-4.08
P1	5	5.026	5_565115	3.21	11.4	4.317-5.026	2.79-3.21
	8	66.109	Un_17625284	2.89	10.2	66.109	2.89

		13	30.651	13_9013857	2.06	7.1	30.651-31.651	2.06-2.04
Guerrieri-Rizzardi 2021	P2	4	0	4_359553	5.51	11	0-2	5.51-4.79
		5	11.41	mu_5_3513954	5.2	10.4	10.7-11.41	4.12-5.2
		5	17.114	5_4902985	4.98	9.9	17.114	4.98
		7	67.936	Un_13943695	2.53	4.7	66.036-67.936	1.95-2.53
		7	79.451	lb_7_15393500	4.59	8.9	77.552-79.451	3.3-4.59
		9	0	9_137176	4.89	9.7	0	4.89
		9	5.684	9_2912792	6.24	12.8	4.975-5.684	4.95-6.24
		10	15.04	Un_26075992	2.87	5.4	13.389-15.04	1.91-2.87
		10	38.2	10_2730976	3.57	6.9	37.491-38.2	3.57
		10	44.594	10_5115610	3.38	6.5	43.884-44.594	2.81-3.38
		10	74.443	10_10642694	3.16	6.1	70.792-74.443	2.48-3.16
		12	16.479	12_3847351	2.99	5.6	15.77-16.479	2.99
		12	29.402	12_7526034	3.39	6.4	28.693-29.402	1.88-3.39
		12	34.407	12_9610810	2.93	5.5	33.979-34.407	2.65-2.93
		15	35.796	15_15813590	1.67	3.1	35.796	1.67
		15	47.389	15_17087784	1.96	3.6	47.389	1.96

Budbreak-Flowering Interval

Table 42. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Budbreak-Flowering Interval.**

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	5	2.159	5_1177392	2.5	6.3	0-2.159	2.31-2.5
		9	11.762	9_2912792	1.84	5.4	8.985-11.762	1.71-1.84
		10	51.375	10_8993108	2.42	5.9	51.217-51.375	2.32-2.42
		10	56.350	10_12144397	3.12	8.2	56.350-57.350	3.12-2.84
		10	64.182	10_16530827	3.7	10.1	63.754-64.182-65.182	3.39-3.7-3.6
		18	18.636	18_3312313	5.86	16	18.636	5.86
		18	19.345	18_3994437	5.41	14.6	19.345	5.41
		18	80.878	18_3397299	2.21	5.3	80.878-82.878	2.21
	P2	4	39.296	4_11091087	5.28	13	37.867-39.296-41.296	5.14-5.28-4.79
		4	47.159	4_16310583	7.08	18.2	46.450-47.159-48.159	7.08-6.9
		4	68.902	4_22019059	3.34	7.9	68.193-68.902	3.14-3.34
		7	42.916	5_20326151	3.54	8.4	41.487-42.916	3.23-3.54
		14	42.133	14_24930332	4.51	10.9	42.133	4.51
		16	51.688	16_21396785	3.34	7.9	51.688	3.34
		17	14.310	17_4315095	2.67	6.2	13.601-14.310	2.07-2.67
		17	40.716	17_9363399	4.45	10.7	40.716	4.45
Gozzo 2021	P1	2	36.794	2_7969863	2.26	6.1	36.085-36.794	1.82-2.26
		12	10.053	12_1884573	3.22	9	10.053	3.22
		12	60.772	12_22537667	2.56	6.5	57.121-60.772	2.18-2.56

		17	52.825	17_14654402	5.19	14.5	51.665-52.825	4.63-5.19
		17	54.963	17_15491016	5.76	16.4	54.963-55.963	5.76-5.43
		18	67.258	18_19509312	1.91	5.6	66.549-67.258	1.86-1.91
		18	72.952	18_23143805	1.93	5.6	72.243-72.952	1.93
	P2	4	74.617	4_23731066	2.64	5.4	73.908-74.617	2.15-2.64
		7	4.985	7_2205196	5.42	11.7	4.985	5.42
		8	0	8_130628	2.86	6.1	0	2.86
		9	9.293	9_3446582	4.7	9.9	9.293	4.7
		9	15.901	9_4412220	4.54	9.7	15.901	4.54
		10	10.680	10_328404	4.7	10.1	9.970-10.680-11.389	3.62-4.7-3.28
		10	15.040	Un_26075992	7.08	16.1	13.389-15.040-17.040	6.72-7.08-6.75
		16	5.027	16_2545201	3.16	6.4	4.318-5.027	2.48-3.16
		17	20.036	17_5098307	3.49	7.3	17.877-20.036	2.84-3.49
		18	44.613	18_10312345	4.84	10.4	41.961-44.613	3.81-4.84
		18	60.966	18_19574196	2.44	4.9	60.966	2.44
Guerrieri-Rizzardi 2021	P1	1	54.553	1_10913571	2.23	5.4	53.125-54.553	1.73-2.23
		2	40.340	2_8828693	3.04	7.5	40.340	3.04
		6	3.546	6_1774386	2.79	6.8	3.546	2.79
		13	17.104	13_5512630	2.95	7.3	16.395-17.104	2.63-2.95
		14	41.081	14_18999008	5.37	16.9	41.081	5.37
		14	45.409	14_20465458	5.09	16.9	45.409	5.09
		18	62.231	18_18685180	4.33	11	61.522-62.231	4-4.33
	P2	2	6.048	2_1424334	2.34	7.6	5.339-6.048	2.09-2.34
		3	30.389	3_4196090	1.99	6.4	30.389	1.99

3	38.221	3_8283428	2.35	7.6	38.221-39.221	2.35-2.08
10	0	10_208152	3.22	10.6	0	3.22
10	7.133	Un_42945165	1.89	6	6.975-7.133	1.85-1.89
18	33.710	18_8245578	1.91	6.1	33.710	1.91

Derivative maturity

Table 43. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Derivative maturity.**

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	1	60.237	lb_1_14453117	2.96	5	59.809-60.237-61.237	2.8-2.96-2.95
		1	66.651	1_16864011	7.69	14.7	66.651	7.69
		5	24.289	5_6372886	4.15	7.2	22.86-24.289	3.68-4.15
		5	29.316	5_6925484	3.48	5.9	28.606-29.316	3.39-3.48
		8	42.622	8_14172672	3.79	6.5	40.464-42.622	2.89-3.79
		13	11.399	13_4472603	5.48	9.8	10.24-11.399	5.28-5.48
		13	21.4	13_6233266	5.37	9.6	20.972-21.4	5.07-5.37
		15	40.26	15_13737462	3.56	6.1	39.832-40.26-41.26	3.25-3.56-3.05
		15	43.847	15_14386692	3.52	6	42.689-43.847	2.96-3.52
		18	18.636	18_3749154	6.88	12.8	18.636	6.88
		18	22.203	SNP1519_47	6.5	12	22.203	6.5
		18	64.39	18_19427319	5.64	10.2	64.39	5.64
		18	72.243	18_22933646	2.96	5	72.243	2.96
	P2	1	29.71	1_9242867	4.94	6.6	29.71	4.94
		1	41.453	1_17139876	7.69	11.1	40.035-41.453	6.23-7.69
		2	39.578	2_16652759	3.88	5.1	39.578	3.88
		2	53.71	2_18360408	14.01	24.3	52.551-53.71	13.68-14.01
		5	4.996	5_272058	8.12	11.9	4.996	8.12
		5	7.832	5_3144283	10.02	15.5	7.832	10.02
		7	3.567	7_1996363	2.85	3.6	3.567	2.85

		7	42.916	5_20326151	7.42	10.7	41.487-42.916	6.29-7.42
		10	82.254	10_17403870	2.82	3.6	81.545-82.254	2.52-2.82
		14	12.901	14_4861406	3.49	4.5	12.191-12.901	3.39-3.49
		14	17.176	14_7311742	2.6	3.3	17.176	2.6
		15	59.56	15_19075876	4.11	5.4	59.56	4.11
		17	8.626	17_1572984	6.15	8.5	4.975-8.626	5.31-6.15
		19	3.577	19_1722935	5.88	8.1	2.868-3.577	5.47-5.88
Gozzo 2021	P1	3	27.815	3_7056491	2.57	6.5	27.105-27.815	2.37-2.57
		11	5.705	11_1370629	7.61	20.9	4.996-5.705	6.14-7.61
		11	9.971	11_1811149	8.27	23.1	9.971-10.971	8.27
		14	26.51-26.51	14_8895523	1.74	4.6	25.082	1.71-1.74
		14	45.409	14_20465458	4.94	15.1	44.98-45.409-46.409	4.76-4.94-4.4
		14	56.967	14_22485384	5.28	15.9	56.967	5.28
		15	15.78	15_10077073	3.12	7.4	14.62-15.78-16.78	2.53-3.12-2.99
		15	25.072	15_11426499	3.93	9.4	25.072	3.93
		18	22.203	18_4381193	2.06	4.8	20.774-22.203	1.95-2.06
		18	35.699	18_9551907	2.32	5.4	35.699	2.32
	P2	4	29.993	4_7662501	6.6	9.9	29.284-29.993	6.6
		7	21.537	7_5855280	5.52	8.1	19.638-21.537	5.13-5.52
		7	23.675	mu_7_6484762	4.88	7	23.675	4.88
		7	97.305	7_19894211	5.18	7.5	95.876-97.305	4.11-5.18
		8	27.846	8_11350976	4.98	7.2	27.846	4.98
		8	29.264	8_11679690	6.35	9.5	31.264-29.264	5.29-6.35
		14	1.418	14_855835	2.79	3.8	0.709-1.418	2.69-2.79
		14	6.476	14_1326554	4.47	6.4	5.318-6.476	4.14-4.47

		17	8.626	17_1632106	12.54	21.8	8.626	12.54
		17	54.973	17_108934	3.73	5.2	54.973	3.73
		18	55.991	18_14227977	4.75	3.2	55.991	4.75
		18	60.966	18_19574196	3.36	3.3	61.966-60.966	3.253.36
		18	75.243	18_26843102	5.32	7.7	75.243	5.32
Guerrieri-Rizzardi 2021	P1	6	21.117	6_4099263	4.83	6.6	21.117	4.83
		6	24.684	6_5012264	3.36	4.6	24.255-24.684	3.063.36
		7	57.203	Un_17790089	4.71	7.2	56.494-57.203	4.23-4.71
		8	66.109	Un_17625284	3.2	5.5	65.4-66.109	3.2
		9	0	9_159487	4.87	11	0	4.87
		11	0	11_391566	7.17	13	0	7.17
		11	9.971	11_1811149	7.69	15.1	9.971-10.971	7-69-7.68
		11	21.4	11_3521933	2.01	3.6	21.4	2.01
		15	34.514	15_12545262	1.68	2.8	34.514	1.68
		15	40.26	15_13737462	2.68	4.7	39.832-40.26-41.26	2.56-2.68-2.24
		15	73.859	15_17464241	6.52	7.9	73.859	6.52
		15	88.804	15_20153730	3.99	3.5	87.376-88.804	3.08-3.99
		16	7.917	16_1259589	3.45	5.8	7.917	3.45
	P2	4	24.998	4_5502978	2.19	4.6	24.998	2.19
		4	27.866	4_6495602	3.01	6.3	27.866	3.01
		8	43.875	8_14384882	4.44	7.8	40.685-43.875	3.77-4.44
		10	25.288	Un_26456876	12.51	27	20.098-25.288	12.2-12.51
		10	31.786	Un_19275293	11.65	24.6	31.077-31.786	11.65
		13	35.24	Un_17274294	2.91	4.8	34.531-35.24	2.91
		17	36.42	17_8738687	9.59	18.9	34.261-36.42	9.81-9.59

	18	73.085	be_18_26115662	5.35	9.5	73.085	5.35	

Table 44. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Berry color.**

Berry color

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	2	43.907	2_15113892	7.62	25.4	43.197-43.907	6.48-7.62
		3	44.019	3_233171	6.24	19.7	39.429-44.019	5.42-6.24
		4	32.780	4_17856784	2.16	5.9	32.070-32.780	2.16
		7	0.000	7_601012	3.74	10.8		
		7	2.837	Un_7297439	3.98	11.6	2.837-3.837	3.98-9.97
		7	45.784	7_652010	5.03	15.2	44.625-45.784	4.54-5.03
		7	56.494	Un_17857429	4.59	13.6	55.065-56.494	4.58
		7	0.000	7_601012	3.74	10.8		
		13	27.814	13_7595444	2.24	6.2	27.105	2.24
		18	5.016	18_582557	3.06	8.6	3.587-5.016	3.06
	P2	2	39.578	2_16280570	8.36	16.8	38.869-39.578	8.36
		2	47.995	2_17308518	5.01	8.9	43.578-47.995	4.23
		2	51.551	2_18023815	4.60	8.1	51.551	4.60
		6	32.671	6_13398369	2.63	4.3	32.671-33.671	2.63-2.60
		15	10.764	15_10338742	4.19	7.3	10.055-10.764	3.86-4.19
		16	10.002	cn_16_6630270	1.87	3.0	10.002	1.87
		17	31.414	17_7502059	5.32	9.6	30.705-31.414	5.32
		17	36.420	17_8738687	5.04	9.0	35.261-36.420-37.420	4.37-5.04
		17	52.126	17_11348276	2.98	5.0	51.416-52.126	2.38
Gozzo 2021	P1	2	46.065	2_17340911	7.32	23.6	43.907-46.065	7.61-7.32
		5	0.000	5_1580613	3.69	10.4		
		5	17.176	5_5079107	4.56	14.4	16.467-17.176	3.04-.56
		5	22.151	5_6166103	2.15	6.6	22.151	2.15

		11	28.524	11_5122694	1.84	5.0	27.815-28.524	1.84
		11	0.000	11_391566	1.88	5.3	0.000	1.88
		12	52.855	12_19007153	3.11	8.7	50.696-52.855	2.63-3.11
		12	55.702	12_20528808	3.27	9.3	54.993-55.702	3.16
	P2	2	39.578	2_16280570	14.35	40.7	38.869-39.578	14.35
		6	6.419	6_1948723	3.90	8.2	4.260-6.419	3.74
		11	14.430	11_2568420	1.85	3.6	14.430	1.85
		14	25.770	Vvi_1222	2.10	4.2	25.770	2.10
		17	27.149	17_6320545	2.30	4.5	26.720-27.149	2.16
Guerrieri-Rizzardi 2021	P1	2	46.065	2_17340911	7.40	22.0	43.907-46.065	7.01-7.40
		3	12.139	3_2582243	2.97	7.9	9.240-12.139	2.1-2.97
		3	33.530	3_8827121	3.33	9.1	32.820-33.530	3.33
		5	37.931	5_11498692	1.95	4.9	37.931	1.95
		18	14.308	18_2853425	4.36	13.0	12.150-14.308-15.308	3.96-4.36-3.89
		18	67.258	18_19509312	3.46	9.7	66.549-67.258	2.39-3.46
		18	77.978	18_24704392	2.85	7.9	77.978	2.85
	P2	2	39.578	2_16280570	12.97	28.8	38.869-39.578	12.97
		6	0.000	6_164244	5.48	10.2	0.000	5.48
		7	56.526	Un_15024930	4.44	8.1	55.816-56.526	4.22
		7	67.936	Un_13943695	6.47	12.3	67.036-67.936	6.03
		7	90.891	7_17802712	2.21	3.6	88.733-90.891	1.89-2.21
		17	27.149	17_6320545	2.81	4.2	25.720-27.149	2.36-2.81
		17	36.420	17_8738687	2.52	3.9	35.261-36.420	2.02-2.52
		17	54.264	be_17_11981751	2.88	4.9	52.835-54.264	1.99-2.88

Table 45. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Polyphenol content.**

Polyphenol content

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	3	28.524	3_7472539	6.46	11.9	28.524	6.46
		3	33.530	3_8827121	10.06	22.1	32.820-33.530	10.06
		9	11.762	9_2912792	5.28	10.7	10.985-11.762-14.762	5.27-5.28-4.27
		11	0.000	11_391566	6.59	12.5	0.000	6.59
		11	43.625	11_9131853	8.50	17.5	42.196-43.625-44.65	8.50-7.98
		11	52.896	11_18318407	6.26	12.0	52.187-52.896	6.26-5.74
		13	32.080	13_9206375	3.34	2.8	31.651-32.080-34.080	3.24-3.34
		15	30.196	15_12175417	4.31	7.2	29.487-30.196	4.31
		15	46.695	15_14635285	4.47	8.1	46.266-46.695	4.31
		16	0.000	16_141454	2.51	5.0	0.000	2.51
	P2	6	38.396	6_15434265	1.90	8.4	37.967-38.396	1.85
		7	11.430	7_3841674	4.32	22.4	10.272-11.430	3.99-4.32
		7	21.537	7_5855280	2.54	11.7	20.638-21.537	2.38-2.54
		8	33.679	8_11896722	2.51	11.5	33.679-35.679	2.51
		8	43.875	8_14384882	2.70	12.5	40.685-43.875	2.01
		10	15.040	Un_26075992	2.21	9.9	11.389-15.040-17.040	2.21
		12	32.979	12_9212393	2.50	11.5	30.820-32.979-33.979	2.5
		12	63.926	SNP1119_176	3.46	16.9	63.926	3.46
Gozzo 2021	P1	8	7.154	8_2111324	6.34	7.1	4.995-7.154	5.8-6.34
		8	54.021	8_18032958	23.65	60.7	53.312-54.021	23.65
		8	66.109	Un_17625284	18.53	33.5	65.400-66.109	18.53
		11	0	11_391566	6.09	6.8	0	6.09
		12	0	12_473337	4.76	6.2	0	4.76

		12	10.763	12_2020210	6.31	7.7	10.763-12.763	6.31-5.47
		14	0.709	14_691403	6.01	5.5	0.000-0.709	6.01
		14	6.424	14_2063666	8.66	8.8	5.266-6.424	8.32-8.66
		14	67.786	14_25093632	4.16	4.9	67.076-67.786	4.16
		14	73.470	14_25700297	5.16	6.5	73.041-73.470	5.01-5.16
		17	47.796	17_11764553	2.28	3.3	47.087-47.796	2.04-2.28
		17	58.550	17_16347597	3.13	4.1	57.391-58.550	2.31
		18	51.521	18_13410273	4.99	5.8	51.521	4.99
		18	56.527	18_18119009	4.85	5.6	56.098-56.527	4.53
		19	0	19_17827	3.6	4.4	0	3.6
		19	18.574	19_4957496	4.88	6.5	15.675-18.574	4.46-4.88
		19	57.575	19_22424512	2.03	2.5	57.575-57.575	2.02-2.03
	P2	2	27.449	2_7970900	3.09	6.6	26.021-27.449	2.88-3.09
		6	36.967	6_15154592	6.12	15	36.967-37.967	6.12-5.6
		6	45.519	6_17411193	5.57	13.4	45.090-45.519	5.42-5.57
		11	14.430	11_2568420	4.21	9.5	9.318-14.430	3.38
		11	39.767	11_7673741	2.17	4.4	38.339-39.767	2.04
		17	17.877	17_4853813	2.09	6.8	17.877	2.09
		17	20.036	17_5098307	4.5	10.3	20.036	4.5
		17	28.567	17_6636733	2.82	5.9	28.567	2.82
		18	40.961	18_9912038	3.42	8.5	40.252-40.961	3.24-3.42
		18	50.307	18_11985539	5.63	14.8	49.598	4.38-5.63
		18	68.819	cn_18_23222456	6.04	15.2	66.660-68.819	5.03-6.04
		18	88.476	18_3406716	5.64	12.3	85.698-88.476	4.71-5.64
		18	94.319	18_29175197	3.11	6.4	90.904-94.319	2.83
Guerrieri-Rizzardi 2021	P1	5	43.688	5_14739326	6.23	17	43.688	6.23
		5	58.131	5_19488276	4.81	12.3	57.421-58.131	4.48-4.81
		6	24.684	6_5012264	3.45	8.4	23.255-24.684	2.54-3.45
		15	4.296	cn_15_3239866	2.61	6.1	2.867-4.296-6.296	2.61-2.42
		16	12.172	16_5315939	3.5	8.5	12.172	3.5
		16	32.884	16_16811470	3.08	7.3	32.455-32.884	2.79

	16	39.372	16_18325657	2.65	6.2	38.663-39.372	2.65
	19	5.694	19_1490246	3.59	8.8	5.694	3.59
	19	7.112	19_1771829	6.28	17.2	7.112	6.28
P2	1	0	1_164167	13.18	11.5	0.000-1.418	13.18-10.53
	2	47.995	2_17308518	4.38	2.6	42.578-47.995	3.41-4.38
	4	64.626	Un_30937816	6.11	3.9	64.626-65.626	6.11
	9	0.709	9_716568	16.87	17.9	0.000-0.706	15.21-16.87
	9	38.467	9_13407497	2.93	1.6	38.467	2.93
	10	15.040	Un_26075992	10.75	1.9	12.389-15.040-17.040	9.62-10.75
	10	31.786	Un_19275293	10.46	1.7	31.077-31.786	10.46
	10	65.076	10_7426153	9.49	7.1	63.646-65.076	9.49
	10	80.836	10_16418120	4.42	2.6	79.407-80.836	4.42
	12	35.826	ae_12_10296309	2.68	1.5	35.117-35.826-37.826	2.35-2.68
	13	19.171	13_5643893	12.59	10.7	19.171-21.171	12.59-10.32
	13	23.488	13_6319782	10.72	8.4	22.329-23.488	10.06
	13	38.848	13_8069517	7.05	1.7	36.949-38.848	6.58-7.05
	13	44.553	13_13485153	8.38	2.6	43.844-44.553	8.38
	16	2.899	16_1259589	5.92	3.8	0.000-2.899	5.62-5.92
	18	62.395	18_19607453	11.6	9.6	60.966-62.395	11.6
	18	80.270	Un_16231057	7.48	5.1	78.111-80.270	6.53-7.48
	18	88.476	18_3406716	3.18	1.8	88.476	3.18

Table 46. Complete list of QTL regions targeted from **Multiple QTL Mapping (MQM)** on the trait **Anthocyanin content.**

Dataset	Map	LG	Peak Position	Marker	LOD max	% var	Confidence interval	
							Position	LOD
Gozzo 2020	P1	8	38.326	8_13093399	7.75	44.9	38.326-37.616	7.75
		8	49.756	8_16155311	5.77	28.8	49.756	5.77
		10	32.026	10_4083113	6.69	37.9	30.867-32.026	6.35-6.69
		10	42.718	10_6919630	7.49	43.7	42.718	7.49
		12	0.000	12_473337	5.52	29.5	0.000	5.52
		12	10.053	12_1715344	6.44	35.7	10.053	6.44
		12	54.993	12_20213782	3.60	8.6	54.993	3.60
		18	15.737	18_2987719	3.34	7.8	14.308-15.737	3.14-3.34
	P2	2	29.587	2_8757865	1.85	8.1	29.587	1.85
		5	8.542	5_3425645	4.31	21.9	7.832-8.542-9.542	3.87-4.31-4.16
		7	11.430	7_3841674	4.30	21.4	11.430	4.30
		16	19.295	16_14060841	2.26	10.0	19.295	10.0
		16	29.995	16_17703435	3.53	16.8	29.995	16.8
		16	41.445	16_20799346	3.82	18.7	40.006-41.445	3.52-3.82
		16	51.688	16_21396785	2.58	11.4	51.688	2.58
Gozzo 2021	P1	1	84.164	1_22498675	15.72	27.9	83.198-84.164	15.09-15.72
		6	18.958	6_3589490	13.57	21.8	15.092-18.958-19.958	12.72-13.57-13.30
		6	24.684	6_5012264	13.18	20.5	24.255-24.684	13.03-13.18
		8	37.616	8_12974771	7.51	0.8	35.639-37.616	7.19-7.51
		8	47.618	8_15889641	5.83	0.1	47.618-49.327	5.83-5.13
		14	0.000	14_329145	4.26	1.8	0.000	4.26

		15	20.776	15_10885751	11.94	15.3	20.776	11.94
		15	84.539	15_19318357	7.54	3.0	83.829-84.539	6.5-7.54
		16	0.000	16_141454	7.24	2.7	0.000	7.24
		16	4.266	16_1182689	3.65	1.5	4.266	3.65
		16	13.591	li_14_4403150	5.81	2.7	13.591	5.81
		16	20.015	16_12177014	13.61	14.7	20.015	13.61
		16	26.460	16_14884281	2.09	0.8	25.301-26.460	1.81-2.09
		16	37.954	16_17780135	3.09	1.3	37.954	3.09
		17	39.213	17_9404993	7.00	3.4	39.213	7.00
		17	48.505	17_364206	3.27	1.3	48.505	3.27
	P2	2	33.884	2_13059712	3.64	20.6	31.725-33.884	3.64
		3	34.654	3_7223431	2.34	12.4	33.226-34.654	3.34
		3	41.787	3_9524678	1.80	9.3	41.787	1.80
Guerrieri-Rizzardi 20021	P1	16	18.586	16_11988588	5.90	29.5	18.586	5.90
		16	24.301	16_14533665	4.31	20.1	24.301	4.31
		18	36.408	18_10464360	2.16	9.2	36.408	2.16
		19	25.351	19_7326716	4.72	22.4	20.574-25.351-27.351	4.19-4.72-4.39
	P2	1	56.254	1_22498675	4.56	21.3	54.323-56.254	4.51-4.56
		11	38.339	11_65847	2.23	8.9	38.339	2.23
		16	26.428	16_16229581	6.01	27.9	26.428	6.01
		16	30.704	16_17780135	4.71	20.7	30.704	4.71
		18	47.460	18_11080029	2.95	12.3	46.751-47.460	2.64-2.95
		18	58.828	18_18851921	2.35	9.7	58.119-58.828-59.828	

Gene ID	*Functional annotation*
VIT_12s0035g00290	Sugar transporter ERD6-like 6
VIT_12s0035g00300	Carbohydrate transmembrane transporter SFP1
VIT_12s0035g00310	Protein kinase SPK-3 ASK1 (SnRK-6)
VIT_16s0039g00610	Serine/threonine-protein kinase ppk15
VIT_16s0039g00660	RNA polymerase nonessential primary-like sigma factor SIGA
VIT_16s0050g01880	Receptor kinase homolog LRK10
VIT_18s0072g00140	CRK1 protein
VIT_18s0072g00170	BRI1 (brassinosteroid insensitive 1)
VIT_18s0001g03430	Flavonol synthase
VIT_18s0001g03470	Flavonol synthase
VIT_18s0001g03490	Flavonol synthase
VIT_18s0001g03510	Flavonol synthase XM_002284374.1
VIT_18s0001g03540	Auxin transporter protein 4
VIT_18s0001g03610	Auxin-independent growth promoter
VIT_18s0001g03670	Zinc finger (C2H2 type) family

Table 47 – Gene ID and functional annotation of genes surrounding flanking region of the three SNPs validated with Sanger sequencing.

- *Supplementary table* 1: available upon request.

Acknowledgments

I would especially like to thank our partners at CREA-VE in Conegliano, Dr. Manna Crespan and Dr. Daniele Migliaro, for the materials, hospitality and advice.

Acknowledgments goes also to our collaborators of the Edmund Mach Foundation, Prof. Maria Stella Grando and Dr. Laura Costantini.

I would like to thank Prof. Diana Bellin for having followed and advised me in this path and research project, once again.

Thanks to all the members of my lab, Giada, Martina, Chiara, Bita and Pietro, who in recent years have made me smile and helped in everything.
I also thank all the guys who have contributed to this project, helping us in the field and in laboratory.

Thanks to all my friends, who have never stopped making these years fantastic.