









Scaling the grape berry developmental stage by molecular phenology: Application details and caveats

Giovanni Battista Torielli^{1,2,3}  | Ron Shmulevitz³ | Alessandra Amato³  |
 Marco Sandri⁴  | Paola Zuccolotto^{4,5}  | Mario Pezzotti³  |
 Nick Dokoozlian⁶  | Marianna Fasoli³  | Sara Zenoni³ 

¹Department of Agronomy, Food, Natural Resources, Animals and Environment, University of Padova, Legnaro, Italy

²Interdepartmental Centre for Research in Viticulture and Enology (CIRVE), University of Padova, Conegliano, Italy

³Department of Biotechnology, University of Verona, Verona, Italy

⁴Big and Open Data Innovation Laboratory (BODal-Lab) and Data Methods and Systems Statistical, University of Brescia, Brescia, Italy

⁵Department of Economics and Management, University of Brescia, Brescia, Italy

⁶GALLO, Modesto, CA, USA

Correspondence

Marianna Fasoli, Department of Biotechnology, University of Verona, 37134 Verona, Italy.

Email: marianna.fasoli@univr.it

Funding information

NextGenerationEU, Grant/Award Number: SPATIALVINEproject(PRIN-PNRR2022, P20222XJKY),CUPB53D23031910001; NextGenerationEU, Grant/Award Number: AGRITECHSpoke1-VALFRUITproject, CUPB83C22002840001; Winegrowing Research department (GALLO)

[Correction added on 28 April 2024, after first online publication: The sections “Water deficit” and “Temperature,” both under “3.2 Application of MPhS,” have been reordered.]

Societal Impact Statement

Global climate change is adversely affecting grape quality by accelerating developmental processes. We developed an advanced method to define grape berry phenological stages based on gene expression information and made it accessible for widespread research. By precisely identifying these growth stages we can improve our understanding of how environmental conditions and agricultural practices influence grape quality. This has significant implications for developing mitigation strategies in viticulture and emphasizes the need for adaptive policies that can sustain grape production in the face of the changing climate.

Summary

- Determining the developmental stage of grape berries is essential to understand the effect of environmental and/or cultivation factors. Our work aims to illustrate thoroughly the application of the recently established Molecular Phenology Scale (MPhS) to any custom berry transcriptomic dataset to map the ontogenic development of the fruit with high precision.
- We detail the code components and instructions to run the MPhS package in R for berry transcriptomic datasets. We describe the output of the application of this tool on berry samples from plants subjected to various cultivation and environmental factors (e.g. cluster thinning, defoliation, water limitation, and varying temperature regimes). We illustrate the procedure for unveiling molecular responses uniquely related to the tested factor by comparing differentially expressed genes upon alignment of fruit samples by MPhS stage.
- We demonstrated that the MPhS application allows defining the shifts of fruit development driven by various agronomic and environmental factors. Moreover, by performing statistical analysis on grape samples aligned according to the MPhS, we clearly highlighted some modulation of secondary metabolism specifically triggered by crop load manipulation, beyond the anticipation or delay of developmental progression.

Giovanni Battista Torielli, Ron Shmulevitz and Alessandra Amato share equal contribution.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Plants, People, Planet* published by John Wiley & Sons Ltd on behalf of New Phytologist Foundation.

- The application of the MPhS to align time-series samples has proven to be an advanced method to define the developmental stage of grape berries. This is particularly important when the effect of cultivation/environmental factors is studied through gene expression analysis. The method is now fully available to all users for their own research scopes and applications.

KEYWORDS

agronomic practices, environmental factors, fruit development, grapevine, molecular phenology scale scripts, transcriptomics

1 | INTRODUCTION

The grape berry development involves a complex interplay of genetic regulation and environmental cues, such as climate, water availability, agronomical practices, and source-sink balance (Ollat et al., 2002). The precise definition of berry developmental stages is crucial to determine the effect of environmental and/or cultivation factors on berry ripening progression, especially in the context of climate change, which is severely impacting grapevine yield and final berry quality (Kuhn et al., 2014; van Leeuwen et al., 2019).

As in many other tree crop species, grape berry growth stages are generally defined by the assessment of visual/physical traits and classified according to widely shared phenological scales, i.e. the modified Eichhorn and Lorenz (E-L) and the extended BBCH systems (Coombe, 1995; Lorenz et al., 1995). However, the significant influence of various environmental and agronomical factors on the fruit traits used for stage delineation, makes these systems not entirely reliable for a precise definition of berry developmental stages.

We recently proposed to refine the existing development classification systems by scaling the progression of the berry transcriptional program with the aim of integrating the derived molecular information with the phenotype-based classifications (Tornielli et al., 2023). Previous transcriptomic analysis approaches highlighted that a core set of genes is consistently modulated during berry development across cultivars and growing conditions (Dal Santo et al., 2013; Fasoli et al., 2018; Massonnet et al., 2017; Theine et al., 2021). We selected about ten thousand genes representing the core developmental transcriptional program of the grape berry to build a molecular phenology scale (MPhS) and successfully map any developmental stage of the fruit with high precision (Tornielli et al., 2023). As for any phenological classification system, the MPhS stages are not time units but conceptual steps of berry development, which can take longer or shorter, depending on the complex dynamics of the berry transcriptome. The highest rate of variation across MPhS stages by time was recorded just before veraison as assessed phenotypically, while progressing through the MPhS stages characterizing the late-ripening period resulted rather slower.

The performance of the MPhS scale on different datasets was also explored by using a reduced set of genes, showing that the number of expression signals necessary for mapping samples can be

drastically reduced to a few dozen without substantial loss of precision.

In this work, we present the procedure to apply the MPhS R-script to potentially any grape berry transcriptomic dataset and provide further examples of the potential of this tool to map the developmental shift of grape berries in response to various agronomical and environmental factors. In addition, we emphasize how precisely aligning berry transcriptional stages allowed us to inspect the impact of any varying environmental or agronomical components besides their influence on developmental progression.

2 | MATERIALS AND METHODS

2.1 | Getting started with the MPhS package

As for any R package, the MPhS application will need a version of R to operate (<https://www.r-project.org>). RStudio can also be downloaded to allow a user-friendly interface (<https://posit.co/download/rstudio-desktop/>).

The usage of the MPhS tool is illustrated for Dataset S1, which serves as an example of a dataset format compatible with the scale. This dataset is a subset of the RNA-seq dataset from Massonnet et al. (2017), providing transcriptome information for 10 Italian grapevine varieties sampled at 4 berry growth stages during the same season. Berry transcriptomic samples from one red (Sangiovese) and one white grapevine variety (Passerina) were selected and included in Dataset S1, totaling 24 samples.

The MPhS package can be used with grape berry transcriptomic datasets based on RNA-seq with different expression metrics, as well as microarray analyses.

Users will need to apply a few basic R scripts to set up a working directory and instruct R to read and load their transcriptomic datasets. The R scripts and markdown for running the MPhS package are described in the results section.

2.2 | Cluster thinning datasets

The cluster thinning case studies presented with the application of the MPhS were based on Pinot noir (Green et al., 2024) and

Cabernet Sauvignon transcriptomic datasets. For both varieties, cluster thinning treatments were applied immediately following the fruit set as required to remove either 50% (CT50) or 75% (CT75) of the clusters on each vine. Unthinned vines were not manipulated (C). The experimental vineyards were organized in a randomized complete block design and the trial was conducted during three vintages (2012, 2013, 2014). Berries were collected at 10-day intervals in 2012, and weekly in 2013 and 2014, beginning at the fruit-set and continuing to harvest (total soluble solids [TSS] = 24.5°Brix). The berry sample collections amounted to 255 for Pinot noir and 324 for Cabernet Sauvignon. A list of samples can be viewed in the Dataset S2. Samples were processed for RNA extraction as described in Fasoli et al. (2012). Library preparation and RNA sequencing analysis were conducted as described in Fasoli et al. (2018).

2.3 | Statistical analysis

Prior to performing statistical analysis on the Pinot noir transcriptomic datasets, a filter of 1-RPKM (Reads Per Kilobase per Million mapped reads) value was applied as a threshold to define gene expression: gene expression < 1 RPKM across the entire dataset was considered very low/absent expression as 1 RPKM corresponds to approximately one transcript per cell (Mortazavi et al., 2008). Differentially expressed genes (DEGs) were determined by performing a *t*-test using the TMeV software (<https://sourceforge.net/projects/mev-tm4/>) with a *p*-value of 0.05.

2.4 | Gene annotation

As described in Tornielli et al. (2023), the MPhS tool was developed using a transcriptomic dataset mapped onto the PN40024 reference genome annotation V1 (available at <https://grapedia.org/files-download/>). The MPhS can handle transcriptomic datasets based on both V1 and VCOST.V3 genome annotations (see conversions at <http://www.vitviz.tombsbiolab.com/conversions/>; Navarro-Paya et al., 2022): the input argument 'geneID' in MPhStimepoints entails the options 'VIT' or 'Vitvi' for the V1 and the V3 annotations, respectively.

3 | RESULTS

3.1 | A user manual for applying the MPhS tool

MPhS is an R package that implements the statistical methods described in the work of Tornielli et al. (2023). The proposed statistical pipeline involves an unsupervised learning process that integrates semiparametric methods, smoothing techniques, and dimensionality reduction tools in an innovative way. The data mining process involves four key steps:

1. Screening—removing genes that are not expressed under certain experimental conditions or whose expression is unrelated to berry development,
2. Smoothing—applying parametric models to specific data subsets to reduce noise in gene expression patterns,
3. Principal Component Analysis—identifying stage-specific principal components for the MPhS definition, while filtering out genotype- and vintage-specific information, and
4. Molecular Phenology Scale (MPhS) definition—projecting points onto a one-dimensional Bézier curve with linear graduation.

This package allows users to map their grape berry transcriptomic datasets onto the molecular phenology scale proposed in the study. The mapping function in the MPhS package is named MPhStimepoints. To illustrate the input arguments and usage of this tool, we present a practical example, based on an RNA-seq dataset in Microsoft Excel format derived from Massonnet et al. (2017), with expression levels of 29,971 genes for 8 samples, each entailing three replicates (Dataset S1). Hereafter, we show step by step how to prepare the data and how to use the functions contained in MPhS to generate the output plot of the data projected onto the transcriptomic scale.

The procedure starts with the installation of the MPhS package by downloading from GitHub (<https://github.com/sndmrc/MPhS>) and use of the remotes package:

```
library(remotes)
install_github("sndmrc/MPhS")
```

Next, load packages MPhS, tidyr, and dplyr:

```
library(MPhS)
library(tidyr)
library(dplyr)
```

and proceed with loading the data file:

```
data("RPKMdata")
```

The structure of this dataset consists of genes by rows and samples by columns. However, MPhStimepoints require the input data being organized such that each column of the data frame represents a gene, with additional columns included to specify the experimental conditions and the berry developmental stage. Therefore, it will be necessary to transpose the data matrix.

First, the variables that define the experimental conditions and the variable that defines the maturation stage need to be created. In this dataset, such information is contained in the column names. With a few simple steps, we can construct the columns related to the cultivar, stage, and replicate, gathering them in the data frame `dts_vars`.

```
exp_cond <- names(RPKMdata)[-1]
genes <- RPKMdata$gene_id
dts_vars <- data.frame(exp_cond) %>%
separate(exp_cond,
into=c("Cultivar", "Stage", "Replicate"), sep="_")
```

At this point, we transpose the gene expression matrix and add the newly derived variables.

```
dts <- t(RPKMdata[, -1])
dts <- cbind(dts, dts_vars)
names(dts) <- c(genes, names(dts_vars))
```

Now, for each stage and each cultivar, we calculate the mean value of the three replicates.

```
dts_means <- dts %>%
group_by(Cultivar, Stage) %>%
summarize(across(all_of(genes), mean))
```

The data are now ready to be mapped onto the transcriptomic scale. The input arguments that the user must specify in `MPhStimepoints` include the data frame with the transcriptomic data, the string vector `strata_var` containing variables from the data frame that define the experimental conditions, and the string `stage_var` that specifies the name of the variable containing the development stage.

```
MPhS_out <- MPhStimepoints(data=dts_means,
strata_var="Cultivar", stage_var="Stage")
```

The object `MPhS_out` is of class `MPhStimepoints` (a list). The most relevant element contained in `MPhS_out` is the data frame `PCscores`, which is detailed below.

```
MPhS_out$PCscores
```

	PC1	PC2	PC5	timepoint	Cultivar	Stage	strata
1	-74.1	-15.1	-1.78	21	Passerina	Harv	Passerina
2	70.0	16.9	12.7	7	Passerina	Pea	Passerina
3	-48.5	-41.2	-7.70	18	Passerina	Soft	Passerina
4	57.1	19.7	-9.36	8	Passerina	Touch	Passerina
5	-88.7	19.0	8.12	24	Sangiovese	Harv	Sangiovese
6	62.4	9.69	13.9	8	Sangiovese	Pea	Sangiovese
7	-39.3	-32.5	-4.88	18	Sangiovese	Soft	Sangiovese
8	52.3	14.9	-9.22	9	Sangiovese	Touch	Sangiovese

In the timepoint column, we have the position on the MPhS scale associated with each sample. For instance, sample number 4 (namely, cultivar *Passerina* and stage *Touch*) is positioned at timepoint 8 on the

scale. The columns `PC1`, `PC2`, and `PC5` contain the principal component scores, which are the projections of the original data onto the principal component axes of the MPhS (Jolliffe, 2002).

An important input argument for the `MPhStimepoints` command is `scaling_type`. This argument defines the type of scaling applied to the expression levels of each gene. The possible options are “none” (no scaling), “scale” (standardization using the estimated mean and standard deviation for each gene), and “means_SDs” (standardization using the means and standard deviations of genes estimated in Tornielli et al., 2023). By “standardization” we mean the transformation $(x - m)/SD$, where m and SD are the mean and standard deviation, respectively.

The `MPhS_out` object can be used to visualize the position of the samples on the transcriptomic scale:

```
p <- plot(MPhS_out)
print(p)
```

The resulting graph is shown in Figure S1. The object returned by the plot command is generated using the `ggplot2` package, a powerful tool for data visualization that enables easy customization of many graphical elements. In the example below, we show how to add a title to the figure and different text for each of the subplots (facets). The resulting plot is presented in Figure S2.

```
library(ggplot2)
df_txt <- data.frame(x=c(28, 28), y=c(0.005, 0.005),
Cultivar=c("Passerina", "Sangiovese"),
txt=c("Text1", "Text2"))

p + labs(title=paste("MPhS mapping of", MPhS_out
$no_genes, "genes")) +
geom_text(data=df_txt, aes(x=x, y=y, label=txt),
color="red")
```

The commands in the MPhS package can also easily project the data on a sample-by-sample basis providing the detail of the biological replicates. Moreover, our tool can handle more recent genome sequencing and annotation versions: an unpublished RNA-seq dataset with V3 liftoff annotation on the T2T reference genome (<https://grapedia.org/files-download/>) is provided as a sample within the package. R code and resulting plots are presented in Figures S3 and S4.

For more details on the use of `MPhStimepoints` and the plot function, it is recommended to refer to the help `?MPhStimepoints` and `?MPhS:::plot.MPhStimepoints`.

3.2 | Application of MPhS

The performance of MPhS was tested on several berry transcriptomes describing berry development under different environmental conditions or following specific agronomic practices, to highlight possible shifts in the berry transcriptomic maturation program. We relied on

both previously published and novel time series datasets of berry from different grapevine varieties (Dataset S3). Each dataset was narrowed to the core set of 10,129 genes defining the berry maturation process as described in Tornielli et al. (2023) and was projected onto the MPhS. The number of genes used for each specific projection could be further reduced if any genes of the core set are not expressed or modulated across the tested transcriptomic samples.

3.2.1 | Water deficit

Two datasets that monitored berry transcripts following different irrigation regimes were explored. The first contained a total of 36 berry samples collected from potted vines of cv. Sangiovese and cv. Montepulciano (Dal Santo et al., 2016). The vines were maintained at 90% and 40% of maximum water availability, from fruit-set to veraison, to create well-watered (WW) and water-stress (WS) conditions, respectively. Transcriptomic data was analyzed at 2, 6, and 27 days after treatment (DAT), in triplicates. The MPhS projection showed no differences between the classification of the WW and the WS samples, for both cultivars (Figure 1a).

The second dataset derives from a field experiment where cv. Merlot vines were subjected to two contrasting water regimes from approximately 25 days after anthesis (DAA) (Savoi et al., 2017). Weekly irrigation, which resulted in water stem potential (Ψ_{stem}) above -0.6 MPa, was defined as control (CT). Water deficit (WD) was generated limiting irrigation to maintain the Ψ_{stem} between -1.0 and -1.4 MPa. Transcriptomic analyses were performed at five berry developmental stages: before the onset of ripening, 26 and 53 DAA; at the onset of ripening, 67 DAA; and during ripening, 81 and 106 DAA. Similarly to the previous case, the projection of the dataset onto the MPhS showed a substantially similar transcriptomic progression for CT and

WD. In fact, we observed a little shift in the MPhS stage between WD and CT berries at 26 DAA and 81 DAA, whereas the other samples were aligned (Figure 1b). Overall, the berry developmental progression appears to be little or not affected by water supply limitation.

3.2.2 | Temperature

To investigate the temperature effect on the berry maturation program we projected on the MPhS a collection of berry transcriptomic samples from potted plants of cv. Sangiovese under two divergent thermal regimes (Pastore et al., 2017). Low-temperature (LT) and high-temperature (HT) regimes, respectively characterized by an average air temperature of 21.3 and 26.4°C throughout ripening, were compared. In both conditions, berry samples were collected 7 days before veraison, at veraison, and 10, 20, and 32 days after veraison (DAV). The alignment of the sampling points on the MPhS showed a clear advanced transcriptomic maturation for the HT berries at 10 DAV and even more clearly at 20 DAV (Figure 2). Interestingly, at 32 DAV, the misalignment in the transcriptomic maturation of the HT berries was no longer detectable.

3.2.3 | Defoliation

We explored the effect of defoliation on the evolution of cv. Sangiovese berry transcriptome leveraging the dataset published by Pastore et al. (2013) and Zenoni et al. (2017). The first study compared berries from plants defoliated at pre-bloom (PB) and at veraison (V) to berries from non-defoliated plants (C). Berry samples were taken at the beginning of veraison (BV), the end of veraison (EV), and harvest (H) for the PB and C treatments, whereas, for the V treatments the transcripts were analyzed only at EV and H. The MPhS projection showed that samples derived

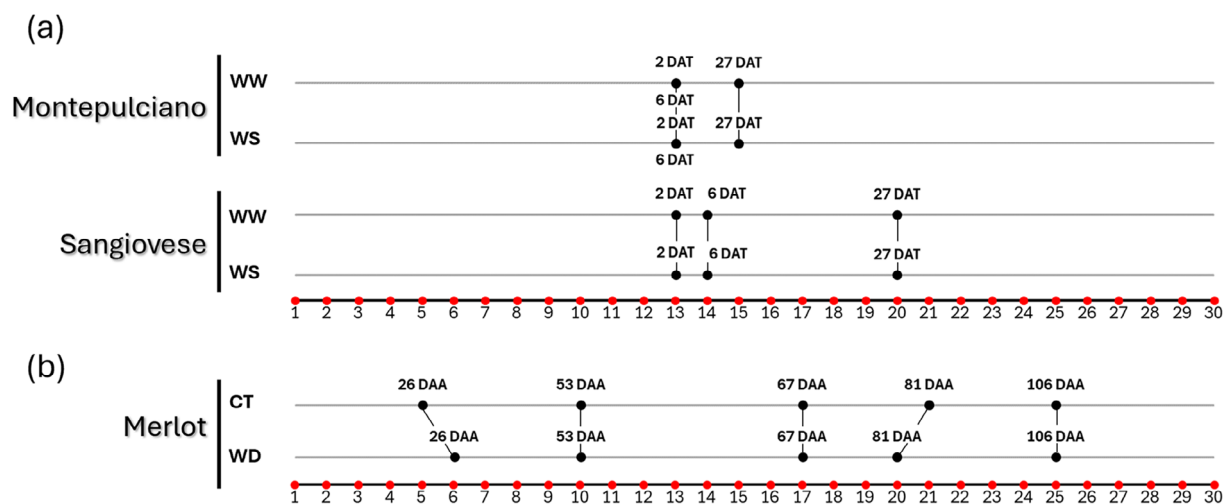


FIGURE 1 Projection of transcriptomic datasets of berries collected from three cultivars following different irrigation regimes on the molecular phenology scale (MPhS). (a) Samples of cv. Montepulciano and cv. Sangiovese under well-watered (WW) and water stress (WS) conditions. (b) Samples of cv. Merlot under control (CT) and water deficit (WD) conditions. Red dots correspond to the 30 stages of the MPhS. Black dots indicate berry samples. DAT = days after treatment; DAA = days after anthesis.

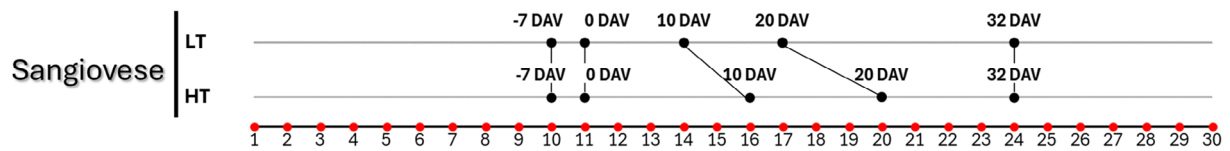


FIGURE 2 Projection of transcriptomic datasets of berries collected from cv. Sangiovese following different thermal regimes on the molecular phenology scale (MPhS). The conditions included a low temperature regime (LT) and a high temperature regime (HT). Red dots correspond to the 30 stages of the MPhS. Black dots indicate berry samples. DAV = days after veraison.

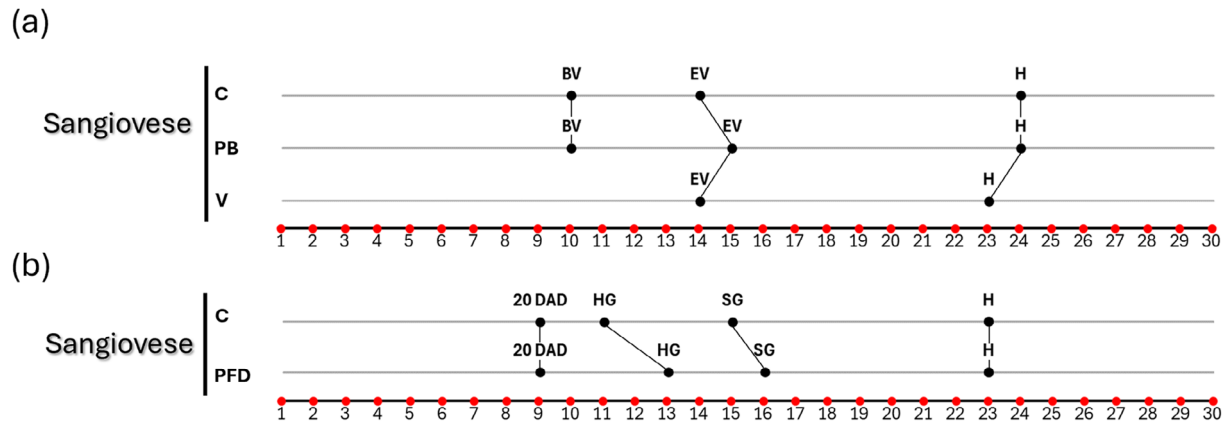


FIGURE 3 Projection of transcriptomic datasets of berries collected from cv. Sangiovese following defoliation on the molecular phenology scale (MPhS). (a) No defoliation (C), pre-bloom defoliation (PB), and defoliation at veraison (V). (b) No defoliation (C) and pre-flowering defoliation (PFD). Red dots correspond to the 30 stages of the MPhS. Black dots indicate berry samples. BV = beginning of veraison; EV = end of veraison; H = harvest; DAD = days after defoliation; HG = hard green; SG = soft green.

from different defoliation implementation timing had alignment differences (Figure 3a). The projection of the PB samples on the MPhS indicated an advanced maturation at EV compared to the C and V treatments. An opposite response was observed for the samples of V defoliation, which presented a slightly delayed maturation at H, compared to the other treatments.

The second transcriptome dataset refers to an experiment comparing pre-flowering defoliation (PFD) and non-defoliation (C) treatments (Zenoni et al., 2017). The sampling series included berries collected 20 days after defoliation (DAD), hard and green berries at veraison (HG), soft yet still not colored berries at veraison (SG), and berries at a sugar concentration of about 18°Brix (H). The transcriptomic samples projection on the MPhS revealed that the defoliation anticipated the development of berry samples collected at the HG and SG stages (Figure 3b). However, the shift was not maintained until harvest when the C and the PFD samples realigned.

3.2.4 | Cluster thinning

Three different transcriptomic datasets were used to characterize the berry development following cluster thinning practice applied in the vineyard. The first dataset was retrieved from the work of Pastore et al. (2011) performed on cv. Sangiovese, in which the removal of 50% of total clusters per vine at veraison (CT), was compared to non-thinned vines (C). Berries were sampled at the beginning of veraison (BV), the

end of veraison (EV), and harvest (H). The CT treatment accelerated the berry transcriptomic evolution during veraison as the EV sample of CT was mapped four MPhS stages ahead with respect to berries of C at the same sampling point (Figure 4a). However, the advanced maturation detected for CT at the end of veraison was no longer evident at harvest.

The two additional transcriptomic datasets mapped on the MPhS were retrieved from experiments that compared vines at two thinning levels and unthinned ones in cv. Pinot noir and cv. Cabernet Sauvignon (unpublished) and featured highly recurring berry sampling. For both cultivars, berries were collected at 10-day intervals in 2012, and weekly in 2013 and 2014, beginning at fruit set and continuing until harvest. Samples were in randomized eight-vine block designs replicated along three rows for each cluster thinning treatment, i.e. unthinned control (C), 50% cluster removal (CT50), and 75% cluster removal (CT75) (Dataset S2). The MPhS projection of the Pinot noir dataset revealed, in all three years, a marked shift forward in the alignment of most of the samples collected from CT50 thinned vines after MPhS stage 8, compared to the control, indicating an advanced maturation in response to cluster thinning (Figure 4b). The projection also revealed that intensifying the treatment (CT75) did not substantially further advance the MPhS stage of the samples. In 2013, the effect of thinning was less evident than what was observed in 2012 and 2014.

Grapes collected from thinned Cabernet Sauvignon vines showed, starting from stage 8, advanced MPhS stages for most sampling points compared to those of unthinned plants (Figure 4c). However, the shifts were clearly much smaller than in thinned Pinot noir vines,

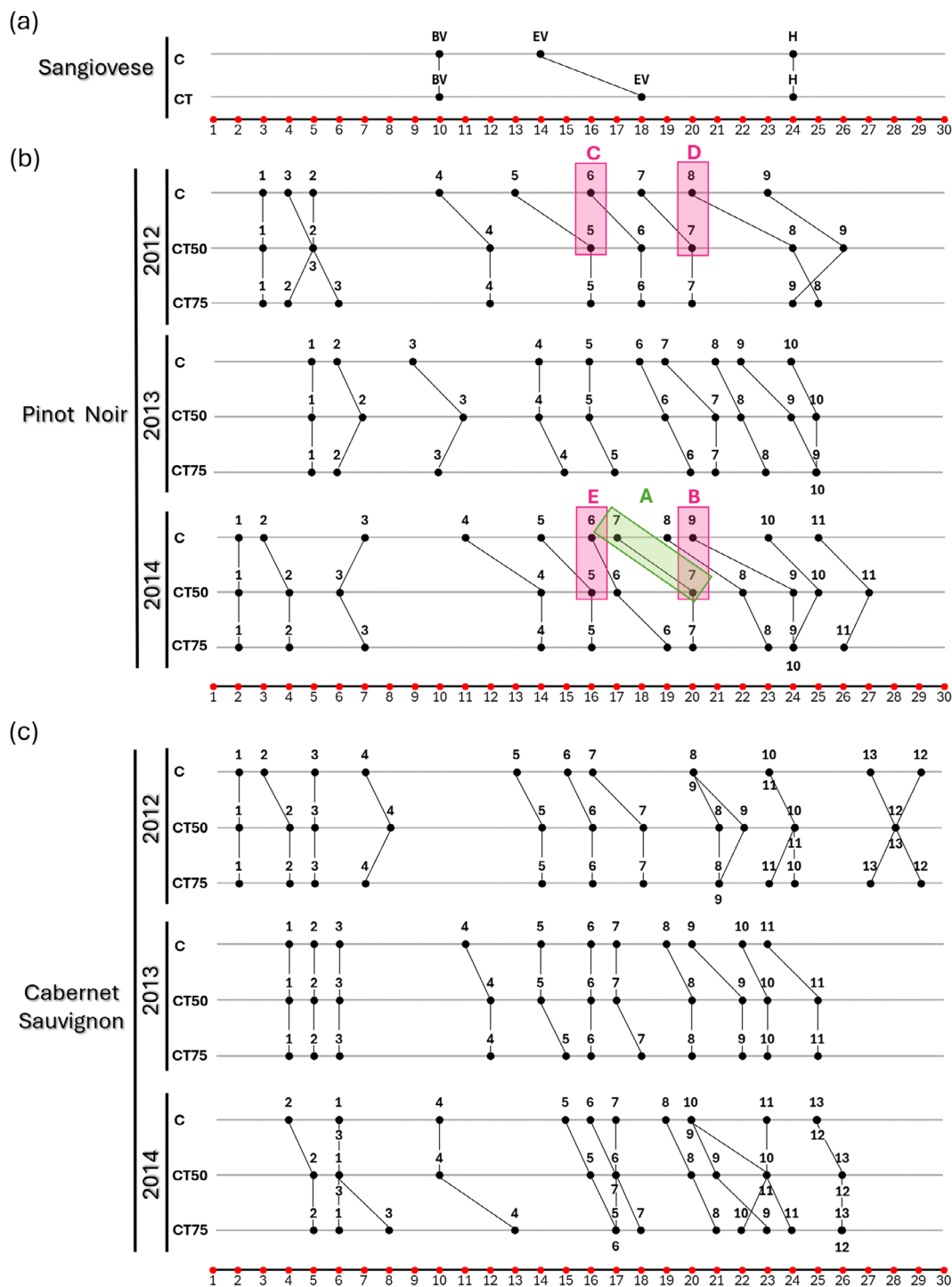


FIGURE 4 Projection of transcriptomic datasets of berries collected from three cultivars following cluster thinning on the molecular phenology scale (MPHS). (a) Samples of cv. Sangiovese following no thinning (C) and 50% cluster thinning at veraison (CT). (b) Samples of cv. Pinot noir following no thinning (C), 50% thinning (CT50), and 75% thinning (CT75) at fruit set. (c) Samples of cv. Cabernet Sauvignon following no thinning (C), 50% thinning (CT50), and 75% thinning (CT75) at fruit set. Red dots correspond to the 30 stages of the MPHS. Black dots indicate berry samples. BV=beginning of veraison; EV = end of veraison; H = harvest. Numbers indicate sampling points which are described in dataset S3. Colored rectangles indicate samples taken at the same sampling time (A; green) and samples aligned on the same MPHS stage (B-E; pink) which were investigated for DEGs employing a *t*-test.

indicating that cv. Cabernet Sauvignon was less responsive to crop load reduction. Likewise, for Pinot noir, the CT75 treatments of Cabernet Sauvignon vines in 2012 and 2013 did not boost the effect

of CT50 on the MPHS stage advancement, whereas in 2014 the severe treatment determined an onward shift of the MPHS stage compared to the CT50 samples.

4 | THE MPHS DEFINES MOLECULAR EVENTS IN GRAPE BERRIES BEYOND DEVELOPMENT PROGRESSION

The abovementioned cv. Pinot noir transcriptomic dataset, whose mapping on the MPhS is represented in Figure 4b, was further investigated by statistical analysis to assess the MPhS potential for emphasizing molecular responses specific to the crop load modulation. A

preliminary analysis focused on samples collected over the 2014 season in which a more accentuated shift was observed (Figure 4b, comparisons A and B). In detail, we determined and intersected the DEGs between sample 7 of the CT50 sets (CT50_7) with i) its time-based control (C7; comparison A) and ii) its MPhS-based control (C9; comparison B). The standard time-based comparison allowed us to identify 4,338 DEGs, whereas 2,907 were found in the MPhS-based one (Figure 5a; Dataset S4). By focusing on genes specifically

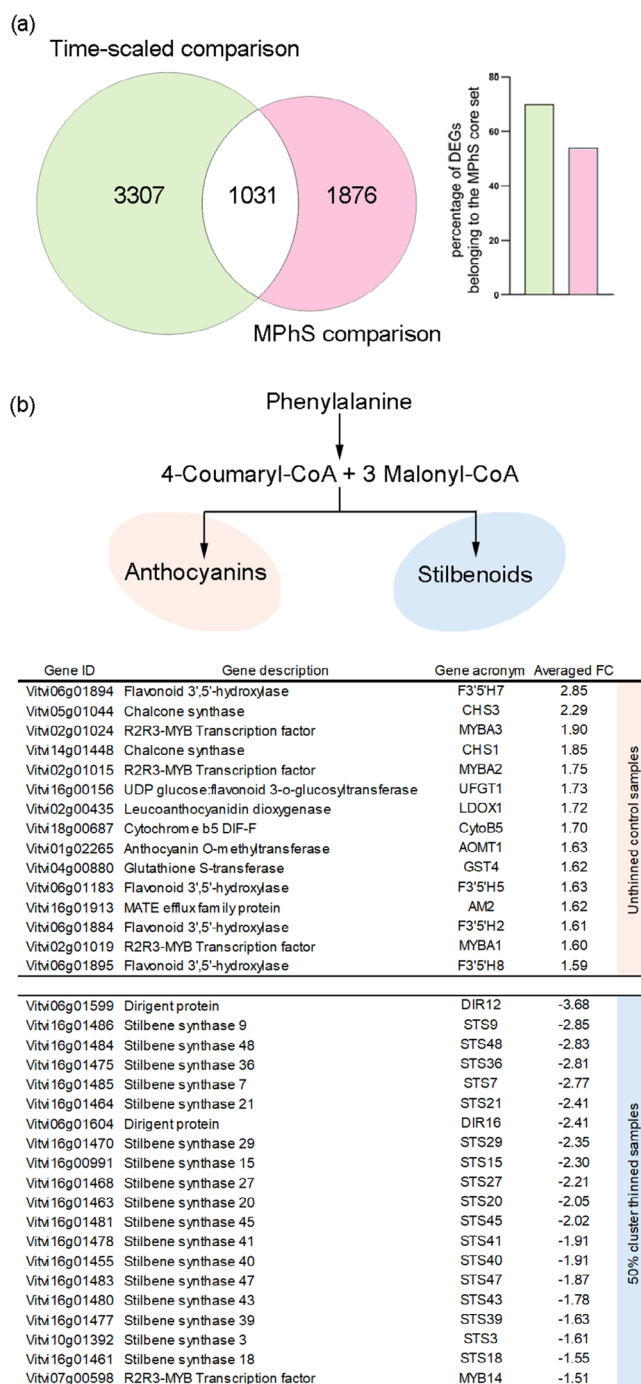


FIGURE 5 Transcriptomic inspection of cv. Pinot noir berry samples subjected to cluster thinning. (a) Venn diagram of DEGs identified comparing samples picked at the same sampling time (A; green) and those aligned on the same molecular phenology scale (MPhS) stage (B; pink). The histogram reports the percentage of specific DEGs belonging to the MPhS core set of genes. (b) Differentiation of the phenylpropanoid metabolic branches for cluster thinned (light blue) and control (light salmon) samples.

modulated in each comparison, we noticed that genes previously identified as conserved molecular markers of fruit development (MPhS core genes; Tornielli et al., 2023) represent 70% of the time-based DEGs, while in the MPhS-based comparison, the percentage drops to 54% (Figure 5a). These observations lead us to speculate that by comparing grape samples aligned according to the MPhS, it is possible to identify molecular responses genuinely related to the condition tested (i.e., heat stress), beyond any anticipation or delay in the developmental progression. We then deeply inspected differential genes that emerged in four different MPhS-aligned comparisons (Figure 4b, B-E comparisons) to highlight the molecular responses genuinely triggered by cluster thinning treatments. Genes differently modulated in at least one comparison were listed and filtered according to the average fold change (FC) value ($|FC| > 1.5$). With this approach, we found 242 genes more expressed in the CT50 samples ($FC > 1.5$) and 124 genes more expressed in the control samples ($FC < -1.5$) (Dataset S5). Among the genes up-regulated in the CT50 samples, we highlighted 14 heat shock encoding genes, three auxin-related, and many others involved in secondary metabolism. Secondary metabolism-related genes were also largely featured among those up-regulated in control samples, indeed contributing to the most represented functional category. Of note, differentiation of the phenylpropanoid pathway branches was observed between the two sample sets. Genes acting in the flavonoid branch (such as two CHSs, four F3'5'Hs, the UFGT and its regulator MYBA1) (Rinaldo et al., 2015; Walker et al., 2007) seem to be positively regulated in grape ripening samples collected under CT treatment, while genes belonging to the stilbene branch, such as 17 STSs, their regulator MYB14 and two DIRs (Holl et al., 2013; Pilati et al., 2021), were distinctly found in the unthinned samples (Figure 5b).

5 | DISCUSSION

The precise definition of the berry developmental stage is essential to understand the effect of environmental and/or cultivation factors on the ripening progression. This is mostly important when the effect of these factors is studied through gene expression analysis, to properly assign the stage of the collected berry samples and highlight the molecular responses specifically attributable to the condition tested in the experimental design (water stress, crop load, etc.). This work aims to illustrate the informative potential and the detailed procedure to apply the recently established MPhS to any custom berry transcriptomic dataset and precisely map the ontogenic development of the fruits with high detail (Tornielli et al., 2023). The MPhS relies on a core set of 10,129 genes modulated during berry development excluding those impacted by genotype, growing condition, and vintage. Therefore, genes involved in genotype-specific metabolism (e.g. anthocyanin in the red-skinned cultivars) or specifically modulated by environmental cues (e.g. disease-responsive or stress-related genes) do not contribute to the definition of the MPhS stage classification.

The R-script pipeline includes a step where the initial core set might be further reduced to the genes that are truly expressed and modulated in the input dataset thus limiting the background noise of the transcriptomic data.

We have previously shown that the MPhS was successful in mapping the berry transcriptomic developmental stage of different grapevine genotypes grown in different sites (Tornielli et al., 2023), as well as of Cabernet Sauvignon vines impacted by intra-vineyard vigor variability (Shmulevitz et al., 2024). Here we show the ability of the MPhS to discriminate samples collected from vines subjected to different agronomical practices or environmental conditions. In the case of water-stressed vines, interestingly the MPhS mapping of three different cultivars highlighted the absence of any substantial shift in the stressed vines. This may indicate that the advancement of sugar accumulation observed by the authors in the berry ripening samples (Savoi et al., 2017) could be the result of solute concentration triggered by the water limitation rather than a true developmental acceleration.

Concerning the projection of berry transcriptomes from plants grown under different thermal regimes, the sample distribution on the MPhS revealed a clear anticipation of the middle ripening steps for the berries maintained at higher temperatures. This well reflects the trends of the technological maturity parameters described in the original paper, with a significant advancement in berries subjected to elevated temperature conditions (Pastore et al., 2017).

The effect of photosynthetic stress related to the agronomical practice of defoliation has been widely described. The pre-bloom application is known to reduce the fruit set level and, therefore, to improve the grape ripening traits (VanderWeide et al., 2020). On the contrary, applying defoliation at veraison is known to reduce the carbohydrate supply to the berries, thus delaying the sugar accumulation during ripening (Poni et al., 2013). Our attempt to test the performance of MPhS on berry transcriptomic samples from defoliated and non-defoliated plants was overall in line with such an outlook. In fact, the two cases of pre-bloom defoliations evidenced an anticipated MPhS stage during early ripening, whereas the veraison treatment delayed the development of the last sampling point.

Another source-sink manipulation practice adopted by viticulturists to improve grape quality at ripening is represented by cluster thinning, which consists of the removal of a sizable percentage of bunches before the onset of ripening (VanderWeide et al., 2024). We explored the effect of cluster thinning by mapping berry transcriptome datasets from three varieties and unveiled a general developmental anticipation effect for this practice. The MPhS mapping of the Pinot noir and Cabernet Sauvignon datasets allowed to highlight (i) a greater response of Pinot noir to the thinning, (ii) a generally weak or negligible effect of the 75% compared to the 50% bunch removal, and (iii) differences across vintages of cluster thinning application for both varieties. Furthermore, the weekly sampling of these very detailed datasets allowed us to determine the MPhS stage at which the developmental advancement effect could be detected. Focusing on Pinot noir dataset, we could identify pairs of fruit samples aligned by MPhS

stage and extract differentially expressed genes to unveil those responses uniquely related to the CT50 treatment, beyond the anticipation of developmental progression. By this approach, we clearly highlighted the modulation of secondary metabolism specifically triggered by crop load manipulation consisting of the induction of genes related to anthocyanin synthesis in berries from thinned vines and the induction of genes related to stilbenoid metabolism in control berries.

When comparing samples re-aligned by MPhS, it should be considered that DEGs related to the timing of collection, most likely characterized by different climatic conditions, could also emerge. In our case, the several heat shock proteins upregulated in berries from CT50 vines could be the result of a response to warmer conditions as berries were collected about two weeks earlier than the control. On the other hand, the comparison of samples collected at the same time and not re-aligned on MPhS evidenced a higher number of DEGs including a greater percentage of genes belonging to the MPhS core set (i.e., ideally, genes that are consistently modulated during development in most cultivars), once more evidencing that the MPhS-alignment of grape samples allows identifying metabolisms and processes specifically affected by the treatment, beyond the developmental shift.

We believe that making the MPhS scripts and procedure fully available to all users for their research purposes will provide the scientific community with a valuable tool for interpreting the impact of agronomical, environmental, and genetic variables on berry ripening dynamics.

AUTHOR CONTRIBUTIONS

Giovanni Battista Tornielli, Sara Zenoni, and Marianna Fasoli conceived and designed the study, discussed the data, and wrote/edited the manuscript. **Ron Shmulevitz** tested the performance of the MPhS, discussed the data, and wrote the manuscript. **Alessandra Amato** investigated transcriptomic changes, discussed the data, and wrote the manuscript. **Marco Sandri** and **Paola Zuccolotto** created and described the R-based MPhS scripts. **Mario Pezzotti** and **Nick Dokoozlian** designed the experimental plan regarding the effect of crop load modulations.

ACKNOWLEDGEMENTS

We thank the staff of the Winegrowing Research department (GALLO) for technical assistance and support: Chandra Richter for guidance and leadership; Donnie Katayama and Octavio Viveros for supporting the fieldwork and for sharing their expertise on grapevine growth and physiology, and Nona Ebisuda for facilitating the grape sample collection and harvest. Open access publishing facilitated by Università degli Studi di Verona, as part of the Wiley - CRUI-CARE agreement.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the supplementary material of this article.

The berry transcriptomes projected onto the MPhS referred to: Dal Santo et al., 2016 (data accession number GSE70670) and Savoi et al., 2017 (data accession number PRJNA348618) for water deficit conditions; Pastore et al., 2017 (data accession number GSE92864) for heat stress conditions; Pastore et al., 2013 (data accession number GSE40487) and Zenoni et al., 2017 (data accession number GSE92980) for defoliation treatments; Pastore et al., 2011 (data accession number GPL13936), Fasoli et al., 2018 (data accession number GSE98923), data accession numbers GSE101532 (Pinot noir samples), and GSE104316 (Cabernet Sauvignon samples) for cluster thinning treatments.

ORCID

Giovanni Battista Tornielli  <https://orcid.org/0000-0001-5027-0269>

Alessandra Amato  <https://orcid.org/0000-0001-8178-6102>

Marco Sandri  <https://orcid.org/0000-0002-1422-5695>

Paola Zuccolotto  <https://orcid.org/0000-0003-4399-7018>

Mario Pezzotti  <https://orcid.org/0000-0002-7430-6147>

Nick Dokoozlian  <https://orcid.org/0000-0002-1525-1967>

Marianna Fasoli  <https://orcid.org/0000-0003-1101-8129>

Sara Zenoni  <https://orcid.org/0000-0002-0496-8161>

REFERENCES

- Coombe, B. G. (1995). Adoption of a system for identifying grapevine growth stages. *Australian Journal of Grape and Wine Research*, 1, 100–110. <https://doi.org/10.1111/j.1755-0238.1995.tb00086.x>
- Dal Santo, S., Fasoli, M., Negri, S., D'Inca, E., Vicenzi, N., Guzzo, F., Tornielli, G. B., Pezzotti, M., & Zenoni, S. (2016). Plasticity of the berry ripening program in a white grape variety. *Frontiers in Plant Science*, 7, 970. <https://doi.org/10.3389/fpls.2016.00970>
- 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 Biology*, 14(6), r54. <https://doi.org/10.1186/gb-2013-14-6-r54>
- 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(9), 3489–3505. <https://doi.org/10.1105/tpc.112.100230>
- Fasoli, M., Richter, C. L., Zenoni, S., Bertini, E., Vitulo, N., Dal Santo, S., Dokoozlian, N., Pezzotti, M., & Tornielli, G. B. (2018). Timing and order of the molecular events marking the onset of berry ripening in grapevine. *Plant Physiology*, 178(3), 1187–1206. <https://doi.org/10.1104/pp.18.00559>
- Green, E., Shmulevitz, R., Amato, A., Tornielli, G. B., Dokoozlian, N., & Fasoli, M. (2024). Unraveling the key molecular events of pinot noir berry ripening under varying crop load. *Scientia Horticulturae*, 338, 113644. <https://doi.org/10.1016/j.scienta.2024.113644>
- Holl, J., Vannozzi, A., Czemmel, S., D'Onofrio, C., Walker, A. R., Rausch, T., Lucchin, M., Boss, P. K., Dry, I. B., & Bogs, J. (2013). The R2R3-MYB transcription factors MYB14 and MYB15 regulate stilbene biosynthesis in *Vitis vinifera*. *Plant Cell*, 25(10), 4135–4149. <https://doi.org/10.1105/tpc.113.117127>
- Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). Springer series in statistics. Springer.

- Kuhn, N., Guan, L., Dai, Z. W., Wu, B. H., Lauvergeat, V., Gomes, E., Li, S. H., Godoy, F., Arce-Johnson, P., & Delrot, S. (2014). Berry ripening: Recently heard through the grapevine. *Journal of Experimental Botany*, 65(16), 4543–4559. <https://doi.org/10.1093/jxb/ert395>
- Lorenz, D. H., Eichhorn, K. W., Bleiholder, H., Klöse, 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, 100–103. <https://doi.org/10.1111/j.1755-0238.1995.tb00085.x>
- 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 Physiology*, 174(4), 2376–2396. <https://doi.org/10.1104/pp.17.00311>
- Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 5(7), 621–628. <https://doi.org/10.1038/nmeth.1226>
- Navarro-Paya, D., Santiago, A., Orduna, L., Zhang, C., Amato, A., D'Inca, E., Fattorini, C., Pezzotti, M., Tornielli, G. B., Zenoni, S., Rustenholz, C., & Matus, J. T. (2022). The grape gene reference catalogue as a standard resource for gene selection and genetic improvement. *Frontiers in Plant Science*, 12, 803977. <https://doi.org/10.3389/fpls.2021.803977>
- Ollat, N., Carde, J.-P., Gaudillère, J.-P., Barrieu, F., Diakou-Verdin, P., & Moing, A. (2002). Grape berry development: A review. *OENO One*, 36(3), 22. <https://doi.org/10.20870/oeno-one.2002.36.3.970>
- Pastore, C., Dal Santo, S., Zenoni, S., Movahed, N., Allegro, G., Valentini, G., Filippetti, I., & Tornielli, G. B. (2017). Whole plant temperature manipulation affects flavonoid metabolism and the transcriptome of grapevine berries. *Frontiers in Plant Science*, 8, 929. <https://doi.org/10.3389/fpls.2017.00929>
- Pastore, C., Zenoni, S., Fasoli, M., Pezzotti, M., Tornielli, G. B., & Filippetti, I. (2013). Selective defoliation affects plant growth, fruit transcriptional ripening program and flavonoid metabolism in grapevine. *BMC Plant Biology*, 13, 30. <https://doi.org/10.1186/1471-2229-13-30>
- Pastore, C., Zenoni, S., Tornielli, G. B., Allegro, G., Dal Santo, S., Valentini, G., Intrieri, C., Pezzotti, M., & Filippetti, I. (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>
- Pilati, S., Malacarne, G., Navarro-Payá, D., Tomè, G., Riscica, L., Cavecchia, V., Matus, J. T., Moser, C., & Blanzieri, E. (2021). *Vitis* OneGenE: A causality-based approach to generate gene networks in *Vitis vinifera* sheds light on the laccase and Dirigent gene families. *Biomolecules*, 11(12), 1744. <https://doi.org/10.3390/biom11121744>
- Poni, S., Gatti, M., Bernizzoni, F., Civardi, S., Bobeica, N., Magnanini, E., & Palliotti, A. (2013). Late leaf removal aimed at delaying ripening in cv. Sangiovese: Physiological assessment and vine performance. *Australian Journal of Grape and Wine Research*, 19(3), 378–387. <https://doi.org/10.1111/ajgw.12040>
- Rinaldo, A. R., Cavallini, E., Jia, Y., Moss, S. M., McDavid, D. A., Hooper, L. C., Robinson, S. P., Tornielli, G. B., Zenoni, S., Ford, C. M., Boss, P. K., & Walker, A. R. (2015). A grapevine anthocyanin acyltransferase, transcriptionally regulated by VvMYBA, can produce Most Acylated anthocyanins present in grape skins. *Plant Physiology*, 169(3), 1897–1916. <https://doi.org/10.1104/pp.15.01255>
- Savoi, S., Wong, D. C. J., Degu, A., Herrera, J. C., Bucchetti, B., Peterlunger, E., Fait, A., Mattivi, F., & Castellari, S. D. (2017). Multi-omics and integrated network analyses reveal new insights into the systems relationships between metabolites, structural genes, and transcriptional regulators in developing grape berries (*Vitis vinifera* L.) exposed to water deficit. *Front. Plant Science*, 8, 1124. <https://doi.org/10.3389/fpls.2017.01124>
- Shmulevitz, R., Amato, A., Previtali, P., Green, E., Sanchez, L., Alsina, M. M., Dokoozlian, N., Tornielli, G. B., & Fasoli, M. (2024). Spatial variability of grape berry maturation program at the molecular level. *Horticulturae*, 10(3), 238. <https://doi.org/10.3390/horticulturae10030238>
- Theine, J., Holtgrawe, D., Herzog, K., Schwander, F., Kicherer, A., Hausmann, L., Viehover, P., Topfer, R., & Weisshaar, B. (2021). Transcriptomic analysis of temporal shifts in berry development between two grapevine cultivars of the pinot family reveals potential genes controlling ripening time. *BMC Plant Biology*, 21(1), 327. <https://doi.org/10.1186/s12870-021-03110-6>
- Tornielli, G. B., Sandri, M., Fasoli, M., Amato, A., Pezzotti, M., Zuccolotto, P., & Zenoni, S. (2023). A molecular phenology scale of grape berry development. *Horticulture Research*, 10(5), uhad048. <https://doi.org/10.1093/hr/uhad048>
- van Leeuwen, C., Destrac-Irvine, A., Dubernet, M., Duchêne, E., Gowdy, M., Marguerit, E., Pieri, P., Parker, A., de Ressaiguier, L., & Ollat, N. (2019). An update on the impact of climate change in viticulture and potential adaptations. *Agronomy*, 9(9), 514. <https://doi.org/10.3390/agronomy9090514>
- VanderWeide, J., Gottschalk, C., Schultze, S. R., Nasrollahiazar, E., Poni, S., & Sabbatini, P. (2020). Impacts of pre-bloom leaf removal on wine grape production and quality parameters: A systematic review and meta-analysis. *Frontiers in Plant Science*, 11, 621585. <https://doi.org/10.3389/fpls.2020.621585>
- VanderWeide, J., Nasrollahiazar, E., Schultze, S., & Sabbatini, P. & Castellari, S. D. (2024). Impact of cluster thinning on wine grape yield and fruit composition: A review and meta-analysis. *Australian Journal of Grape and Wine Research*, 1, 2504396. <https://doi.org/10.1155/2024/2504396>
- Walker, A. R., Lee, E., Bogs, J., McDavid, D. A., Thomas, M. R., & Robinson, S. P. (2007). White grapes arose through the mutation of two similar and adjacent regulatory genes. *The Plant Journal*, 49(5), 772–785. <https://doi.org/10.1111/j.1365-3113.2006.02997.x>
- Zenoni, S., Dal Santo, S., Tornielli, G. B., D'Inca, E., Filippetti, I., Pastore, C., Allegro, G., Silvestroni, O., Lanari, V., Pisciotta, A., Di Lorenzo, R., Palliotti, A., Tombesi, S., Gatti, M., & Poni, S. (2017). Transcriptional responses to pre-flowering leaf defoliation in grapevine berry from different growing sites, years, and genotypes. *Frontiers in Plant Science*, 8, 630. <https://doi.org/10.3389/fpls.2017.00630>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Tornielli, G. B., Shmulevitz, R., Amato, A., Sandri, M., Zuccolotto, P., Pezzotti, M., Dokoozlian, N., Fasoli, M., & Zenoni, S. (2025). Scaling the grape berry developmental stage by molecular phenology: Application details and caveats. *Plants, People, Planet*, 1–11. <https://doi.org/10.1002/ppp3.70018>