TOWARDS THE DEFINITION OF A DETAILED TRANSCRIPTOMIC MAP OF GRAPE BERRY DEVELOPMENT

Authors:Marianna FASOLI[1], Chandra L. RICHTER[1], Sara ZENONI[2], Marco SANDRI[2], Paola ZUCCOLOTTO[3], Mario PEZZOTTI[2], Nick DOKOOZLIAN[1], and Giovanni Battista TORNIELLI[2,*]

[1]*E&J Gallo Winery, Modesto, CA 95353, USA*
[2]*Department of Biotechnology, University of Verona, 37134 Verona, Italy*
[3]*Big & Open Data Innovation Laboratory, University of Brescia, 25123 Brescia, Italy*

*Corresponding author: *giovannibattista.tornielli@univr.it*

Abstract:

Context and purpose of the study - In the last years the application of genomic tools to the analysis of gene expression during grape berry development generated a huge amount of transcriptomic data from different varieties and growing conditions. This information set the stage to understand the molecular basis of crucial developmental and metabolic rearrangements occurring during grape berry formation and ripening. It is now clear that the variation of a portion of berry transcriptome is conserved across cultivars and growing conditions, and thus may be used universally to describe the stage of berry development. In this work we explore the possibility of using the transcriptomic data generated from two cultivars to define a very detailed developmental map of the grape berry.
Material and methods - To map the molecular events associated with berry development at very high temporal resolution, we performed RNA-seq analysis of berry samples collected every week from fruit-set to maturity from Pinot noir and Cabernet Sauvignon vines grown in the same location. The experiment was replicated across three consecutive years (2012, 2013, 2014) resulting in 219 samples overall. Applying multivariate analyses to the most variable portion of the transcriptome, we built a transcriptomic model of berry development based on the molecular information obtained from samples of both cultivars.
Results - The Pinot noir and Cabernet Sauvignon samples mostly aligned in a 3D transcriptomic map (~80% of the variance described by Principal Component Analysis), allowing to define a general model of berry development based on gene expression. The performance of the model in describing the development of other grape varieties was accessed projecting RNA-seq samples of fruit development of ten Italian cultivars onto the model. Both red and white-skin berry samples mapped on the transcriptomic map and revealed alignment by standard ripening parameters (e.g. total soluble solids) as well as unrelated to any of these. Moreover, we validated that berry maturation of the same cultivar cultivated in different International growing regions can be well represented and aligned by means of our transcriptomic map.
These results showed that the transcriptomic information can be accessed to precisely define a model of "molecular phenology" that can be used to map the ontogenetic development of the fruit with high precision and to align the stage of berry development of different grapes.

Keywords: Grapevine, Berry development, Ripening, Molecular Phenology, Transcriptomics.

1. Introduction

The Modified E-L and the Extended BBCH systems (Coombe, 1995; Lorenz et al., 1995) are the phenology scale systems most adopted by viticulturists. These systems describe the annual phenology of the plant, including grape berry development from fruit set to maturity, and number the main developmental stages by increasing order. The use of phenological scales is extremely helpful to describe and compare the timing of grape development and ripening, which can be very different by vintage and/or cultivation site. However, although some stages are well defined and clearly mark developmental events from a physiological point of view (e.g. fruit set, veraison), defining a comparable developmental stage for grapes of the same cultivar when grown in different conditions or for grapes of different cultivars can very likely generate mistakes, in particular after the onset of ripening. The ripening degree is normally described by the sugar content that represents a major technological

parameter for winemaking. However, it is well known that this parameter is highly influenced by several factors like climate, water availability, canopy management practices and crop load. The uncoupling between sugar concentration and other ripening parameters such as berry acidity or skin anthocyanin content have been reported (Sadras and Moran, 2012; Bobeica et al., 2015), evidencing that sugar content may not solely denote the physiological stage of berry ripening progression.
In the last years the application of genomic tools to the analysis of gene expression during grape berry development have generated a huge amount of transcriptomic data from different varieties and growing conditions (Fasoli et al., 2012; Massonnet et al., 2017; Dal Santo et al., 2018). It has been shown that the variations of a portion of the transcriptome (the core transcriptome) along berry development seem to be conserved across cultivars and growing condition of grapevines, and thus may be used to describe the developmental stage of berry development (Dal Santo et al., 2013; Massonnet et al., 2017; Dal Santo et al., 2018).
In this work we explore the possibility of using the transcriptomic data generated from grape berries weekly sampled from Cabernet Sauvignon and Pinot noir vines grown in the same location over three consecutive vintages to map the development of the grape berry. We used the most variable portion of the transcriptome to build a preliminary transcriptomic model of berry development, which allowed to precisely define the progression of berry development during both formation and ripening phases.

2. Material and methods

2.1 - Vineyard features and plant material growth conditions
Vitis vinifera cultivars Cabernet Sauvignon (clone FPS 8 grafted on 5C rootstock and planted in 1997) and Pinot noir (clone FPS 23 grafted on Freedom rootstock and planted in 2001) were used in this study. Both varieties were grown in sandy clay loam in an east–west orientation with 10-foot row spacing and 5-foot vine spacing.

2.2 - Sampling strategy
Sampling was performed as described by Fasoli et al. (2018). Berries were collected at 10-day intervals in 2012, and weekly in 2013 and 2014, beginning at fruit-set and continuing until harvest (24.5 °Brix). Veraison was defined as 50% colored berries per cluster. All samples were collected at the same time of day (8:00 am) in randomized block designs for each cultivar: eight-vine blocks for Pinot noir and six-vine blocks for Cabernet Sauvignon. The block designs were replicated along three rows for each cultivar, to allow the collection of biological triplicates. We therefore collected 219 samples in total: 120 for Cabernet Sauvignon (39, 42 and 39 during vintages 2012, 2013 and 2014, respectively) and 99 for Pinot noir (30, 33 and 36 during vintages 2012, 2013 and 2014, respectively). Each sample replicate comprised 26 clusters of berries from each vine block.

2.3 - RNA extraction, Library preparation and RNA sequencing analysis
Sixty berries, from six isolated clusters randomly selected from the vine blocks, were ground under liquid nitrogen. Seeds were removed before grinding. Frozen powder was divided into 400-mg aliquots for RNA extraction. Total RNA was isolated using the Spectrum Plant Total RNA kit (Sigma-Aldrich, St Louis, MO, USA) following the manufacturer's instructions with some modifications (Fasoli et al., 2012).
We prepared 219 non-directional cDNA libraries from 2.5 µg total RNA using the Illumina TruSeq RNA Sample preparation protocol (Illumina Inc., San Diego, CA, USA). Library quality was determined using the Agilent High Sensitivity DNA kit on the Agilent 2100 Bioanalyzer, and the quantity was determined by quantitative PCR using the KAPA Library Quantification kit (Kapa Biosystems, Roche Diagnostics, Basel, Switzerland). Single-end reads of 100 nucleotides were obtained using an Illumina HiSeq 1000 sequencer and sequencing data were generated using the base-calling software Illumina Casava v1.8 (32,027,722 ± 7,628,415 per sample). The reads were aligned to the grapevine 12x reference genome PN40024 Jaillon et al., 2007 using TopHat v2.0.6 with default parameters Kim et al., 2013. An average of 86.7% of reads was mapped for each sample. Mapped reads were used to reconstruct the transcripts using Cufflinks v2.0.2 (Roberts et al., 2011) and the reference genome annotation V1 (http://genomes.cribi.unipd.it/DATA). The normalized expression of each transcript was calculated as RPKM (reads per kilobase of transcript per million mapped reads) for each sample.

2.4 - Multivariate analysis and model creation

PCA was carried out using R Studio software (https://www.rstudio.com/) and the Spearman statistical metric was chosen to create the correlation matrix.

To focus on the most meaningful genes, filtering steps were applied to the Cabernet Sauvignon and Pinot noir dataset, and the red-skin Italian varieties survey (Massonnet et al., 2017). In addition to eliminating genes with no expression across all samples, we targeted and removed genes showing low expression (RPKM < 1; Mortazavi et al., 2008) at each time point except one in at least one year and one cultivar. A second filter was applied to remove genes showing low variability (almost constant expression) across the experimental conditions. A final polish aimed at removing genes with overall low expression level (RPKM < 1) and random isolated peaks considered expression outliers. This multistep approach allowed concentrating the creation of the transcriptomic model on 9,819 meaningful genes.

3. Results and discussion

***2.1** - Berry development and ripening*

We collected berry samples from fruit set to full maturity from Cabernet Sauvignon and Pinot noir vines grown in the same location over three consecutive vintages. Samples were collected every 7–10 days using a randomized block approach to account for field variability. This resulted in 219 samples. The recording of heat accumulation (growing degree days) showed that the 2013 and 2014 seasons were warmer than 2012 during March to August (data not shown). Veraison therefore occurred 10–20 days earlier in Pinot noir and 4–11 days earlier in Cabernet Sauvignon. Total Soluble Solids (TSS by °Brix) were measured from the sampling time point following 50% veraison to harvest (target °Brix value for commercial harvest was 24.5). Pinot noir berries accumulated sugar more rapidly than Cabernet Sauvignon berries, resulting in a shorter development time and an earlier harvest in all three vintages (Figure 1).

In 2014 the estimation of veraison stage may have been slightly anticipated and veraison samples (time point 3) likely collected at an actual earlier stage for both varieties compared to the previous vintages, as suggested by the lower sugar content of the first post-veraison samples (time point 4). Later maturation stages occasionally showed TSS fluctuation, in particular for Cabernet Sauvignon in 2012, which made harvest prediction particularly challenging. At this developmental phase, sugar content accumulation is considerably slowed down, and the irrigation schedule is the most responsible for the evidenced variation in TSS. Alternation of irrigation and dry days is an approach used to manage the harvest schedule of a vineyard block and adjust the TSS to the desired value. This is a clear example showing that sugar content may not be a suitable parameter when used alone to determine the stage of grape berry development during the ripening period.

***2.2** - Creation of the transcriptomic model*

RNA-Seq was used to monitor the expression of all grapevine genes (http://srs.ebi.ac.uk). Cabernet Sauvignon showed a longer development time and thus the preliminary transcriptomic grape ripening model was developed on this variety. To focus on the most meaningful genes, 9,819 genes were selected over non- or poorly-expressed genes across the Cabernet Sauvignon 120 samples. Principal component analysis (PCA) was used to evaluate changes in gene expression during berry development, using the average expression value of the sample triplicates (Figure 2). The first principal component (Dim 1, 66.8%) separated samples according to development and/or time. The second principal component (Dim 2, 19.3%) added resolution in separating samples from time points 4 to 13. The third component (Dim 3, 5.6 %) allowed to distinguish early time point samples (from 0 to 3). The PCA did not reveal sample separation by vintage.The transcriptomic route of the Cabernet Sauvignon berry during its development could be summarized by the two plots showed in Figure 2. Pinot noir RNA-seq triplicates were averaged and then projected onto the Cabernet Sauvignon transcriptomic maps (Dim 1-Dim 2, Dim 1-Dim 3; Figure 2). The alignment of the Pinot noir samples on the Cabernet Sauvignon transcriptomic route evidenced the suitability of the model in comparing grape developmental data of different varieties (not shown). We therefore performed a PCA analysis using RNA-seq data of both varieties. All samples distributed along an ideal transcriptomic route well described by three components, and could be interpolated by a curve line that was marked by 30 equally spaced stages of transcriptomic development (Figure 3).

2.3 - *Model performance with five red-skin Italian varieties*

To verify the performance of the transcriptomic model for other varieties and different growing regions, we used a RNA-Seq dataset describing berry development transcriptome profiles of five grapevine red-berry varieties (Sangiovese, Barbera, Negroamaro, Refosco, and Primitivo) and five white-berry varieties (Garganega, Glera, Moscato bianco, Passerina, Vermentino) cultivated during 2011 in the same experimental vineyard (Conegliano, Veneto region, Italy; Massonnet et al., 2017). This dataset was generated from berries collected at four phenological stages: pea-size (BBCH 75) at 20 days after flowering (Pea), beginning to touch (BBCH 77) just prior to veraison (Touch), softening (BBCH 85) at the end of veraison (Soft), and ripe (BBCH 89) (Harv). After averaging the biological triplicates, the forty RNA-seq averaged samples were plotted onto the transcriptomic map (Figure 4). The ten-variety samples aligned on the map and clustered by phenological stage rather than cultivar, likewise previously shown by Massonnet et al. (2017). As expected, the Dim 1-Dim 2 plot separated the green-phase and the maturation stages, and Soft and Harvest stages from each other (Figure 4, left plot). The alignment seemed to mirror the grapes maturation degree in terms of TSS accumulated by each variety at the specific stage. Indeed, the berry samples collected from the Italian varieties never reached TSS content as high as the Cabernet Sauvignon or Pinot noir grapes (24.5 Brix or higher) from Fasoli et al. (2018). The Dim 1-Dim 3 plot confirmed the potentiality of resolving the phenological stages of the pre-veraison phase (not shown). Pea-size and Touch berry samples from both red- and white-berry varieties well separated along Dim 3, also suggesting some variability by cultivar at the beginning-to-touch stage. Interestingly, while Soft samples confirmed their position along the model, the Harvest ones, particularly in the case of re-berry varieties, showed a shift towards the later stages. This could correspond to a specific transcriptomic state of the berry at late ripening stages that is independent on the TSS level. A similar approach was used to align the Cabernet Sauvignon samples, collected in two years from vines grown in different Italian locations (Dal Santo et al., 2018)

4. Conclusions

In this work we used the very detailed analysis of the transcriptomic changes occurring during the entire development of grape berry (cvs Cabernet Sauvignon and Pinot noir) to define a model of molecular phenology that can be useful to map the ontogenetic development of the fruit with high precision. We showed that this approach could be successful to align the developmental stage of berries of the same cultivar sampled in different vintages and of different cultivars characterized by different phenology and/or grown in different sites.

These results showed that the transcriptomic information can be accessed to precisely define a model of molecular phenology that can be used to map the ontogenetic development of the fruit with high precision and to align the stage of berry development of different grapes.

5. Acknowledgments

We thank the members of Plant Genetics Laboratory (University of Verona) and the staff of the Viticulture Chemistry and Enology Department (E. & J. Gallo Winery) for technical assistance and support.

6. Literature cited

Bobeica N., Poni S., Hilbert G., Renaud C., Gomes E., Delrot S., Dai Z., 2015. Differential responses of sugar, organic acids and anthocyanins to source-sink modulation in Cabernet Sauvignon and Sangiovese grapevines. Front Plant Sci, 6 382

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

Dal Santo S., Tornielli G.B., Zenoni S., Fasoli M., Farina L., Anesi A., Guzzo F., Delledonne M., Pezzotti M., 2013. The plasticity of the grapevine berry transcriptome. Genome Biol, 14 r54

Dal Santo S., Zenoni S., Sandri M., De Lorenzis G., Magris G., De Paoli E., Di Gaspero G., Del Fabbro C., Morgante M., Brancadoro L., Grossi D., Fasoli M., Zuccolotto P., Tornielli G.B., Pezzotti M., 2018. Grapevine field experiments reveal the contribution of genotype, the influence of environment and the effect of their interaction (GxE) on the berry transcriptome. Plant J, 93 1143-1159

Fasoli M., Dal Santo S., Zenoni S., Tornielli G.B., Farina L., Zamboni A., Porceddu A., Venturini L., Bicego M., Murino V., Ferrarini A., Delledonne M., Pezzotti M., 2012. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program. Plant Cell, 24 3489-3505

Fasoli M., Richter C.L., Zenoni S., Bertini E., Vitulo N., Dal Santo S., Dokoozlian N., Pezzotti M., Tornielli G.B., 2018. Timing and Order of the Molecular Events Marking the Onset of Berry Ripening in Grapevine. Plant Physiol, 178 1187-1206

Jaillon O., Aury J.M., Noel B., Policriti A., Clepet C., Casagrande A., Choisne N., Aubourg S., Vitulo N., Jubin C., Vezzi A., Legeai F., Hugueney P., Dasilva C., Horner D., Mica E., Jublot D., Poulain J., Bruyere C., Billault A., Segurens B., Gouyvenoux M., Ugarte E., Cattonaro F., Anthouard V., Vico V., Del Fabbro C., Alaux M., Di Gaspero G., Dumas V., Felice N., Paillard S., Juman I., Moroldo M., Scalabrin S., Canaguier A., Le Clainche I., Malacrida G., Durand E., Pesole G., Laucou V., Chatelet P., Merdinoglu D., Delledonne M., Pezzotti M., Lecharny A., Scarpelli C., Artiguenave F., Pe M.E., Valle G., Morgante M., Caboche M., Adam-Blondon A.F., Weissenbach J., Quetier F., Wincker P., French-Italian Public Consortium for Grapevine Genome C., 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature, 449 463-467

Kim D., Pertea G., Trapnell C., Pimentel H., Kelley R., Salzberg S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol, 14 R36

Lorenz D.H., Eichhorn K.W., Bleiholder H., Klose R., Meier U., Weber E., 1995. Growth Stages of the Grapevine: Phenological growth stages of the grapevine (Vitis vinifera L. ssp. vinifera)—Codes and descriptions according to the extended BBCH scale. Australian Journal of Grape and Wine Research, 1

Massonnet M., Fasoli M., Tornielli G.B., Altieri M., Sandri M., Zuccolotto P., Paci P., Gardiman M., Zenoni S., Pezzotti M., 2017. Ripening Transcriptomic Program in Red and White Grapevine Varieties Correlates with Berry Skin Anthocyanin Accumulation. Plant Physiol, 174 2376-2396

Mortazavi A., Williams B.A., McCue K., Schaeffer L., Wold B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods, 5 621-628

Roberts A., Pimentel H., Trapnell C., Pachter L., 2011. Identification of novel transcripts in annotated genomes using RNA-Seq. Bioinformatics, 27 2325-2329

Sadras V.O., Moran M.A., 2012. Elevated temperature decouples anthocyanins and sugars in berries of Shiraz and Cabernet Franc. Australian Journal of Grape and Wine Research, 18 115-122

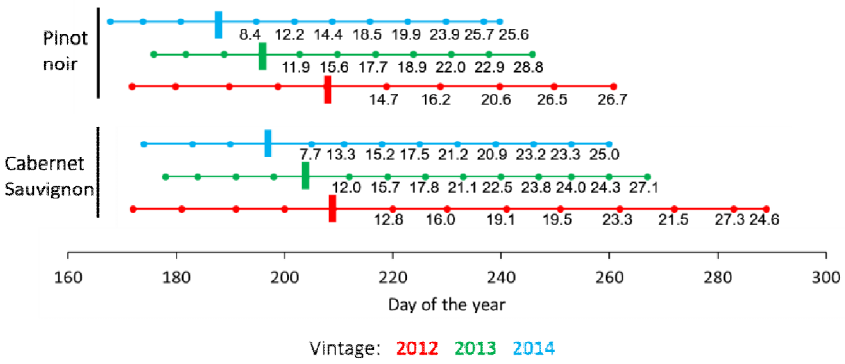

Figure 1: Sample collection details. Numbers under each round marker represent total soluble solids (°Brix). Solid vertical bars represent the 50% veraison samples.

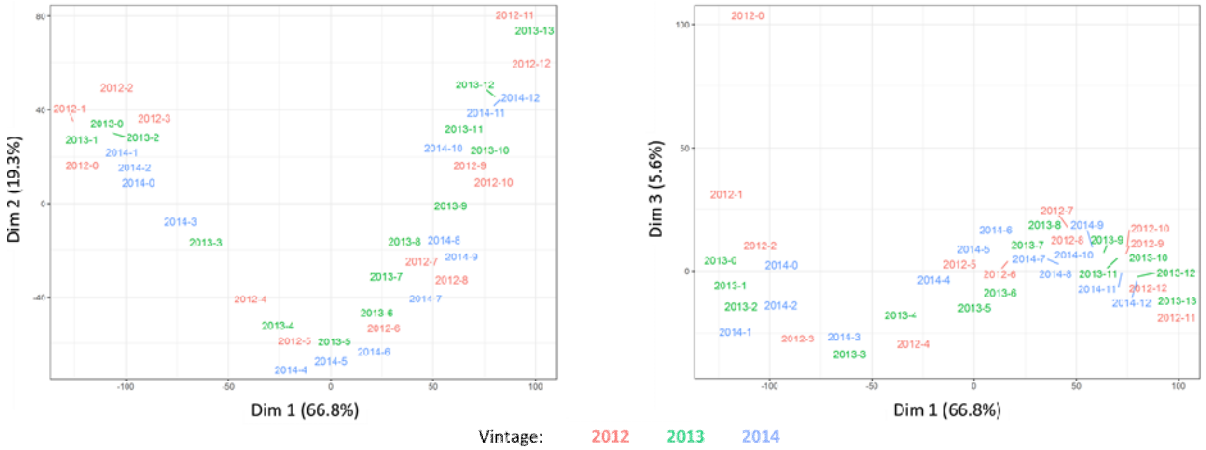

Figure 2: Principal component analysis showing the distribution of gene expression in Cabernet Sauvignon berries for all three vintages. Dim 1 and Dim 2 explain 86.1% of the variance (top plot). Dim 1 and Dim 3 explain 72.4% of the variance (bottom plot). Samples represent averaged biological triplicates.

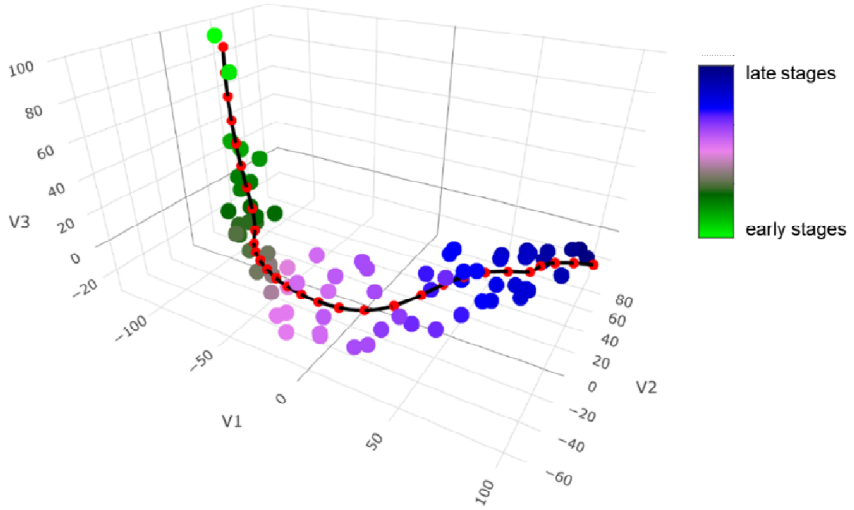

Figure 3: 3D principal component analysis showing the transcriptional ripening model generated using Cabernet Sauvignon and Pinot noir samples for all three vintages. The scale bar describes the transcriptional time from early to late stages over berry development. Equally spaced stages of transcriptomic development marked by red circles are evidenced on the curve that interpolates the data.

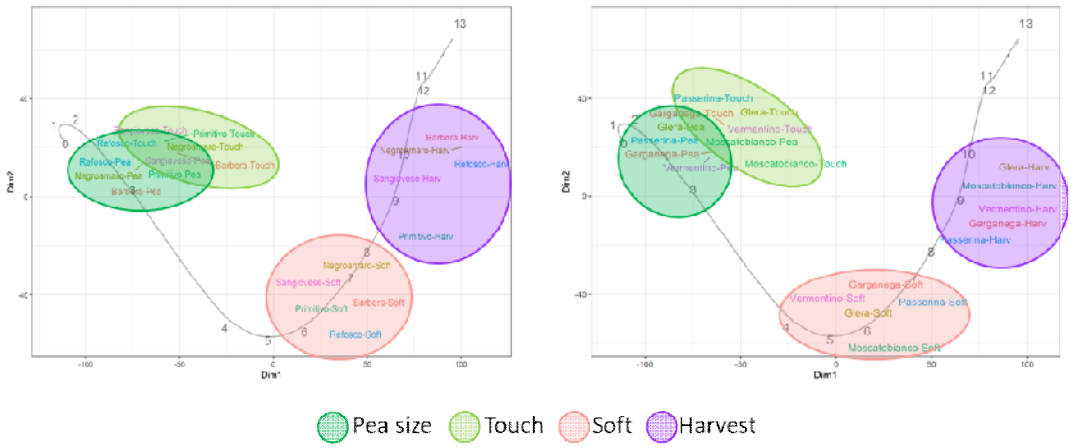

Figure 4: Principal component analysis showing the distribution of gene expression in five red-skin (left) and five white-skin (right) Italian varieties during berry development onto the transcriptomic map (Dim 1-Dim 2) from the dataset of Fasoli et al. (2018). The transcriptomic route is represented by a gray line. Berry samples of Barbera, Negroamaro, Primitivo, Refosco, Sangiovese, Garganega, Glera, Moscato bianco, Passerina, Vermentino. represent averaged biological triplicates, as reported by Massonnet et al. (2017). Stages of berry development and ripening: Pea = pea size berries (BBCH 75); Touch = beginning of bunch closure (BBCH 77); Soft = berry softening (BBCH 85); Harvest = ripe berries (BBCH 89).

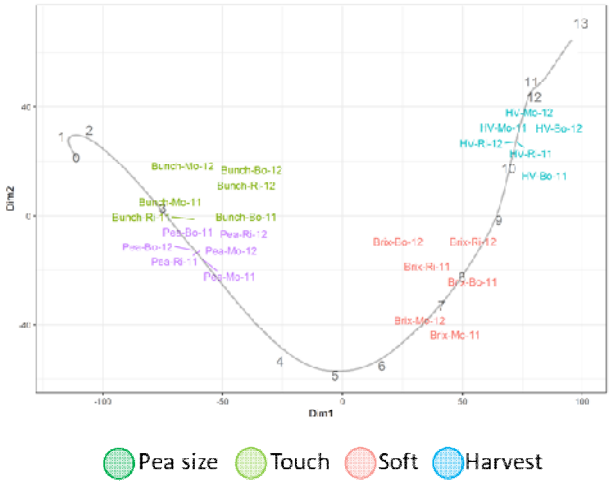

Figure 5: Principal component analysis showing the distribution of gene expression during berry development in Cabernet Sauvignon samples collected in Italy by Dal Santo and et al. (2018), onto the transcriptomic map (Dim 1-Dim 2) from the dataset of Fasoli et al. (2018), developed in this study. The transcriptomic route is represented by a gray line. Stages of berry development and ripening: Pea = pea size berries (BBCH 75); Touch = beginning of bunch closure (BBCH 77); Soft = berry softening (BBCH 85); Harvest = ripe berries (BBCH 89).