Transcriptomic analysis of temporal shifts in berry development between two grapevine cultivars of the Pinot family reveals potential ripening-regulative genes

Jens Theine1, Daniela Holtgräwe1, Katja Herzog2, Florian Schwander2, Anna Kicherer2, Ludger Hausmann2, Prisca Viehöver1, Reinhard Töpfer2, Bernd Weisshaar1*

1 Bielefeld University, Genetics and Genomics of Plants, Faculty of Biology & Center for Biotechnology, Bielefeld

2 Julius Kühn-Institute, Institute for Grapevine Breeding Geilweilerhof, Siebeldingen

*corresponding author:

BW Bernd Weisshaar (0000-0002-7635-3473); bernd.weisshaar@uni-bielefeld.de

JT Jens Theine (0000-0002-6042-373X); jens.theine@uni-bielefeld.de
DH Daniela Holtgräwe (0000-0002-1062-4576); dholtgra@cebitec.uni-bielefeld.de
KH Katja Herzog; katja.herzog@julius-kuehn.de
FS Florian Schwander; florian.schwander@julius-kuehn.de
AK Anna Kicherer; anna.kicherer@julius-kuehn.de
LH Ludger Hausmann (0000-0002-4046-9626); ludger.hausmann@julius-kuehn.de
PV Prisca Viehöver; viehoeve@cebitec.uni-bielefeld.de
RT Reinhard Töpfer (0000-0003-1569-2495); reinhard.toepfer@julius-kuehn.de

Abstract

Background

Grapevine cultivars of the Pinot family represent in the broader sense clonally propagated mutants with clear-cut phenotypes, such as different color or shifted ripening time, that result in major phenotypic and physiological differences as well as changes in important viticultural traits. Specifically, the cultivars 'Pinot Noir' (PN) and 'Pinot Noir Precoce' (PNP, early ripening) flower at the same time, but vary for the beginning of berry ripening (véraison) and consequently for the harvest time. Apart from the genotype, seasonal climatic conditions (i.e. high temperatures) also affect ripening times. To reveal possible ripening-regulatory genes affecting the timing of the start of ripening, we investigated differences in gene expression profiles between PN and PNP throughout berry development with a closely meshed time series and in two years.

Results

The difference in the duration of berry formation between PN and PNP was quantified to be about two weeks under the growth conditions applied, using plant material with a proven clonal relationship of PN and PNP. Clusters of co-expressed genes and differentially expressed genes (DEGs) were detected which reflect the shift in the beginning of ripening at the level of gene expression profiles. Functional annotation of these DEGs fits to phenotypic and physiological changes during berry development. In total, we observed between PN and PNP 3,342 DEGs in 2014 and 2,745 DEGs in 2017. The intersection of both years comprises 1,923 DEGs. Among these, 388 DEGs were identified as véraison-specific and 12 were considered as candidates for a regulatory effect on berry ripening time. The expression profiles revealed two candidate genes for Ripening Time Control, designated *VviRTIC1* and

VviRTIC2 (VIT_210s0071g01145 and VIT_200s0366g00020, respectively) that may contribute to controlling the phenotypic difference between PN and PNP.

Conclusions

Many of the 1,923 DEGs identified show highly similar expression profiles in both cultivars as far as accelerated berry formation of PNP is concerned. Putative ripening-regulatory genes differentially expressed between PNP and PN as well as véraison-specific genes were identified. We point out potential connections of these genes to molecular events during berry development and discuss potential ripening time controlling candidate genes, two of which are already differentially expressed in the early berry development phase. Several down-regulated genes are annotated to encode auxin response factors / ARFs. Conceivably, changes in auxin signaling may realize the earlier ripening phenotype of PNP.

Key words

Vitis vinifera, Pinot Noir, Pinot Noir Precoce, grapevine, berry ripening, fruit development, differential gene expression, transcriptome profiling, ripening time control, véraison

Background

Vitis vinifera subsp. vinifera (grapevine) belongs to the family *Vitaceae*. With 6,000 to 11,000 cultivars, it is one of the most important perennial crops worldwide [1]. Grapevine fruit development can be divided into two physiological phases, berry formation and berry ripening. The time of véraison refers to the transition of both phases whereby each phase is represented by a sigmoidal growth curve of development [2]. The progress through development is described by stages referred to as "BBCH stages" that have been defined for several crops including grapevine [3, 4]. The first physiological phase described as berry formation (berry initiation and growth with cell divisions) is lasting from the end of flowering (BBCH71) until the majority of berries are touching each other (BBCH79), approximately 60 days later. The developmental stage of véraison (BBCH81) depicts the end of berry formation and the start of berry ripening [2]. Phenotypically, véraison is the developmental switch when the berries start to soften, accompanied by the onset of accumulation of phenylpropanoids. In red grapevine cultivars, véraison is also indicated by a color change of the berries that is caused by the beginning of accumulation of anthocyanins which are one class of phenylpropanoids. Members of the well-studied protein superfamily of R2R3-MYB transcription factors (TFs) are considered to be mainly accountable for controlling anthocyanin accumulation [5-7]. After véraison, berry ripening continues until harvest (BBCH89) with cell enlargement, sugar accumulation and acidity decline.

Anthropogenic climate change is resulting in successively earlier ripening of grapes with a significant impact on berry quality and on typicity of a desired wine style [8]. In addition, the time of véraison and for harvest of a given cultivar may differ greatly, driven by regional and/or year-depending differences in weather conditions. Obviously, this calls for a better molecular understanding of the control of ripening time in grapevine.

Comparison of grapevine cultivars grown at the same environmental conditions often uncovers differences in ripening time. However, the underlying genetic factors are mostly unknown. Previous studies have elucidated how ripening time is affected by external factors. For example, the effect of phytohormones on berry ripening has been widely studied [1]. In general, fruit growth is discussed to be controlled by several phytohormones, which play essential roles to trigger or delay ripening processes [9]. In grapevine as a non-climacteric fruit, effects of abscisic acid (ABA) have been investigated in many studies and ABA is considered to trigger ripening [10-12]. Furthermore, it was shown that ABA is involved in controlling leaf senescence [13], responses to drought [14] and pathogen defense [15]. In grapevine, although not as central as in climacteric fruits like tomato (*Solanum lycopersicum*), the phytohormone ethylene is also involved in the control of berry ripening [1, 11, 16, 17]. Application studies on grapevine berries indicate that auxin has a negative effect on berry ripening in that it induces a delay of ripening [18, 19].

Fruit development of both, dry and fleshy fruits, has been studied very intensively for the obvious reason that fruits are central to human nutrition [20, 21]. The main model system for studies on fleshy fruits is tomato, because of established genetics and molecular biology, access to mutants, and well advanced transgenic approaches to gene function identification [22, 23]. Berry development of grapevines has also been studied intensively [1, 24] and often at the level of the transcriptome. In quite some of the studies, predominantly late berry development stages were sampled to put véraison into the focus [25-28]. In addition, whole berry development was studied with coarse time point distribution [29-33].

To monitor gene expression changes at a high resolution throughout grapevine berry development, starting from flowering until berries are matured, we sampled a comprehensive time series from two Pinot cultivars and in two years. The samples were

collected from the grapevine cultivar 'Pinot Noir' (PN) and the comparably earlier ripening cultivar 'Pinot Noir Precoce' (PNP) that is expected to be closely related to PN. The cultivar PNP is listed in the *Vitis* International Variety Catalogue (*V*IVC; [34]) and described to flower at the same time as PN but to reach véraison significantly earlier than PN [35]. Quantitative data for transcript levels, interpreted as values for gene expression, were generated by RNA-Seq. We studied the general course of gene expression patterns throughout berry development in both years and cultivars, and identified a number of differentially expressed genes (DGEs) between PN and PNP prior to véraison. These DEGs can be considered as important candidates for either delaying or pushing forward berry development. Our main aim was the identification of genes controlling the speed of development, to offer an entry point into characterization of the relevant molecular functions in grapevine, and to facilitate breeding in the future that addresses traits relevant to and affected by climate change.

Results

Phenotypical comparison between two Pinot cultivars

To study ripening shifts, we used samples of two closely related grapevine cultivars. The cultivar PNP is an earlier ripening clonal variant of its ancestor PN. Clonal relation of PN and PNP was confirmed by a set of 24 SSR markers that all displayed the identical allele status for both cultivars (Additional file 1: Table S1). To confirm and validate the phenotypic differences between PN and PNP, detailed BBCH developmental stages were determined and documented (Figure 1). PN and PNP display similar phenotypical properties during development and flower (BBCH65) at the same time. However, véraison (BBCH81) is shifted to about two weeks earlier for PNP and similar shifts were detected in four documented years (Table 1). In addition, Figure 1A shows an overview over the time points at which

samples were taken. For a visual documentation of the phenotypic differences between PN and PNP, images of developing berries were taken between onset of berry formation and véraison (Figure 1B and Additional file 1: Table S2 and Table S3). Again, véraison (BBCH81), visible on the images as the begin of anthocyanin accumulation, is detected approximately two weeks later in PN compared to PNP.

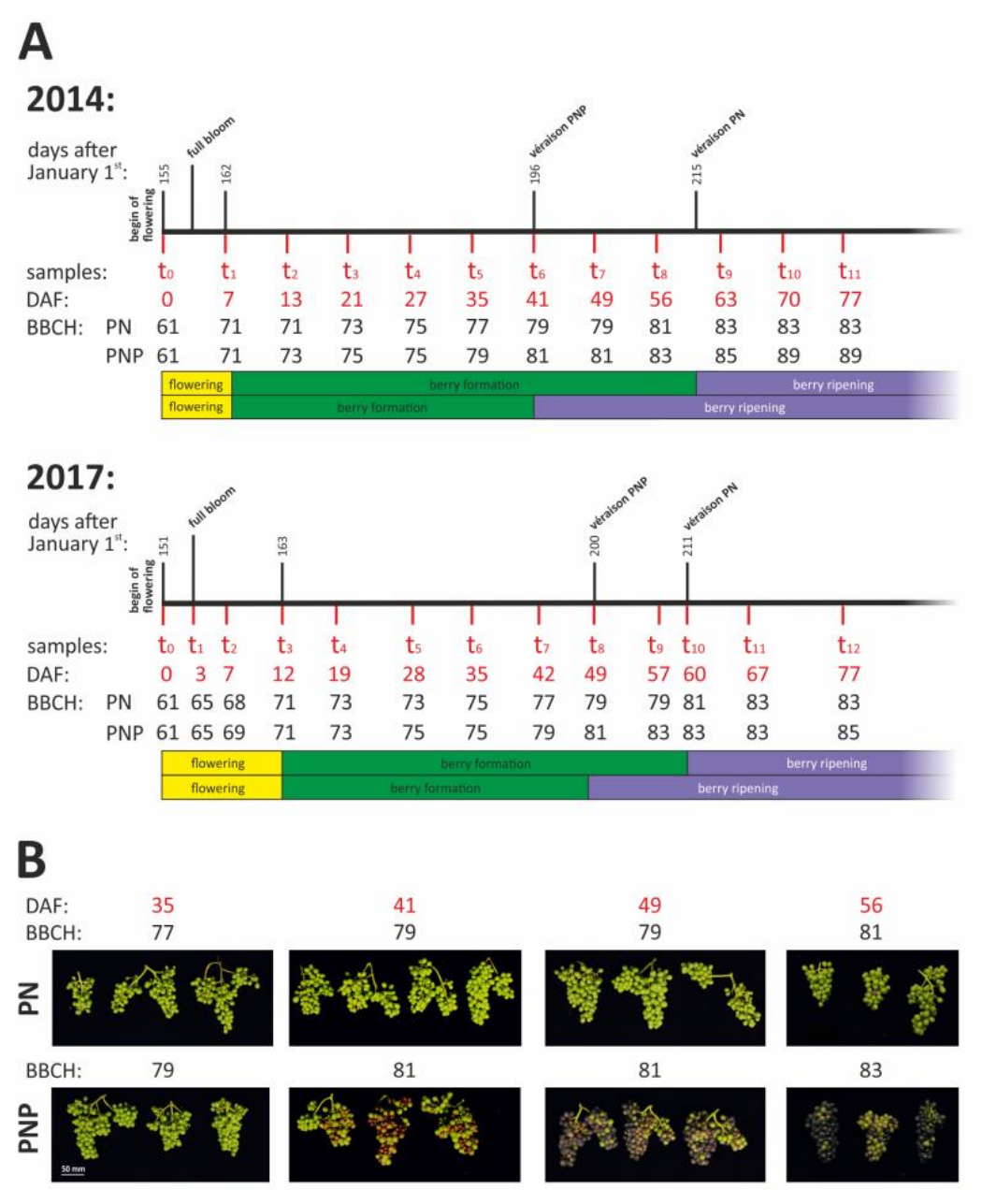


Figure 1: Phenotypical observations and sampling scheme. (A) Sampling time points and days after begin of flowering (DAF) are indicated in red. The developmental stage observed is shown in the BBCH stages [3, 4]. Berry development is depicted schematically and categorized into the phases flowering (yellow), berry formation (green), and berry ripening (purple) for both cultivars. The junction between green and purple indicates véraison (BBCH81). As an orientation for time of the year, days after January 1st are shown for distinct transitions of the phases. (B) Images of grape bunches and developing berries taken in 2014 to document the difference between PN and PNP. Images were taken 35, 41, 49 and 56 DAF. Scale bar: 50 mm.

Table 1: Observed flowering- and berry development shifts between the cultivars PNP and PN in 2014, 2015, 2016 and 2017 at the vineyards Geilweilerhof, Siebeldingen (in days after January 1st).

Year	Cultivar	Start of flowering (BBCH61)	End of flowering period / Start of berry formation (BBCH71)	End of berry formation / véraison (BBCH81)	Flowering time [Δ days]	Berry formation time [Δ days]
2014	PNP	155	162	196	0	19
2014	PN	155	162	215		
2015	PNP	159	166	201	7	14
2015	PN	159	173	222		
2016	PNP	171	180	215	0	14
2016	PN	171	180	229		
2017	PNP	151	163	200	0	11
2017	PN	151	163	211		

Global view on gene expression patterns

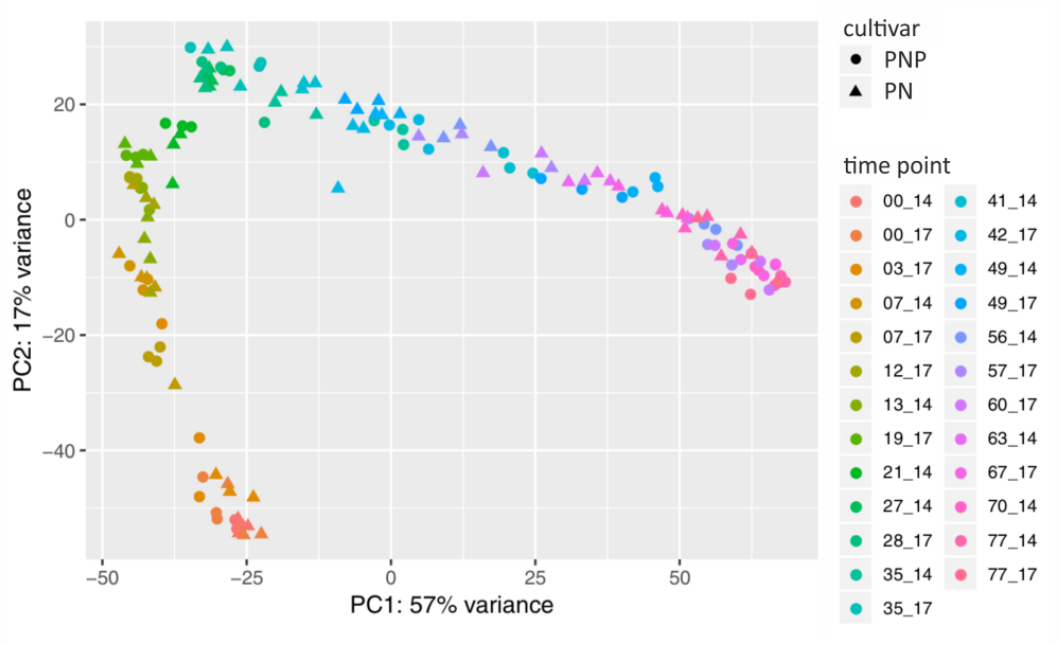
We harvested the samples in 2014 and 2017 (for time points see Figure 1 and Additional file 1 Tables S2 and S3) in triplicates (individual harvests are referred to below as subsamples) and analyzed them by RNA-Seq. After preprocessing of the raw data (see Methods), the reads derived from each subsample were mapped to the reference sequence from PN40024

and analyzed with respect to the CRIBI V2.1 annotation dataset. For 2014, approximately 19.7 million reads per subsample were obtained from each of the 72 libraries. An overall alignment rate of 79 % to the grape reference genome sequence was reached. For 2017, approximately 43.5 million reads per subsample were obtained from each of the 78 libraries. From these, an overall alignment rate of 92 % to the reference was calculated. Expression values were initially detected as Transcripts Per Kilobase Million (TPM) and averaged over the three subsamples for each sample. Considering both years separately, a total of 28,692 genes were detected as expressed in both cultivars and in both years. In contrast, 2,152 CRIBI V2.1 genes were found not to be expressed.

The correlation between gene expression data, determined as TPM values per sample, of the datasets from both years over all genes was r = 0.5095 (Pearson correlation coefficient) for PN and r = 0.6557 for PNP, respectively. For PN and PNP, 10,205 and 16,226 genes, respectively, expression values were significantly correlated (p-value < 0.05) between the years 2014 and 2017. A list of the correlation strength of the eight time points with the same BBCH stage is provided in Additional file 1: Table S4.

To visualize global trends and similarity of the gene expression values obtained from all subsamples, a Principal Component Analysis (PCA) of both years was performed with keeping the subsamples separate. The first component PC1 explains 57% of the variance, whilst the second component PC2 explains 17% (Figure 2). Almost all data points of the subsamples (triplicates within a sample) from both years cluster near to each other. The data follow a track of time in a nearly consecutive and continuous way. Main actors, which influence most of the variance in the data, were genes related to cell wall modification, secondary metabolism, wounding-response and hormone signaling. The top 500 genes describing most of the variance in PC1 and PC2 are listed in Additional file 1: Table S5 and

185 Table S6, together with functional information for each listed gene from MapMan/Mercator
186 and RefSeq.
187



188

189 **Figure 2**: Principle component analysis of gene expression values from all subsamples. Each
190 data point represents a single subsample of the triplicates for each time point of both years
191 (2014 and 2017 as indicated by [DAF]_14 and [DAF]_17 with the color code) and for both
192 cultivars (PN as triangles, PNP as circles).
193

194 **Cluster analysis for identification of co-expressed genes**

195 The four gene expression time series profiles of all genes (two years, both cultivars,
196 combined data from the subsamples/triplicates for each gene) were compared using the
197 clustering tool clust. The goal was the characterization of the data with respect to similarity
198 and/or differences among years and cultivars throughout berry development. Over all four
199 datasets, 13 PN/PNP clusters of genes with similar gene expression patterns (C1-C13) were
200 obtained (Additional file 2: Figure S1A). In these clusters, 3,316 (12.2 %) of the 27,139 genes
201 expressed during berry development show co-expression among both years and cultivars

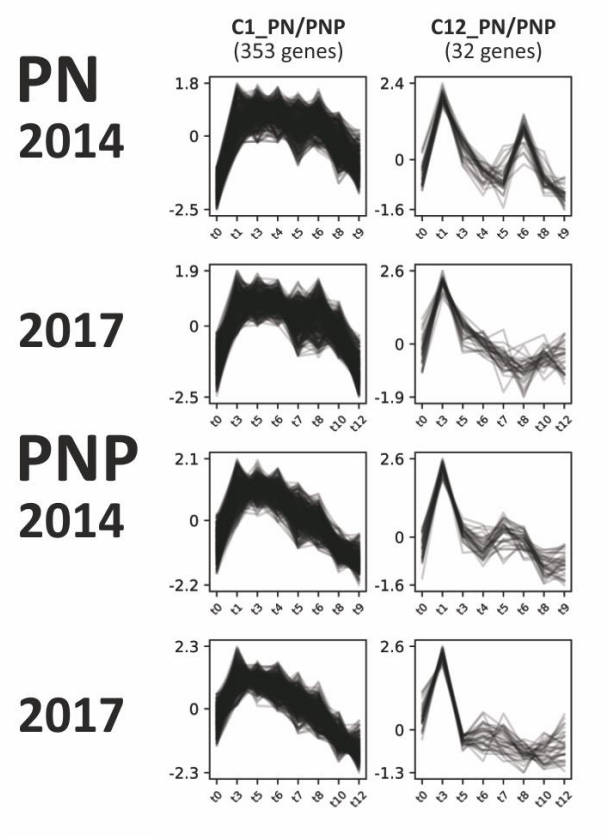
(note that clust uses criteria to define expressed and not expressed genes that differ from the ones applied above, see Methods). The observed expression profiles differ clearly between the clusters, which was in part the result of the restricted the number of clusters that clust extracts. Manual inspection of the clusters revealed little deviation of individual gene expression profiles within each individual cluster for a given year or cultivar. All cluster gene memberships, also those for the additional cultivar-specific cluster analyses (see below), are available in Additional file 1: Table S7-S9.

The PN/PNP clusters C2, C5, C6 and C12 (C12_PN/PNP selected as example, see Figure 3) reveal a small but detectable difference in the gene expression profile between both sampled years, but are almost identical for both cultivars. Thus, the genes in these clusters may display dependence on environmental factors in their expression patterns, potentially due to differences in the weather conditions between the two years studied. The PN/PNP clusters C1, C7 and C11 (C1_PN/PNP selected as example, see Figure 3) show similar expression profiles over the two years, but stand out by shifted expression peaks that distinguish PN and PNP.

To characterize the clusters with respect to potential functions of the co-expressed genes included in a given cluster, GO term enrichment for biological processes was calculated. The full list of enriched GO terms for all clusters is listed in Additional file 1: Table S10-S12. Two examples for GO terms showing up with highly significant incidence were 'response to oxidative stress' in cluster C11_PN/PNP (term GO:0051276) and 'regulation of defense response' in cluster C5_PN/PNP (term GO:0031347).

Two additional cluster analyses were performed, one for the PN data from both years (Additional file 2: Figure S1B) and one for the PNP data from both years (Additional file 2: Figure S1C). These analyses revealed a high abundance of genes from the expansin gene

226 family showing a similar expression profile in clusters C0_PNP and C6_PN. Moreover, the
227 cluster C16_PN showed a highly significant enrichment for 'vegetative to reproductive phase
228 transition' (GO:0010228). Cluster gene memberships for the cultivar-specific clustering are
229 available in Additional file 1: Table S8-9, and the corresponding GO term enrichment is
230 summarized in Additional file 1: Table S11-12.
231



232

233 **Figure 3:** Two selected gene expression profile clusters with either a cultivar-specific
234 difference (C1_PN/PNP) or a weather/field condition-specific difference (C12_PN/PNP) after
235 clustering all data (both cultivars and both years). Strength of gene expression (quantile
236 normalization) was plotted over the time course of berry development. Sampling time points
237 are detailed in Figure 1 and were restricted to those eight equivalent time points at which
238 the cultivars display the same BBCH stage (Additional file 1: Table S4). For all PN/PNP
239 clusters see Additional file 2: Figure S1A. IDs of genes that make up the clusters are listed in
240 Additional file 1: Table S7.
241

Analyses of differentially expressed genes

The gene expression time series throughout berry development were analyzed for differentially expressed genes (DEGs) between the two cultivars PN and PNP with DESeq2. Significantly differentially expressed genes were selected by using the filters adjusted p-value (PADJ) < 0.05 and log2fold change (LFC) > 2. The results are summarized in (Table 2) and are detailed at the gene level per time point compared in Additional file 1: Table S13.

Table 2: Filtering steps applied for selecting DEGs, and the number of DEGs that were carried on after each selection step. For details see Methods.

	PN/PNP 2014 [DEGs]	PN/PNP 2017 [DEGs]
adjusted p-value (PADJ) < 0.05: (counted over all sample pairs)	8,206	4,419
log2fold change (LFC) > 2: (counted over all sample pairs)	6,629	4,298
down- / up-regulated (in PNP vs. PN)**:**	3,293 / 3,336	2,130 / 2,168
unique: (non-redundant within time series)	3,342	2,745
intersection: (detected in both years)	1,923	
excluded due to intersection:	*1,419*	*822*
véraison-specific genes: (detected within BBCH79-81 of PNP)	388	
potentially regulatory: (detected within BBCH61-79 of PNP)	12	

In total, 8,206 and 4,419 DEGs were identified with PADJ better than 0.05 for 2014 and 2017, respectively. Almost twice as many DEGs were initially detected for 2014 compared to 2017. By applying the filter for an at least 2-fold difference in expression level (LFC > 2), the number of significant DEGs decreased, mainly for the PN/PNP time series from 2014. In the

beginning of both time series, that is during flowering (BBCH61 to 69), only few DEGs between PN and PNP were observed (see Figure 4). Within berry formation (BBCH71 to 79), the number of DEGs detected increased towards véraison (BBCH81) as both genotypes increasingly vary for their physiological stage. The highest number of DEGs was observed in parallel to the time-shifted véraison of PNP relative to PN that phenotypically is the most prominent difference between the two cultivars. A set of véraison-specific genes was defined by selecting the DEGs from time points DAF35 and DAF41 from 2014 that show also up at DAF42 and DAF49 from 2017. These criteria identified 388 véraison-specific DEGs. This set of véraison-specific genes was compared to results from similar studies and matches very well (e.g. 81% [25] and 52% [26]; IDs of the 388 genes, the genes that match results from the other studies and their functional annotation are included in Additional file 1: Table S14). During the subsequent phase of berry ripening (BBCH81 to 89), the number of DEGs detected decreases.

We developed a visualization for the numbers of DEGs detected and the changes with respect to which genes are newly appearing as differentially expressed at a given time point (sample pair PN/PNP) in the time series. Figure 4 shows this visualization. Groups of newly appearing DEGs relative to an earlier time point were indicated by a new color shade in the column (bar) for each time point. For members of a given group of DEGs, the attributed color shade was kept for the subsequent columns. This implies that DEGs are initially counted for a pair of PN/PNP samples for each time point individually. If DEGs appearing in several time points are counted only once, 3,342 and 2,745 unique DEGs (different genes) are detected from 2014 and 2017, respectively (compare Table 2).

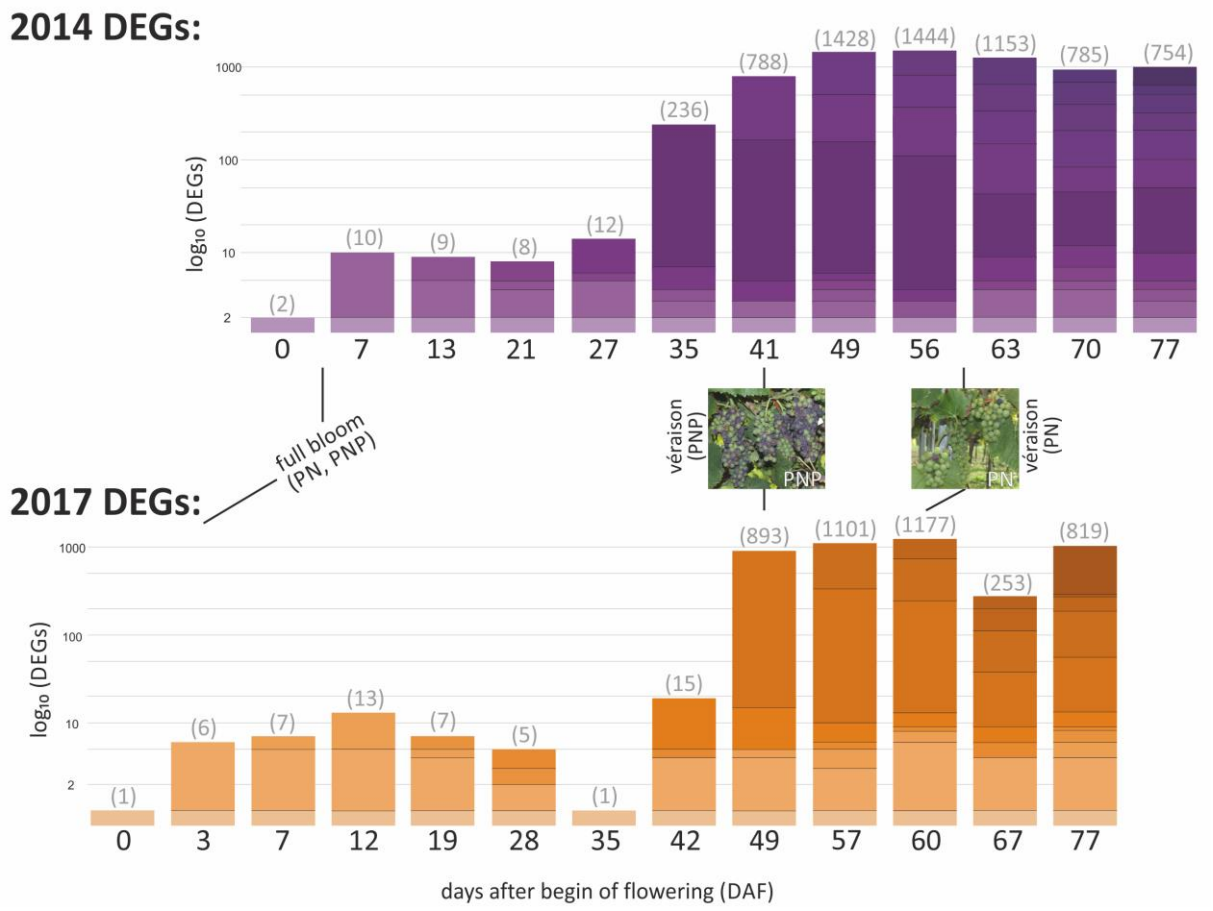


Figure 4: Visualization of the number of DEGs detected between PN and PNP in a logarithmic scale (log_{10}). Results for 2014 are shown in purple, those for 2017 in orange. The time series from the two years were aligned at véraison of PNP; the timeline is given as days after begin of flowering (DAF). The number above each column mentions the number of DEGs detected at the respective time point. Each color shade in each column represents newly appearing DEGs, i.e. those that were additionally detected compared to the earlier sample; the color shade is kept for this set of genes in the subsequent time points (columns/samples).

To further increase the reliability, reproducibility, and relevance of the selected DEGs, the intersection between the DEGs identified in the two years studied was build. In total, 1,923 unique DEGs were obtained (Table 2). To reveal DEGs potentially involved in the regulation of ripening, i.e. genes that might be involved in the trait that mainly distinguishes PN and PNP, only intersecting DEGs which appeared at time points before véraison of PNP were picked. This resulted in a list of 12 putative ripening-regulatory DEGs. It should be noted that

these putative ripening-regulatory DEGs are supposed to be relevant before the set of véraison-specific genes implements the phenotypic changes at véraison. The full list of DEGs, their identity and annotation information as well as their fit to the selection criteria on the way from all (raw) DEGs to ripening-regulatory DEGs is detailed in Additional file 1: Table S13. IDs of the 12 putative ripening-regulatory genes, the genes that match results from related studies (7 DEGs [25], 4 DEGs [33] and 3 DEGs [26]) and their functional annotation are included in Additional file 1: Table S14.

Functional classification of DEGs

To complement the gene lists with functional information from grapevine that might potentially be informative for berry development, the 1,923 intersecting DEGs were analyzed with respect to enrichment of genes that have been assigned to biological pathways already established for grapevine (see Methods). For 46 of the 247 defined grapevine pathways, significant enrichment (permuted p-value <0.1) was detected. The most reliable predictions (permuted p-value <0.001) for pathways that might be relevant were photosynthesis antenna proteins (vv10196; 9 DEGs); nitrogen metabolism (vv10910, 19 DEGs); phenylpropanoid biosynthesis (vv10940, 66 DEGs), tyrosine metabolism (vv10350, 33 DEGs); transport electron carriers (vv50105, 18 DEGs); phenylalanine metabolism (vv10360, 33 DEGs); brassinosteroid biosynthesis (vv10905, 8 DEGs) and flavonoid biosynthesis (vv10941, 30 DEGs). The enrichment results are provided in (Additional file 1: Table S15). The same analysis was also carried out for the 12 putative ripening-regulatory (Additional file 1: Table S16) and the 388 véraison-specific DEGs (Additional file 1: Table S17).

A check of the 1,923 intersecting DEGs revealed that 141 TF genes are included. Of these, 48 DEGs were clearly up- and 93 down-regulated at their first appearance in the time series.

The full list of TF encoding genes that were higher expressed in PNP (up-regulated), or lower expressed in PNP (down-regulated), compared to PN, is shown in Additional file 1: Table S18. For a more detailed view on the expression patterns of selected TF encoding genes, we generated for the TF gene family with the highest abundance among the 141 TF genes, namely the R2R3-MYB-type TFs with 22 cases in the MapMan functional assignment, an expression heatmap (Additional file 2: Figure S2). As a result, *VviMYB24* (VIT_214s0066g01090), which is related to At3g27810/*AtMYB21*, At5g40350/*AtMYB24* and At3g01530/*AtMYB57* according to TAIR/PhyloGenes, was identified as an early appearing DEG that showed its highest expression level at flowering (BBCH61). Prominent *R2R3-MYB* genes known to be relevant for anthocyanin accumulation like *VviMYBA1* (VIT_202s0033g00410), *VviMYBA2* (VIT_202s0033g00390), *VviMYBA3* (VIT_202s0033g00450) and *VviMYBA8* (VIT_202s0033g00380) were detected as expressed starting from véraison (BBCH81) in both cultivars and with a time shift towards earlier expression in PNP. An additional *R2R3-MYB* gene with a similar expression pattern is *VviMYB15* (VIT_205s0049g01020). Other *R2R3-MYB* genes are expressed early during berry formation, these include *VviMYBF1* (VIT_207s0005g01210, related to At2g47460/*AtMYB12*/*AtPFG1*) as well as *VviMYBPA5* (VIT_209s0002g01400) and *VviMYBPA7* (VIT_204s0008g01800, both related to At5g35550/*AtMYB123*/*AtTT2*). According to their related expression patterns visualized in the heatmaps (Additional file 2: Figure S2), the *R2R3-MYB* genes fall into three groups that roughly fit to the three phases marked in Figure 1B, namely flowering, berry formation, and berry ripening (see discussion).

Putative candidates for ripening-regulatory genes

As mentioned above, DEGs detected in both years at time points before véraison of PNP were selected and considered as putative ripening-regulatory genes (Table 2, see full list in Additional file 1: Table S13). The VitisNet enrichment analyses performed for these 12 candidates resulted in 2 pathways that showed significant (permuted p-value < 0.05) enrichment with two genes in the pathway: auxin signaling (vv30003 with *VviEXPA5* (VIT_206s0004g00070) and *VviEXPA14* (VIT_213s0067g02930)) and cell wall (vv40006 with *VviPL1* (VIT_205s0051g00590) and *VviGRIP28* (VIT_216s0022g00960)); see Additional file 1: Table S16).

A detailed check of the data presented in Figure 4, together with results for the putatively ripening-regulatory DEGs, resulted in the identification of two DEGs that stand out from the whole list of DEGs. Both genes almost completely lack expression in the early ripening cultivar PNP while there is clear expression in PN. Therefore, these two genes were detected as DEGs throughout the whole time series in both years. The first of the two, designated *VviRTIC1* for "Ripening TIme Control" (VIT_210s0071g01145, encoding a protein similar to "protein of unknown function DUF789"), is expressed during flowering (BBCH61 - 65) and is more or less continuously down-regulated over time in PN (Figure 5A). The second of the two, designated *VviRTIC2* (VIT_200s0366g00020, encoding a protein similar to "cysteine-rich receptor-like protein kinase"), displays expression in PN during berry formation as well as during berry ripening with a peak before véraison in 2017 (Figure 5B). The expression patterns of the 10 remaining DEGs of the putative ripening-regulatory gene set are shown in Additional file 2: Figure S3.

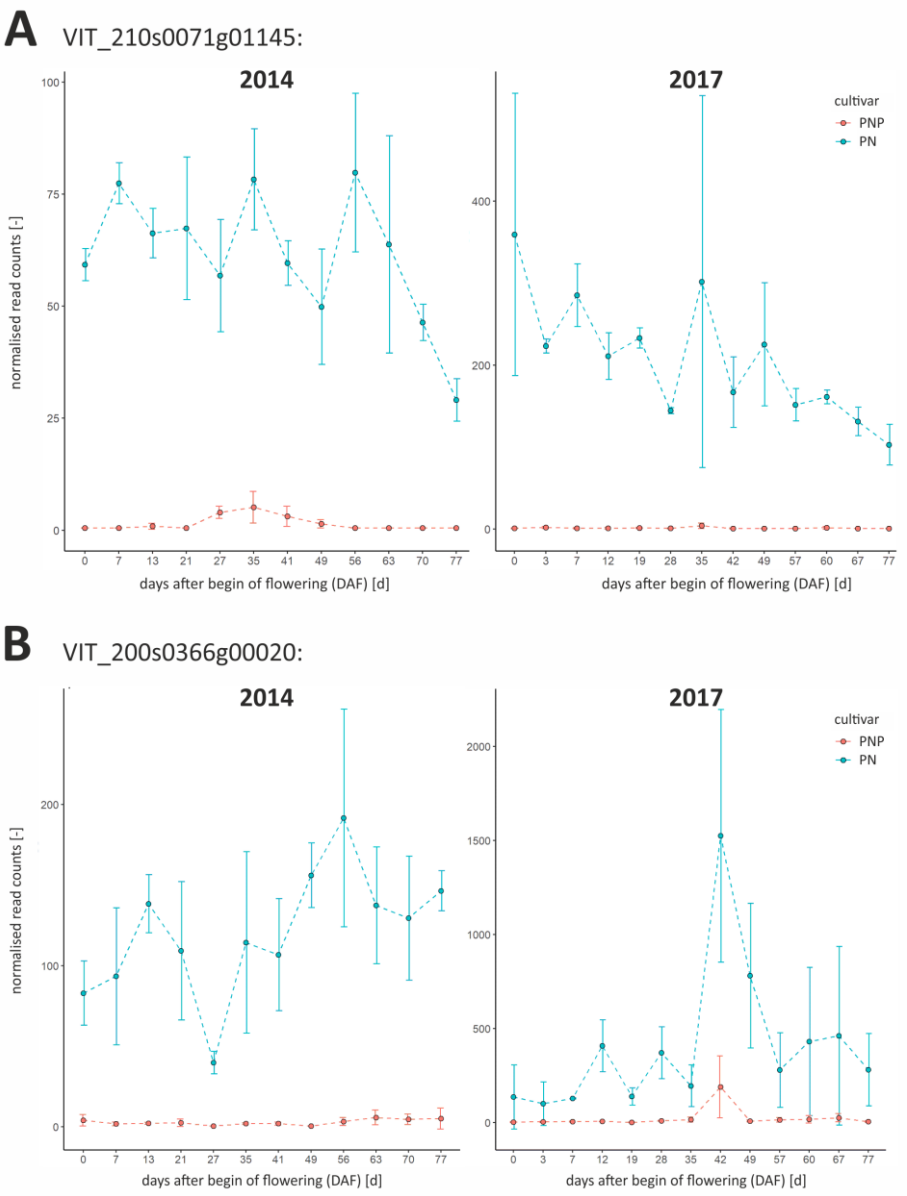


Figure 5: Expression patterns of *VviRTIC1* (VIT_210s0071g01145) in (A) and *VviRTIC2* (VIT_200s0366g00020) in (B) from RNA-Seq data of PN (blue) and PNP (red). Error bars display the standard deviation of triplicates. Left, expression profile from 2014. Right, expression profile from 2017. The y-axis represents the read counts from the output of DESeq2. The x-axis represents the development stages in days after begin of flowering (DAF).

Discussion

One of the first detectable mentions of the cultivar 'Pinot Precoce' in connection with the synonym '(German) Früh Burgunder Traube' (PNP) is in the French book "Ampelographie retrospective" [36]. We have confirmed the clonal relationship of PN and PNP by 24

374 genomically well distributed SSR markers. Although this does not prove that PN is the
375 ancestor, it is very likely that PNP was derived from the cultivar PN by somatic mutation [35].
376 We used these two isogenic cultivars, that are distinguished by a clear phenotype regarding
377 the duration of berry formation, to analyze changes in gene expression throughout berry
378 development. The aim was to identify candidates for genes controlling the speed of berry
379 development and the timing of véraison. Samples from inflorescences as well as from
380 forming and ripening berries were collected from the begin of flowering until after véraison
381 of PN and PNP in 2014 and 2017. These samples were subjected to RNA-Seq analyses in two
382 dense time series.
383
384 **Phenotypic differences between the cultivars PN and PNP**
385 The data from 4 years of careful assessment of the BBCH developmental stages of PN and
386 PNP at the same location validate earlier observations from viticulture [35] that lead to
387 establishment of PNP as a distinct grapevine cultivar in north European wine growing
388 countries. Berry formation lasts about two weeks less in PNP, is clearly accelerated
389 compared to PN and results in PNP entering véraison approximately two weeks earlier than
390 PN (Figure 1A, Table 1). It is reasonable to assume that this acceleration affects berry
391 formation throughout, i.e. from immediately after fruit set until véraison. Functionally, this
392 hypothesis implies that the genes that are responsible for the control of timing of berry
393 development and for the establishment of the phenotypic difference between PN and PNP
394 should be acting already very early in berry development, starting at least shortly after
395 flowering and at or even before BBCH61 to BBCH79. At the end of ripening (harvest), the
396 berries of PNP reach high sugar content earlier within the season when compared to PN.
397

General validation of the RNA-Seq dataset

To estimate overall data quality, the expression profiles obtained from PN and PNP were correlated for the two sampled years 2014 and 2017. Pearson correlation was moderate, but this is expected considering the conditions of the free field environment. Exposure of the vines to external factors like biotic or abiotic stressors, including weather conditions that differ significantly between the years, also affect the transcriptome which reduces the level of correlation [37]. In a PCA, almost all datapoints lie on the same intended track, and biological replicates (subsamples) from both years are located close to each other. The main actors, which predominantly influence the variance in the dataset, are genes related to cell wall modification, secondary metabolism, wounding and hormone signaling. These gene categories fit the expectation since berry development is known to be controlled by hormones, requires new cell walls, and is accompanied by accumulation of secondary/specialized metabolites [22, 24]. These initial results validated the quality of the dataset and indicated clearly that sampling of biologically closely related material for the subsamples/triplicates was successful.

Co-expression analysis shows similar gene expression clusters between cultivars and years

To further validate the data with respect to comparability as well as reproducibility between the two years, related gene expression profiles were identified among all genes by clustering the data from the four different time series. Generally, clusters of the same genes with similar expression patters over time were observed for both cultivars and both years. Also, the cluster analyses for gene expression patterns among the years 2014 & 2017 in only PN and in only PNP, confirmed comparability of the gene expression patterns obtained in these two years. Detailed inspection revealed clusters representing expression profiles (and

clusters of genes) with and without an environmental influence. Especially the cultivar-specific clusters C1_PN/PNP, C7_PN/PNP and C11_PN/PNP stand out. Comparison of the expression profiles for PN to those of PNP in these clusters identified a similar pattern that is moved to a different time in PNP. These findings are congruent with the shifted ripening time phenology of the two cultivars, namely accelerated berry formation in PNP by about two weeks. In contrast, the clusters C2_PN/PNP, C5_PN/PNP, C6_PN/PNP and C12_PN/PNP display more pattern similarity among the two years than among the two cultivars which indicates environmental control of the expression of the genes in these clusters. Strong environmental effects on gene expression patterns have also been described for grapevine berry development at 11 different environments (vineyards) from northern Italy [38]. By approximation, highly similar expression profiles are summarized in the PN/PNP clusters C3, C4, C7, C8, C9 and C10. The genes included in these clusters are probably less affected by environmental factors and/or the genotypic difference between PN and PNP. We conclude that our RNA-Seq results and expression level comparisons between two years are based on valid data.

When the genomic location of the DEGs is analyzed, a genome region on chromosome 16 comes into focus. In this region, 54 of the DEGs from the set of 1,923 intersecting DEGs (Table 2) are located. Of these, 28 encode stilbene synthases [39] that are all up-regulated after véraison of PNP (BBCH83). This genome region fits to a major QTL (Ver1) for "timing of the onset of véraison" on linkage group 16 [40]. It remains to be determined if this aggregation of DEGs is by chance. Potentially, the observation is biased by co-regulation of a large number of closely linked stilbene synthase genes.

Differentially expressed genes throughout berry development and identification of véraison-specific genes

Differential gene expression analysis and subsequent filtering revealed 1,923 DEGs between PNP and PN. DEG detection was based on a comparison of samples taken from the two cultivars at very similar DAF. As expected for the characterized phenotype, PNP reaches véraison when PN is still in the phase of berry formation. Consequently, the strong increase in the number of detected DEGs shortly before and at véraison of PNP results from the different developmental stage of PNP compared to the lagging PN. Subsequently, when also PN enters véraison, the number of DEGs declines (note that Figure 4 uses a logarithmic scale). A list of 388 genes that show up in both years with a véraison-specific expression pattern was extracted and compared to published results. Interestingly, about 81.5% of the 388 PN/PNP véraison-specific genes were also described in the 4,351 differentially expressed genes between the table grape cultivar '8612-CTR' (wild type) and its early ripening bud mutation '8612-ERN' [25]. Also, analyses of berries from the cultivars 'Cabernet Sauvignon' and 'Pinot Noir' by RNA-Seq identified a gene set of 5,404 genes marking the onset of berry ripening [26]. This set covers 51.5% of the 388 PN/PNP véraison-specific genes (Additional file 1: Table S14). Further, several "switch genes" which are supposed to encode key regulators of the developmental transition at véraison [33, 41] are included in the 388 véraison-specific gene set (Additional file 1: Table S14). We conclude that the PN/PNP véraison-specific set of 388 genes represents a core set of genes that are relevant for executing the switch from berry formation to berry ripening. The fact that a relatively small gene set was detected that still displays high overlap to studies that addressed a similar biological question indicates that the specific experimental setup implemented filters acting well against unrelated genes. The comparison of "wildtype to mutant" results in isogenic

469 background, growth in the same vineyard/location reduced environmentally controlled
470 transcriptome differences, and dense sampling together with high RNA-Seq read coverage
471 allowed good resolution power.
472 In order to check for potentially co-expressed genes within the véraison-specific gene set,
473 the memberships for these genes in the PN/PNP cluster analysis were investigated. A total of
474 48 véraison-specific DEGs were detected in cluster C6_PN/PNP (contains 914 genes). These
475 48 genes include several prominent ripening-related genes like *VviGRIP61*
476 (VIT_201s0011g05110), *VviMYBA8*, *VviMRIP1* (VIT_205s0049g00760, [42]), *VviGRIP4*
477 (VIT_205s0049g00520) and *VviGRIP28*. The about 20 *VviGRIP* genes were initially detected
478 by differential cDNA screening as ripening-induced genes in grape [43]. Another relevant
479 cluster is C5_PN/PNP (contains 263 genes) which includes 37 of the 388 véraison-specific
480 DEGs. Among these are *VviMYBA1*, *VviMYBA2*, *VviMYB15* and *VviGRIP22*
481 (VIT_206s0004g02560). The two clusters C5_PN/PNP and C6_PN/PNP show quite similar
482 patterns (Additional file 2: Figure S1A). At first glance it is not obvious which difference has
483 forced clust to put a given gene in either C5_PN/PNP or C6_PN/PNP. A comparison of the
484 expression patterns of *VviMYBA2* (in C5_PN/PNP) and *VviMYBA8* (in C6_PN/PNP; see
485 Additional file 2: Figure S2 for a heatmap) shows that there are borderline cases regarding
486 assignment to either C5_PN/PNP or C6_PN/PNP.
487 In total, 22 genes encoding R2R3-MYB TFs were found among the 1,923 intersecting DEGs.
488 Based on the timing of expression in PN and PNP, the 22 *R2R3-MYB* genes can be classified
489 into three groups (Additional file 2: Figure S2). The first group is represented by *VviMYB24*
490 which is expressed during early flowering (BBCH61) but switched off already at the transition
491 from flowering to berry formation (BBCH71). *VviMYB24* is potentially orthologous to a group
492 of three *A. thaliana R2R3-MYB* genes (*AtMYB21*/*24*/*57*) that are expressed in flowers and

493 function redundantly to regulate stamen development in the context of jasmonate action
494 [44]. It is tempting to speculate that *VviMYB24* has a similar function in grape.
495 The second group covers about 15 *R2R3-MYB* genes that are expressed during berry
496 formation and pre-véraison (BBCH71 to 77). This group includes *VviMYBF1* which regulates
497 flavonol biosynthesis [45], and *VviMYBPA5* as well as *VviMYBPA7* which belong to the clade
498 of *AtTT2*-related genes that control proanthocyanidin (PA, flavan-3-ol) biosynthesis [7, 46,
499 47]. The other *R2R3-MYB* genes in this group are less well characterized although there are
500 functions described for some of them, e.g. *VviMYBC2-L3* (VIT_214s0006g01620) as repressor
501 of specific branches of the phenylpropanoid pathway [48].
502 The third group of *R2R3-MYB* genes is active starting at véraison (after BBCH81) and covers
503 about six genes. Among them are the anthocyanin accumulation controlling genes,
504 *VviMYBA1*, *VviMYBA2, VviMYBA3* and *VviMYBA8* for which there is good evidence that they
505 trigger anthocyanin biosynthesis [49]. Since PN and PNP are red berry cultivars, activity of
506 the TF genes that direct anthocyanin accumulation is expected. In addition, this group
507 includes *VviMYB14* and *VviMYB15* that are supposed to regulate the stilbene biosynthetic
508 pathway [50]. With regard to the heatmaps (Additional file 2: Figure S2) and the analyses of
509 the DEGs in this study in general, it should be noted that while the resolution within the
510 developmental program and time is quite good, our data do not resolve the exact location of
511 gene expression. Therefore, it remains to be determined if the expression detected is
512 derived from berry skin, flesh, the seed or other tissues/cells.
513
514 **Putative ripening-regulatory genes acting early in berry development**
515 To focus on genes that are contributing to the acceleration of berry formation in PNP, and/or
516 to the control of timing of véraison, we selected DEGs detected at time points prior to

véraison of PNP (Table 2, Figure 4). This resulted in a set of only 12 genes that are potentially involved in the regulation of ripening time. According to our hypothesis that the genes relevant for acceleration of berry formation in PNP, which cause the earlier begin of ripening of PNP, should be acting from at least shortly after flowering, we designated this set of genes as "ripening-regulatory". However, genes that encode components of the respective regulatory networks and target genes of regulators including secondarily affected DEGs are surely included as well [51]. The 12 ripening-regulatory genes, i.e. the DEGs detected before véraison of PNP, encode proteins related to auxin action, pectin processing enzymes related to cell wall modification, TFs from the HD-Zip as well as NF-Y/LEC families, a cysteine-rich receptor-like protein kinase, an oleosin, and proteins with domains of unknown function.

The two genes assigned to auxin signaling by VitisNet (vv30003) encode expansins (*VviEXPA5* and *VviEXPA14*, [52]). Expansins are known to be involved in fruit ripening through cell wall expansion and cell enlargement [53]. Auxin can delay the onset of véraison and ripening processes in grapevine [17-19]. Since reduced expression of genes from the auxin signaling pathway may indicate reduced auxin action due to lower auxin levels, the accelerated entry of PNP into véraison might be initiated by reduced auxin levels. Additionally, the genes *VviPL1* (pectate lyase 1 [54]), *VviPME10* (pectin methylesterase 10, VIT_206s0009g02560) *VviGRIP28* (encoding a pectin methylesterase inhibitor precursor-like protein) are also related to cell wall processes, indicating that cell wall modification is an important target process also prior to véraison [53]. The gene *VviGRIP28* was also detected within a véraison-specific meta-QTL designated ver/ph16.1 [28]. It remains to be determined if this correlation has a functional basis.

The two genes in the set of 12 that encode TFs are *VviHDZ28* (VIT_216s0098g01170, [55]) and *VviLEC1* (VIT_200s0956g00020, [56]). The *V. vinifera* gene VIT_216s0098g01170 that has

been designated *VviHDZ28* has also been considered as a homolog of *AtHB12* (At3g61890), but it seems that *VviHDZ07* (VIT_202s0025g02590) and *VviHDZ27* (VIT_215s0048g02870) are more similar to *AtHB12*. In these cases, which lack clearly assignable homologs, transfer of functional information reaches its limits and might be restricted to concluding that *VviHDZ28* is important for organ development in *Vitis*. The gene *VviLEC1* is one of three genes in *V. vinifera* which are homologs of *AtLEC1* (At1g21970, NF-YB9) and *AtL1L* (LEC1-like, At1g21970, NF-YB6). LEC1 and L1L are central regulators of embryo and endosperm development. They are controlling, among other processes, embryo morphogenesis and accumulation of storage reserve [57]. It is tempting to speculate that the reason for the detection of *VviLEC1* among the 12 ripening-regulatory genes is that also seed development needs to be accelerated in PNP compared to PN. This would explain earlier and higher expression of *VviLEC1* in PNP compared to PN as observed (Additional file 2: Figure S3G). Consequently, *VviOLE5* (VIT_216s0013g00880, encoding an oleosin involved in oil body formation [58]) would fit into the picture as relevant for lipid storage during seed development. According to the proposed enzyme function as alcohol acyltransferase of the protein encoded by VIT_209s0018g01490 involved in volatile ester formation [59], this gene could fit in a similar way. For the gene VIT_205s0077g01980 no functional annotation is available (uncharacterized protein), although homologs exist throughout the Magnoliophyta.

Candidates for causal genes explaining the difference between PN and PNP

Among the 12 putative ripening-regulatory genes, of which 10 are discussed above, two are especially interesting. Detailed analyses of the full set of DEGs, visualized in Figure 4, resulted in the identification of *VviRTIC1* and *VviRTIC2*, that could possibly be centrally involved in the accelerated berry development and earlier beginning of ripening in PNP

compared to PN. The special feature of the expression patters of the two genes (Figure 5) is that both are differentially expressed already at the first time point analyzed which was selected to hit the BBCH stage 61 (flowering before full bloom). Also, both genes are only barely expressed in PNP in both years studied, while expression in PN is high at almost all time points. *VviRTIC1* is annotated to encode a protein containing a domain of unknown function (DUF789), while *VviRTIC2* is annotated to encode a "cysteine-rich receptor-like protein kinase". The best blastP hit to *A. thaliana* protein sequences indicates that it is related to At4g23180/AtCRK10, but a closer inspection shows that similarity to At4g05200/AtCRK25, At4g23160/AtCRK8 and At4g23140/AtCRK6 is almost as high. This ambiguity, and also the fact that the *V. vinifera* genome contains several genes related to *VviRTIC2* (e.g. VIT_210s0071g01200, VIT_202s0087g01020 or VIT_203s0017g01550 as listed by PhyloGenes), complicates transfer of functional information. For *AtCRK10*, relevance in cytokinin signaling has been provided by a systematic analysis of the *A. thaliana* CRKs [60]. As pointed out above, it is well possible that the genes we have identified are part of a genetic pathway that realizes control of timing of berry development in *V. vinifera*, but that we have not reached the start of this pathway. The relevance of these two candidate genes in the causal genetic difference between PN and PNP remains to be determined. Phase-separated genome sequences of the cultivars will be required to resolve the genome structure of both alleles of *VviRTIC1 and 2* the genes in PN and PNP for an informative comparison. In future studies, we will address this question, for example by long read DNA sequencing.

Conclusions

588 This study detected 1,923 DEGs between the Pinot cultivars PN and PNP. The two clonal
589 cultivars display a phenotypic difference, namely acceleration of berry development from
590 full bloom to véraison in PNP. We defined 388 DEGs as véraison-specific and 12 DEGs as
591 ripening-regulatory. The relatively small number of véraison-specific genes displays a very
592 high overlap with results published for similar studies (see Additional file 1: Table S14) and
593 could be used for studying a phytohormone network that is similarly realized in PN and PNP,
594 but accelerated by about two weeks in PNP. Additionally, the ripening-regulatory gene set
595 might offer access to a set of genes putatively important for triggering or delaying the start
596 of berry ripening within PNP or PN, respectively. Further investigations are needed on the
597 DNA sequence level to elucidate structural differences in the genomes, the function of the
598 observed DEGs and their role in shifting ripening time in grapevine.
599

Material and Methods

Plant material and analysis of clonal relation

The grapevine (*Vitis vinifera* subsp. *vinifera* L.) cultivar PNP (Pinot Precoce Noir, VIVC No. 9280) is early ripening and has been described to be related to the cultivar PN (Pinot Noir, VIVC No. 9279) [34] that ripens later than PNP. To prove the clonal relation, DNA from both cultivars was genotyped utilizing 24 polymorphic SSR markers (VVS2, VVMD7, VVMD5, VVMD32, VVMD28, VVMD27, VVMD25, VVMD24, VVMD21, VVIV67, VVIV37, VVIQ52, VVIP60, VVIP31, VVIN73, VVIN16, VVIH54, VVIB01, VrZAG83, VrZAG79, VRZAG67, VrZAG62, VMC4F3.1, VMC1B11) as described [61]. The two cultivars used have been identified as accession DEU098_VIVC9280_Pinot_Precoce_Noir_DEU098-2008-076 and DEU098_VIVC9279_Pinot_Noir_DEU098-2008-075, respectively. The tissue used for harvest is indicated below and in Figure 1. Both cultivars do not belong to an endangered species and were obtained and are grown in accordance with German legislation.

Phenotypical characterization and sampling of plant material

Plant material was harvested from PN and PNP grapevines trained in trellis. The plants are growing at the vineyards of JKI Geilweilerhof located at Siebeldingen, Germany (N 49°21.747, E 8°04.678). The grapevine plants were planted with an interrow distance of 2.0 m and spacing of 1.0 m in north-south direction. Inflorescences, developing and ripening berry samples of PNP and PN for RNA extraction were collected in two years with three independent biological replicates (subsamples) each. Sampling took place at systematic time points (12 time points in 2014, 13 time points in 2017), and at approx. 8 a.m. each day. In 2014, harvesting took place regularly every 7 days with only two exceptions (one day deviation, DAF 13 and DAF 27). In 2017, harvesting was adapted to BBCH stages (Figure 1A).

The timeline in both years is described as days after begin of flowering (DAF), with begin of flowering defined as the day at which 10% of the individual flowers have lost their caps [3]. For each subsample within the triplicates, material from two neighboring grapevines was selected. Grapevine plants were weekly phenotyped according to BBCH stage [3, 4]. Phenotyping was performed repeatedly to ensure sampling from vines of the same development stage (e.g. percentage of open flowers during flowering, or berry development stage) to reach uniform subsamples. The phenotypical observations were summarized in Additional file 1 - Table S2 and S3. From these, the durations of flowering, berry formation and berry ripening as well as the resulting shifts between the cultivars were calculated (Table 1). Furthermore, images from berry developmental stages of both cultivars were taken in 2014 for 35, 41, 49 and 56 DAF. The sampled material was directly frozen in liquid nitrogen and stored at -70°C until RNA extraction.

RNA extraction and cDNA library construction

Biological replicates, i.e. the subsamples, were ground separately under liquid nitrogen. Total RNA was extracted using an RNA Isolation Kit (Sigma-Aldrich Spectrum™ Plant Total RNA) according to suppliers' instructions. For RNA-Seq, 500 ng total RNA per subsample were used to prepare sequencing libraries according to the Illumina TruSeq RNA Sample Preparation v2 Guide. For subsamples from 2014 and 2017, 72 and 78 libraries were constructed and sequenced, respectively. Enrichment of poly-A containing mRNA was performed twice, using poly-T oligos attached to magnetic beads included in the Illumina kit. During the second elution of the poly-A+ RNA, the RNA was fragmented and primed for cDNA synthesis. After cDNA synthesis, the fragments were end-repaired and A-tailing was performed. Multiple indexing adapters were ligated to the ends of the cDNA fragments and

the adapter ligated fragments were enriched by 10 cycles of PCR. After quality check and quantification, the libraries were pooled equimolarly.

RNA-Seq

Single end (SE) sequencing of the pooled barcoded libraries from 2014 was performed on an Illumina HiSeq1500 in HighOutput mode generating 100 nt reads. For samples from 2017, sequencing was done using an Illumina NextSeq500 generating 83 nt SE reads; two runs were performed with the same pool of barcoded libraries from 2017.

Processing of RNA-Seq read data

Raw reads were trimmed with Trimmomatic (version 0.36) [62]. For raw reads from the year 2014, the following settings were used: LEADING:10 TRAILING:10 SLIDINGWINDOW:4:15 MINLEN:50. In addition, a collection of all available Illumina adapter sequences was supplied to remove matches within the parameter 2:30:10. For raw reads from the year 2017, trimming settings were set to LEADING:6 TRAILING:6 SLIDING WINDOW:4:15 MINLEN:36. All trimmed reads were quality-checked via FastQC (version 0.11.8) [63]. Thus, possible adapter sequences and low-quality bases were removed. All trimmed reads passing QC were mapped to the reference genome sequence PN40024 (version 12Xv2) [64] using the graph-based alignment tool HISAT2 (version 2.1.0) [65, 66] with no additional soft clipping. Afterwards, all tagged genes (structural gene annotation: CRIBI v2.1) were counted with Featurecounts (Bioconductor package Rsubread version 3.8 [67]). To estimate transcript abundance as a measure for gene expression, counts for Transcripts Per Kilobase Million (TPM, [68]) were determined.

672 **Basic gene expression analyses**

673 TPM counts from the various samples were used for manual gene expression inspection, for
674 determination of the number of expressed and not expressed genes, and to calculate the
675 correlation between gene expression values from both years. Genes with a TPM value > 0
676 added up over all samples from one year were classified as expressed, conversely genes with
677 a TPM value = 0 added up over all samples as not expressed. A custom python script was
678 applied utilizing the function pearsonr from SciPy python package (v. 1.2.3) [69], which
679 calculates the Pearson correlation coefficient and the p-value for all year-to-year
680 comparisons. Expression data pairs for TPM counts per gene from both sampled years,
681 averaged over the three subsamples of each sample, were used. To test correlations and
682 relationships between expression values from the two years, where samples were harvested
683 with slightly different sampling patterns (see Figure 1A), eight equivalent time points with
684 the same BBCH stages between the cultivars of each year were selected (see Additional file
685 1: Table S4).

686

687 **Principal component analysis**

688 To explore data similarity, a Principal Component Analysis (PCA) was calculated over all gene
689 expression values from both years and cultivars for all subsamples. All data points were
690 normalized using variance stabilizing transformation function 'vst' from the R package
691 DESeq2 (v. 1.12.4) [70]. Subsequently, the principal components were generated using
692 'prcomp' from the R package 'stats' (v. 3.5.2) [71]. The resulting PCA object, displaying the
693 main components PC1 and PC2, was plotted and exported. Additionally, genes with the
694 highest variance contribution to PC1 and PC2 were extracted separately.

695

696 **Functional annotation of genes**

697 Transfer of annotation information from other plant species, mainly *A. thaliana,* was
698 calculated using MapMan's sequence annotation tool Mercator (v. 3.6) [72, 73]. Additionally,
699 all open reading frame (CDS from *V. vinifera*/grapevine genes) sequences were aligned to
700 the non-redundant protein sequence data base RefSeq [74] with the basic local alignment
701 tool for proteins BLASTp [75] (e-value ≤ 0.001). Short descriptions of gene functions were
702 extracted and added to the gene lists in Additional file 1: Table S5, S6, S13, S14.
703 GO term enrichment for biological processes was calculated via the R package 'topGO' (v.
704 2.38.1) [76]. Subsequently, statistical reliability was calculated using Fishers exact test. All
705 Gene IDs and their corresponding GO terms were extracted from the CRIBI database
706 (http://genomes.cribi.unipd.it/DATA/V2/annotation/bl2go.annot_with_GO_description.txt).
707 All results of the GO term enrichment are deposited in Additional file 1 - Table S10-12.
708

709 **Cluster analysis**

710 To reveal co-expressed genes over all four datasets, the tool clust (v. 1.10.8) was used with
711 default parameters [77]. As input, raw read counts from eight time points were used. These
712 time points were selected to cover the same BBCH stages of PN and PNP from the years
713 2014 and 2017 (Additional file 1: Table S4). First, all data were pre-processed as described in
714 the clust manual. Values from corresponding subsamples (triplicates) were combined and
715 averaged. To filter out uninformative (very low) gene expression values, an additional filter
716 was applied: genes not reaching a sample expression value > 1 in at least three conditions
717 and in at least one cultivar from one year were discarded (-fil-v 1 -fil-c 3 -fil-d 1). Afterwards,
718 the data were quantile normalized according to the RNA-Seq defaults of clust. Genes
719 showing a flat expression profile were filtered out by applying the default settings [77].

720

721 **Differential gene expression analyses**

722 For analyses of differentially expressed genes, DESeq2 (v. 1.12.4; R Bioconductor) was
723 employed. To test if gene expression differs significantly between two samples, the
724 likelihood ratio test nbinomLRT, included in the DESeq2 package, was used. Normalization
725 factors and dispersion estimates were used as described [70]. The output table contained all
726 differentially expressed genes (DEGs) and the corresponding values for baseMean,
727 log2FoldChange (LFC), lfcSE (LFC standard error), stat (difference in deviation between the
728 reduced model and the full model), p-value and PADJ (adjusted p-value). To focus on
729 significantly differentially expressed genes from the DESeq2 analyses, cut-off filters PADJ ≤
730 0.05 and LFC > 2 were applied.

731

732 **Selection of gene sets potentially relevant for ripening and comparison with literature data**

733 In order to identify gene sets from the DEGs relevant for control and realisation of ripening,
734 an intersection between the DEGs detected at all time points between both years was built.
735 To determine a subset of putatively ripening-regulatory genes, the intersection between
736 both years covering the development stages BBCH61 (begin of flowering) to BBCH79 (one
737 developmental BBCH stage before véraison) was used (time points 2014: DAF0-35; 2017:
738 DAF0-42). Furthermore, a set of véraison specific genes was defined from the the DEGs
739 detected at the intersection of development stages BBCH79 (one developmental BBCH stage
740 before véraison) to BBCH81 (begin of ripening / véraison; time points 2014: DAF35-41; 2017:
741 DAF42-49). To test for biological relevance of the subsets, all DEGs were screened to their
742 occurrence in similar relevant studies [25, 26, 28, 29, 33, 41, 78, 79].

743

744 **Visualization of gene numbers newly appearing as differentially expressed**

745 To visualize appearance of DEGs over time, a stacked bar plot script was set up using the R
746 package 'plotly' (v. 4.9.2.1) [80]. Each bar represents the amount of DEGs of a given time
747 point or condition. In order to track groups of DEGs newly appearing at a given time point
748 throughout the following time points, the colorshade representing the group of DEGs
749 remains the same.

750

751 **Pathway enrichment analysis**

752 To search for possible targets in known pathways of grapevine, a pathway enrichment
753 analysis using the tool VitisPathways [81] was performed. To achieve a reliable enrichment,
754 1000 permutations, a Fisher's exact test of $p < 0.05$ and permuted p-value < 0.1 were set.
755 Thus, all significant enriched pathway genes and their relations can be displayed in VitisNet
756 [82], a specific molecular network for grapevine.

757

758 **Heatmaps**

759 As an extension to assignment of genes to biosynthesis pathways, the genes were also
760 filtered for annotation as coding for transcription factors (TFs). This filter was based on the
761 annotation information transferred from Mercator and RefSeq (see above). To look at the
762 entire family of *R2R3-MYB* TF genes, the list of *MYB* genes identified via MapMan was
763 extended by additional grapevine *R2R3-MYB* gene family members that have been
764 characterized [6, 7]. The *R2R3-MYB* genes detected among the intersecting DEGs were
765 displayed in heatmaps addressing the four individual time series (2 cultivars, 2 years) using
766 the R package 'pheatmap' (v. 2.1.3) [83]. Predictions for phylogenetic relationships were
767 deduced from PhyloGenes v. 2.2 [84].

768

769 **Abbreviations**

770 ABA abscisic acid

771 DAF days after begin of flowering

772 DEGs differentially expressed genes

773 PNP Pinot Noir Precoce

774 PN Pinot Noir

775 LFC log2fold change

776 PADJ adjusted p-value

777 TF transcription factor

778 TPM transcripts per kilobase million

779

780 **Declarations**

781 **Ethics approval and consent to participate**

782 Not applicable

783

784 **Consent for publication**

785 Not applicable

786

787 **Availability of data and material**

788 The FASTQ files containing all RNA-Seq reads (PN2014, PNP2014, PN2017 and PNP2017)

789 have been deposited at the European Nucleotide Archive (ENA) according to the INTEGRAPE

790 guidelines under the accession numbers PRJEB39262, PRJEB39261, PRJEB39264 and

791 PRJEB39263, respectively (see Additional file 1: Tables S19 to S22). All scripts developed for

792 this study are available on GitHub [https://github.com/bpucker;
793 https://github.com/jenthein].
794
795 **Competing interests**
796 The authors declare that they have no competing interests
797
798 **Funding**
799 This project was funded by BMBF/PtJ (FKZ 031A3496).
800
801 **Authors' contributions**
802 Sampling and phenotyping were done at the JKI (Julius Kühn-Institute, Institute for
803 Grapevine Breeding Geilweilerhof, Siebeldingen, Germany). Sequencing and data analysis
804 were performed at Bielefeld University, Faculty of Biology & Center for Biotechnology
805 (CeBiTec). Organization, supervision at Bielefeld University and parts of sampling were done
806 by DH. KH, AK and FS coordinated weekly sampling, phenotyping, and image capture for
807 documentation. Logistic work was done by KH. The design of the experiments was set up by
808 FS, DH and KH. RNA isolation and cDNA synthesis were carried out by DH and PV. PV
809 accomplished library preparation and sequencing. RT and KH supervised the work at JKI
810 Geilweilerhof. LH performed the SSR analysis. BW supervised the work at Bielefeld
811 University. FS, KH, RT and BW acquired project funding and wrote the project proposal. All
812 bioinformatic data analyses, creation of figures, tables and writing of the manuscript were
813 performed by JT with the help of DH. JT and BW drafted the manuscript. All authors have
814 read and approved the final manuscript.
815

Acknowledgements

We would like to thank Boas Pucker for providing several python scripts for handling large data sets as well as helpful comments on the manuscript, Katharina Frey-Sielemann for help visualizing the DEG distributions and Wiebke Halpape for providing the GO annotation enquiry. We also thank Willy Keller for the excellent technical support. We are very grateful to the Bioinformatics Resource Facility support team of the CeBiTec for providing computing infrastructure and excellent technical support. This publication is based in part upon contributions from COST Action CA 17111 INTEGRAPE, supported by COST (European Cooperation in Science and Technology); the authors thank BMBF/PtJ for funding in the context of IPAS (project acronym novisys). The authors from UniBi also wish to thank the other members of the Chair of Genetics and Genomics of Plants for their support. We would also like to thank the Open Access Publication Funds of Bielefeld University for funding the publication charge.

Supplementary material

Additional file 1: Table S1: Clonal relation of PN and PNP confirmed by 24 SSR markers

Additional file 1: Table S2 - Sampling and phenotypical observations (BBCH) for 2014

Additional file 1: Table S3 - Sampling and phenotypical observations (BBCH) for 2017

Additional file 1: Table S4 - Pearsonr correlations of gene expression values over time between both years

Additional file 1: Table S5 - Top 500 genes influencing principal component PC1

Additional file 1: Table S6 - Top 500 genes influencing principal component PC2

Additional file 1: Table S7 - Cluster memberships of clusters obtained from the clusteranalysis PN/PNP

Additional file 1: Table S8 - Cluster memberships of clusters obtained from the cluster analysis PN

Additional file 1: Table S9 - Cluster memberships of clusters obtained from the cluster analysis PNP

Additional file 1: Table S10 - GO term enrichment for the clusters from the cluster analysis of PN/PNP

Additional file 1: Table S11 - GO term enrichment for the clusters from the cluster analysis of PN

Additional file 1: Table S12 - GO term enrichment for the clusters from the cluster analysis of PNP

Additional file 1: Table S13 - Differentially expressed genes between PN and PNP in both years and filtering steps

Additional file 1: Table S14 - Overlap of 12 ripening regulatory DEGs (time points: DAF0-35 from 2014 intersecting with DAF0-42 from 2017) and 388 véraison-specific DEGs (time points: DAF 35+41 from 2014 intersecting with DAF42+49 from 2017) with previous transcriptomic grapevine studies.

Additional file 1: Table S15 - VitisNet enrichment for intersecting DEGs

Additional file 1: Table S16 - VitisNet enrichment for regulatory DEGs

Additional file 1: Table S17 - VitisNet enrichment for véraison specific DEGs

Additional file 1: Table S18 - Up- & Down-regulated TFs among the DEGs

Additional file 1: Table S19 - ENA sample identifier and metadata linked to the study for PNP RNA-Seq reads from 2014 (PRJEB39262)

Additional file 1: Table S20 - ENA sample identifier and metadata linked to the study for PN RNA-Seq reads from 2014 (PRJEB39261)

Additional file 1: Table S21 - ENA sample identifier and metadata linked to the study for PNP RNA-Seq reads from 2017 (PRJEB39264)

Additional file 1: Table S22 - ENA sample identifier and metadata linked to the study for PN RNA-Seq reads from 2017 (PRJEB39263)

Additional file 2: Figure S1 - Cluster analyses of PN/PNP in both years (A), PN in both years (B) and PNP in both years (C)

Additional file 2: Figure S2 - Heatmap of all MYB-TFs gene expression among the DEGs for PN and PNP in both years

873 **Additional file 2: Figure S3** - Expression patterns of genes from the ripening-regulatory gene

874 set

875

876

877 **References**

878

879 1. Kuhn N, Guan L, Dai ZW, Wu BH, Lauvergeat V, Gomès E, Li SH, Godoy F, Arce-Johnson P,
880 Delrot S: **Berry ripening: recently heard through the grapevine.** *Journal of Experimental*
881 *Botany* 2014, **65**(16):4543-4559.

882 2. Coombe BG: **Research on development and ripening of the grape berry**. *American Journal of*
883 *Enology and Viticulture* 1992, **43**(1):101-110.

884 3. Lorenz DH, Eichhorn KW, Bleiholder H, Klose R, Meier U, Weber E: **Phenological growth stages**
885 **of the grapevine (Vitis vinifera L. ssp. vinifera) - Codes and descrptions according to the**
886 **extended BBCH scale**. *Australian Journal of Grape and Wine Research* 1995, **1**(2):100-103.

887 4. Bleiholder H, Weber E, Feller C, Hess M, Wicke H, Meier U, van den Boom T, Lancashire PD,
888 Buhr L, Hack H *et al*: **Growth stages of mono- and dicotyledonous plants - BBCH Monograph**:
889 Federal Biological Research Centre for Agriculture and Forestry; 2001.

890 5. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L: **MYB transcription factors**
891 **in Arabidopsis**. *Trends in Plant Science* 2010, **15**(10):573-581.

892 6. Matus JT, Aquea F, Arce-Johnson P: **Analysis of the grape MYB R2R3 subfamily reveals**
893 **expanded wine quality-related clades and conserved gene structure organization across Vitis**
894 **and Arabidopsis genomes.** *BMC Plant Biology* 2008, **8**:83.

895 7. Wong DCJ, Schlechter R, Vannozzi A, Holl J, Hmmam I, Bogs J, Tornielli GB, Castellarin SD,
896 Matus JT: **A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family**
897 **uncovers new insights into the regulation of stilbene accumulation**. *DNA Research* 2016,
898 **23**(5):451-466.

899 8. Delrot S, Grimplet J, Carbonell-Bejerano P, Schwandner A, Bert P-F, Bavaresco L, Costa LD, Di
900 Gaspero G, Duchêne E, Hausmann L *et al*: **Genetic and Genomic Approaches for Adaptation of**
901 **Grapevine to Climate Change**. In: *Genomic Designing of Climate-Smart Fruit Crops.* Edited by
902 Kole C. Cham (Switzerland): Springer; 2020: 157-270.

903 9. Fenn MA, Giovannoni JJ: **Phytohormones in Fruit Development and Maturation**. *The Plant*
904 *Journal* 2020.

905 10. Osorio S, Scossa F, Fernie AR: **Molecular regulation of fruit ripening**. *Frontiers in Plant Science*
906 2013, **4**:198.

907 11. Sun L, Zhang M, Ren J, Qi J, Zhang G, Leng P: **Reciprocity between abscisic acid and ethylene**
908 **at the onset of berry ripening and after harvest.** *BMC Plant Biology* 2010, **10**:257.

909 12. Pilati S, Bagagli G, Sonego P, Moretto M, Brazzale D, Castorina G, Simoni L, Tonelli C, Guella G,
910 Engelen K *et al*: **Abscisic Acid Is a Major Regulator of Grape Berry Ripening Onset: New**
911 **Insights into ABA Signaling Network.** *Frontiers in Plant Science* 2017, **8**:1093.

912 13. Zhu Z, Li G, Yan C, Liu L, Zhang Q, Han Z, Li B: **DRL1, Encoding A NAC Transcription Factor, Is**
913 **Involved in Leaf Senescence in Grapevine.** *International Journal of Molecular Sciences* 2019,
914 **20**(11):2678.

14. Rossdeutsch L, Edwards E, Cookson SJ, ., Barrieu F, Gambetta GA, Delrot S, Ollat N: **ABA-mediated responses to water deficit separate grapevine genotypes by their genetic background.** *BMC Plant Biology* 2016, **16**(91):s12870-12016-10778-12874.

15. Coelho J, Almeida-Trapp M, Pimentel D, Soares F, Reis P, Rego C, Mithöfer A, Fortes AM: **The study of hormonal metabolism of Trincadeira and Syrah cultivars indicates new roles of salicylic acid, jasmonates, ABA and IAA during grape ripening and upon infection with Botrytis cinerea.** *Plant Science* 2019, **283**:266-277.

16. Chervin C, El-Kereamy A, Roustan J-P, Latché A, Lamon J, Bouzayen M: **Ethylene seems required for the berry development and ripening in grape, a non-climacteric friut**. *Plant Science* 2004, **167**(6):1301-1305.

17. Böttcher C, Burbidge CA, Boss PK, Davies C: **Interactions between ethylene and auxin are crucial to the control of grape (Vitis vinifera L.) berry ripening**. *BMC Plant Biology* 2013, **13**:222.

18. Böttcher C, Harvey K, Forde CG, Boss PK, Davies C: **Auxin treatment of pre-veraison grape (Vitis vinifera L.) berries both delays ripening and increases the synchronicity of sugar accumulation**. *Australian Journal of Grape and Wine Research* 2011, **17**(1):1-8.

19. Ziliotto F, Corso M, Rizzini FM, Rasori A, Botton A, Bonghi C: **Grape berry ripening delay induced by a pre-véraison NAA treatment is paralleled by a shift in the expression pattern of auxin- and ethylene-related genes.** *BMC Plant Biology* 2012, **12**:185.

20. Seymour GB, Ostergaard L, Chapman NH, Knapp S, Martin C: **Fruit Development and Ripening.** *Annual Review of Plant Biology* 2013, **64**:219-241.

21. Giovannoni J, Nguyen C, Ampofo B, Zhong S, Fei Z: **The Epigenome and Transcriptional Dynamics of Fruit Ripening**. *Annual Review of Plant Biology* 2017, **68**:61-84.

22. Karlova R, Chapman N, David K, Angenent GC, Seymour GB, de Maagd RA: **Transcriptional control of fleshy fruit development and ripening**. *J Exp Bot* 2014, **65**(16):4527-4541.

23. Wang R, Tavano E, Lammers M, Martinelli AP, Angenent GC, de Maagd RA: **Re-evaluation of transcription factor function in tomato fruit development and ripening with CRISPR/Cas9-mutagenesis**. *Scientic Reports* 2019, **9**(1):1696.

24. Serrano A, Espinoza C, Armijo G, Inostroza-Blancheteau C, Poblete E, Meyer-Regueiro C, Arce A, Parada F, Santibanez C, Arce-Johnson P: **Omics Approaches for Understanding Grapevine Berry Development: Regulatory Networks Associated with Endogenous Processes and Environmental Responses**. *Frontiers in Plant Science* 2017, **8**:1486.

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

26. Fasoli M, Richter CL, Zenoni S, Bertini E, Vitulo N, Dal Santo S, Dokoozlian N, Pezzotti M, Tornielli GB: **Timing and Order of the Molecular Events Marking the Onset of Berry Ripening in Grapevine**. *Plant Physiology* 2018, **178**(3):1187-1206.

27. Sun R, He F, Lan Y, Xing R, Liu R, Pan Q, Wang J, Duan C: **Transcriptome comparison of Cabernet Sauvignon grape berries from two regions with distinct climate.** *Journal of Plant Physiology* 2015, **178**:43-54.

28. Delfino P, Zenoni S, Imanifard Z, Tornielli GB, Bellin D: **Selection of candidate genes controlling veraison time in grapevine through integration of meta-QTL and transcriptomic data.** *BMC Genomics* 2019, **20**(1):739.

29. Zenoni S, Ferrarini A, Giacomelli E, Xumerle L, Fasoli M, Malerba G, Bellin D, Pezzotti M, Delledonne M: **Characterization of transcriptional complexity during berry development in Vitis vinifera using RNA-Seq**. *Plant Physiology* 2010, **152**(4):1787-1795.

30. Xu L, Yue Q, Xiang G, Bian F, Yao Y: **Melatonin promotes ripening of grape berry via increasing the levels of ABA, H2O2, and particularly ethylene**. *Horticulture Research* 2018, **5**:41.

31. Griesser M, Martinez SC, Eitle MW, Warth B, Andre CM, Schuhmacher R, Forneck A: **The ripening disorder berry shrivel affects anthocyanin biosynthesis and sugar metabolism in Zweigelt grape berries**. *Planta* 2018, **247**(2):471-481.

32. Zarrouk O, Brunetti C, Egipto R, Pinheiro C, Genebra T, Gori A, Lopes CM, Tattini M, Chaves MM: **Grape Ripening Is Regulated by Deficit Irrigation/Elevated Temperatures According to Cluster Position in the Canopy**. *Frontiers in Plant Science* 2016, **7**:1640.

33. Massonnet M, Fasoli M, Tornielli GB, Altieri M, Sandri M, Zuccolotto P, Paci P, Gardiman M, Zenoni S, Pezzotti M: **Ripening Transcriptomic Program in Red and White Grapevine Varieties Correlates with Berry Skin Anthocyanin Accumulation.** *Plant Physiology* 2017, **174**(4):2376-2396.

34. Maul E, Sudharma KN, Ganesch A, Brühl U, Hundemer M, Kecke S, Mahler-Ries A, Marx G, Schreiber T, Walk M *et al*: **30 Years VIVC - Vitis International Variety Catalogue (www.vivc.de)**. In: *XI International Conference on Grapevine Breeding and Genetics; Yanqing, Beijing, China*. 2014.

35. Robinson J, Harding J, Vouillamoz J: **Wine Grapes: A complete guide to 1,368 vine varieties, including their origins and flavours**: Penguin Books; 2012.

36. Roy-Chevrier J: **Ampelographie retrospective**. Paris; 1900.

37. Pucker B, Schwandner A, Becker S, Hausmann L, Viehover P, Topfer R, Weisshaar B, Holtgrawe D: **RNA-Seq Time Series of Vitis vinifera Bud Development Reveals Correlation of Expression Patterns with the Local Temperature Profile**. *Plants* 2020, **9**(11):1548.

38. Dal Santo S, Tornielli GB, Zenoni S, Fasoli M, Farina L, Anesi A, Guzzo F, Delledonne M, Pezzotti M: **The plasticity of the grapevine berry transcriptome.** *Genome Biology* 2013, **14**(6):R54.

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

40. Zyprian E, Ochßner I, Schwander F, Šimon S, Hausmann L, Bonow-Rex M, Moreno-Sanz P, Grando MS, Wiedemann-Merdinoglu S, Merdinoglu D *et al*: **Quantitative trait loci affecting pathogen resistance and ripening of grapevines.** *Molecular Genetics and Genomics* 2016, **291**(4):1573-1594.

41. Palumbo MC, Zenoni S, Fasoli M, Massonnet M, Farina L, Castiglione F, Pezzotti M, Paci P: **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.** *The Plant Cell* 2014, **26**(12):4617-4635.

42. Burger AL, Zwiegelaar JP, Botha FC: **Characterisation of the gene encoding the Merlot ripening-induced protein 1 (mrip1): evidence that this putative protein is a distinct member of the plant proline-rich protein family**. *Plant Science* 2004, **167**(5):1075-1089.

43. Davies C, Robinson SP: **Differential screening indicates a dramatic change in mRNA profiles during grape berry ripening. Cloning and characterization of cDNAs encoding putative cell wall and stress response proteins**. *Plant Physiology* 2000, **122**(3):803-812.

44. Qi T, Huang H, Song S, Xie D: **Regulation of Jasmonate-Mediated Stamen Development and Seed Production by a bHLH-MYB Complex in Arabidopsis.** *The Plant Cell* 2015, **27**(6):1620-1633.

45. Czemmel S, Stracke R, Weisshaar B, Cordon N, Harris NN, Walker AR, Robinson SP, Bogs J: **The grapevine R2R3-MYB transcription factor VvMYBF1 regulates flavonol synthesis in developing grape berries.** *Plant Physiology* 2009, **151**(3):1513-1530.

46. Bogs J, Jaffe FW, Takos AM, Walker AR, Robinson SP: **The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development**. *Plant Physiology* 2007, **143**(3):1347-1361.

47. Koyama K, Numata M, Nakajima I, Goto-Yamamoto N, Matsumura H, Tanaka N: **Functional characterization of a new grapevine MYB transcription factor and regulation of proanthocyanidin biosynthesis in grapes**. *Journal of Experimental Botany* 2014, **65**(15):4433-4449.

48. Cavallini E, Matus JT, Finezzo L, Zenoni S, Loyola R, Guzzo F, Schlechter R, Ageorges A, Arce-Johnson P, Tornielli GB: **The phenylpropanoid pathway is controlled at different branches by a set of R2R3-MYB C2 repressors in grapevine.** *Plant Physiology* 2015, **167**(4):1448-1470.

49. Matus JT, Cavallini E, Loyola R, Holl J, Finezzo L, Dal Santo S, Vialet S, Commisso M, Roman F, Schubert A *et al*: **A group of grapevine MYBA transcription factors located in chromosome 14 control anthocyanin synthesis in vegetative organs with different specificities compared with the berry color locus**. *The Plant Journal* 2017, **91**(2):220-236.

50. Vannozzi A, Wong DCJ, Holl J, Hmmam I, Matus JT, Bogs J, Ziegler T, Dry I, Barcaccia G, Lucchin M: **Combinatorial Regulation of Stilbene Synthase Genes by WRKY and MYB Transcription Factors in Grapevine (Vitis vinifera L.)**. *Plant and Cell Physiology* 2018, **59**(5):1043-1059.

51. Forlani S, Masiero S, Mizzotti C: **Fruit ripening: the role of hormones, cell wall modifications, and their relationship with pathogens**. *Journal of Experimental Botany* 2019, **70**(11):2993-3006.

52. Dal Santo S, Vannozzi A, Tornielli GB, Fasoli M, Venturini L, Pezzotti M, Zenoni S: **Genome-wide analysis of the expansin gene superfamily reveals grapevine-specific structural and functional characteristics**. *PLoS One* 2013, **8**(4):e62206.

53. Du M, Spalding EP, Gray WM: **Rapid Auxin-Mediated Cell Expansion**. *Annual Review of Plant Biology* 2020, **71**:379-402.

54. Guillaumie S, Fouquet R, Kappel C, Camps C, Terrier N, Moncomble D, Dunlevy JD, Davies C, Boss PK, Delrot S: **Transcriptional analysis of late ripening stages of grapevine berry**. *BMC Plant Biology* 2011, **11**:165.

55. Li Z, Zhang C, Guo Y, Niu W, Wang Y, Xu Y: **Evolution and expression analysis reveal the potential role of the HD-Zip gene family in regulation of embryo abortion in grapes (Vitis vinifera L.)**. *BMC Genomics* 2017, **18**(1):744.

56. Cui M, Wang W, Guo F, Fan X, Guan L, Zheng T, Zhu X, Jia H, Fang J, Wang C *et al*: **Characterization and temporal–spatial expression analysis of LEC1 gene in the development of seedless berries in grape induced by gibberellin**. *Plant Growth Regulation* 2020, **90**(3):585-596.

57. Jo L, Pelletier JM, Harada JJ: **Central role of the LEAFY COTYLEDON1 transcription factor in seed development**. *Journal of Integrative Plant Biology* 2019, **61**(5):564-580.

58. Liu Q, Sun Y, Su W, Yang J, Liu X, Wang Y, Wang F, Li H, Li X: **Species-specific size expansion and molecular evolution of the oleosins in angiosperms**. *Gene* 2012, **509**(2):247-257.

59. Qian X, Liu Y, Zhang G, Yan A, Wang H, Wang X, Pan Q, Xu H, Sun L, Zhu B: **Alcohol acyltransferase gene and ester precursors differentiate composition of volatile esters in three interspecific hybrids of Vitis labruscaxV. Vinifera during berry development period**. *Food Chemistry* 2019, **295**:234-246.

60. Bourdais G, Burdiak P, Gauthier A, Nitsch L, Salojarvi J, Rayapuram C, Idanheimo N, Hunter K, Kimura S, Merilo E *et al*: **Large-Scale Phenomics Identifies Primary and Fine-Tuning Roles for CRKs in Responses Related to Oxidative Stress**. *PLoS Genetics* 2015, **11**(7):e1005373.

61. Dallakyan M, Zyprian E, Maul E, Yesayan A, Hovhannisyan N: **Genetic characterization of rape varieties in Armenia**. *Vitis* 2015, **54**:23-26.

62. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data.** *Bioinformatics* 2014, **30**(15):2114-2120.

63. **FastQC, a quality control tool for high throughput sequence data.** [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]

64. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C *et al*: **The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla**. *Nature* 2007, **449**(7161):463-467.

65. Kim D, Langmead B, Salzberg SL: **HISAT: a fast spliced aligner with low memory requirements**. *Nature Methods* 2015, **12**(4):357-360.

66. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL: **Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype**. *Nature Biotechnology* 2019, **37**(8):907-915.

67. Liao Y, Smyth GK, Shi W: **featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.** *Bioinformatics* 2014, **30**(7):923-930.

68. Zhao S, Ye Z, Stanton R: **Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols**. *RNA* 2020, **26**(8):903-909.

69. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J *et al*: **SciPy 1.0: fundamental algorithms for scientific computing in Python**. *Nature Methods* 2020, **17**(3):261-272.

70. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome Biology* 2014, **15**(12):550.

71. **R: A language and environment for statistical computing** [https://www.R-project.org/]

72. Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M: **MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes.** *The Plant Journal* 2004, **37**(6):914-939.

73. Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie AR, Stitt M, Usadel B: **Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data**. *Plant, Cell & Environment* 2014, **37**(5):1250-1258.

74. Pruitt KD, Tatusova T, Maglott DR: **NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins.** *Nucleic Acids Research* 2007, **35**(Database issue):D61-65.

75. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *Journal of Molecular Biology* 1990, **215**(3):403-410.

76. Alexa A, Rahnenfuhrer J, Lengauer T: **Improved scoring of functional groups from gene expression data by decorrelating GO graph structure**. *Bioinformatics* 2006, **22**(13):1600-1607.

77. Abu-Jamous B, Kelly S: **Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data.** *Genome Biology* 2018, **19**:172.

78. Zenoni S, D'Incà E, Tornielli GB: **Genetic dissection of grape berry ripening control: defining a role for NAC transcription factors**. *Acta Horticulturae* 2019(1248):387-402.

79. Wei L, Cao Y, Cheng J, Xiang J, Shen B, Wu J: **Comparative transcriptome analyses of a table grape 'Summer Black' and its early-ripening mutant 'Tiangong Moyu' identify candidate genes potentially involved in berry development and ripening**. *Journal of Plant Interactions* 2020, **15**(1):213-222.

80. Sievert C: **Interactive Web-Based Data Visualization with R, plotly, and shiny**: CRC Press; 2020.

81. Osier MV: **VitisPathway: gene pathway analysis for V. vinifera**. *Vitis* 2016, **55**:129-133.

82. Grimplet J, Van Hemert J, Carbonell-Bejerano P, Díaz-Riquelme J, Dickerson J, Fennell A, Pezzotti M, Martínez-Zapater JM: **Comparative analysis of grapevine whole-genome gene predictions, functional annotation, categorization and integration of the predicted gene sequences.** *BMC Research Notes* 2012, **5**:213.

83. **pheatmap: Pretty Heatmaps. R package version 1.0.12** [https://CRAN.R-project.org/package=pheatmap]

84. Zhang P, Berardini TZ, Ebert D, Li Q, Mi H, Muruganujan A, Prithvi T, Reiser L, Sawant S, Thomas PD *et al*: **PhyloGenes: An online phylogenetics and functional genomics resource for plant gene function inference**. *Plant Direct* 2020, **4**(12):e00293.