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

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

49 **Conclusions**

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

58

59 **Key words**

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

62

63 **Background**

64 *Vitis vinifera subsp. vinifera* (grapevine) belongs to the family *Vitaceae*. With 6,000 to 11,000
65 cultivars, it is one of the most important perennial crops worldwide [1]. Grapevine fruit
66 development can be divided into two physiological phases, berry formation and berry
67 ripening. The time of véraison refers to the transition of both phases whereby each phase is
68 represented by a sigmoidal growth curve of development [2]. The progress through
69 development is described by stages referred to as "BBCH stages" that have been defined for
70 several crops including grapevine [3, 4]. The first physiological phase described as berry
71 formation (berry initiation and growth with cell divisions) is lasting from the end of flowering
72 (BBCH71) until the majority of berries are touching each other (BBCH79), approximately 60
73 days later. The developmental stage of véraison (BBCH81) depicts the end of berry formation
74 and the start of berry ripening [2]. Phenotypically, véraison is the developmental switch
75 when the berries start to soften, accompanied by the onset of accumulation of
76 phenylpropanoids. In red grapevine cultivars, véraison is also indicated by a color change of
77 the berries that is caused by the beginning of accumulation of anthocyanins which are one
78 class of phenylpropanoids. Members of the well-studied protein superfamily of R2R3-MYB
79 transcription factors (TFs) are considered to be mainly accountable for controlling
80 anthocyanin accumulation [5-7]. After véraison, berry ripening continues until harvest
81 (BBCH89) with cell enlargement, sugar accumulation and acidity decline.
82 Anthropogenic climate change is resulting in successively earlier ripening of grapes with a
83 significant impact on berry quality and on typicity of a desired wine style [8]. In addition, the
84 time of véraison and for harvest of a given cultivar may differ greatly, driven by regional
85 and/or year-depending differences in weather conditions. Obviously, this calls for a better
86 molecular understanding of the control of ripening time in grapevine.

87 Comparison of grapevine cultivars grown at the same environmental conditions often
88 uncovers differences in ripening time. However, the underlying genetic factors are mostly
89 unknown. Previous studies have elucidated how ripening time is affected by external factors.
90 For example, the effect of phytohormones on berry ripening has been widely studied [1]. In
91 general, fruit growth is discussed to be controlled by several phytohormones, which play
92 essential roles to trigger or delay ripening processes [9]. In grapevine as a non-climacteric
93 fruit, effects of abscisic acid (ABA) have been investigated in many studies and ABA is
94 considered to trigger ripening [10-12]. Furthermore, it was shown that ABA is involved in
95 controlling leaf senescence [13], responses to drought [14] and pathogen defense [15]. In
96 grapevine, although not as central as in climacteric fruits like tomato (*Solanum*
97 *lycopersicum*), the phytohormone ethylene is also involved in the control of berry ripening
98 [1, 11, 16, 17]. Application studies on grapevine berries indicate that auxin has a negative
99 effect on berry ripening in that it induces a delay of ripening [18, 19].
100 Fruit development of both, dry and fleshy fruits, has been studied very intensively for the
101 obvious reason that fruits are central to human nutrition [20, 21]. The main model system
102 for studies on fleshy fruits is tomato, because of established genetics and molecular biology,
103 access to mutants, and well advanced transgenic approaches to gene function identification
104 [22, 23]. Berry development of grapevines has also been studied intensively [1, 24] and often
105 at the level of the transcriptome. In quite some of the studies, predominantly late berry
106 development stages were sampled to put véraison into the focus [25-28]. In addition, whole
107 berry development was studied with coarse time point distribution [29-33].
108 To monitor gene expression changes at a high resolution throughout grapevine berry
109 development, starting from flowering until berries are matured, we sampled a
110 comprehensive time series from two Pinot cultivars and in two years. The samples were

111 collected from the grapevine cultivar 'Pinot Noir' (PN) and the comparably earlier ripening
112 cultivar 'Pinot Noir Precoce' (PNP) that is expected to be closely related to PN. The cultivar
113 PNP is listed in the *Vitis* International Variety Catalogue (*V*IVC; [34]) and described to flower
114 at the same time as PN but to reach véraison significantly earlier than PN [35]. Quantitative
115 data for transcript levels, interpreted as values for gene expression, were generated by RNA-
116 Seq. We studied the general course of gene expression patterns throughout berry
117 development in both years and cultivars, and identified a number of differentially expressed
118 genes (DGEs) between PN and PNP prior to véraison. These DEGs can be considered as
119 important candidates for either delaying or pushing forward berry development. Our main
120 aim was the identification of genes controlling the speed of development, to offer an entry
121 point into characterization of the relevant molecular functions in grapevine, and to facilitate
122 breeding in the future that addresses traits relevant to and affected by climate change.
123

124 **Results**

125 **Phenotypical comparison between two Pinot cultivars**

126 To study ripening shifts, we used samples of two closely related grapevine cultivars. The
127 cultivar PNP is an earlier ripening clonal variant of its ancestor PN. Clonal relation of PN and
128 PNP was confirmed by a set of 24 SSR markers that all displayed the identical allele status for
129 both cultivars (Additional file 1: Table S1). To confirm and validate the phenotypic
130 differences between PN and PNP, detailed BBCH developmental stages were determined
131 and documented (Figure 1). PN and PNP display similar phenotypical properties during
132 development and flower (BBCH65) at the same time. However, véraison (BBCH81) is shifted
133 to about two weeks earlier for PNP and similar shifts were detected in four documented
134 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.

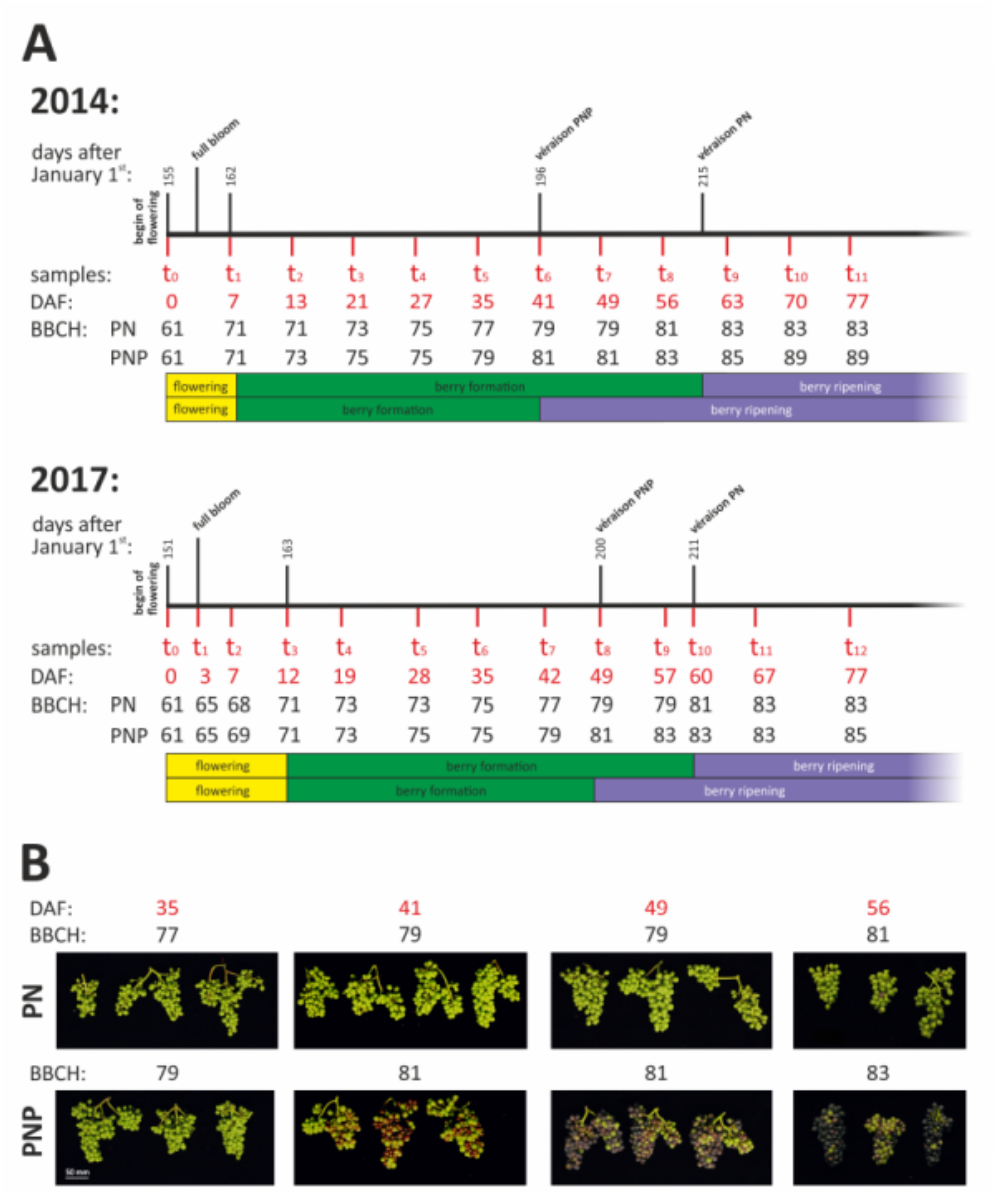


Figure 1: Phenotypical observations and sampling scheme. (A) Sampling time points and days after begin of flowering (DAF) are indicated in red. The developmental stage observed is shown in the BBCH stages [3, 4]. Berry development is depicted schematically and

144 categorized into the phases flowering (yellow), berry formation (green), and berry ripening
145 (purple) for both cultivars. The junction between green and purple indicates véraison
146 (BBCH81). As an orientation for time of the year, days after January 1st are shown for distinct
147 transitions of the phases. (B) Images of grape bunches and developing berries taken in 2014
148 to document the difference between PN and PNP. Images were taken 35, 41, 49 and 56 DAF.
149 Scale bar: 50 mm.
150

151 **Table 1**: Observed flowering- and berry development shifts between the cultivars PNP and
152 PN in 2014, 2015, 2016 and 2017 at the vineyards Geilweilerhof, Siebeldingen (in days after
153 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		

154

155 **Global view on gene expression patterns**

156 We harvested the samples in 2014 and 2017 (for time points see Figure 1 and Additional file
157 1 Tables S2 and S3) in triplicates (individual harvests are referred to below as subsamples)
158 and analyzed them by RNA-Seq. After preprocessing of the raw data (see Methods), the
159 reads derived from each subsample were mapped to the reference sequence from PN40024
160 and analyzed with respect to the CRIBI V2.1 annotation dataset. For 2014, approximately
161 19.7 million reads per subsample were obtained from each of the 72 libraries. An overall

162 alignment rate of 79 % to the grape reference genome sequence was reached. For 2017,
163 approximately 43.5 million reads per subsample were obtained from each of the 78 libraries.
164 From these, an overall alignment rate of 92 % to the reference was calculated. Expression
165 values were initially detected as Transcripts Per Kilobase Million (TPM) and averaged over
166 the three subsamples for each sample. Considering both years separately, a total of 28,692
167 genes were detected as expressed in both cultivars and in both years. In contrast, 2,152
168 CRIBI V2.1 genes were found not to be expressed.
169 The correlation between gene expression data, determined as TPM values per sample, of the
170 datasets from both years over all genes was $r = 0.5095$ (Pearson correlation coefficient) for
171 PN and $r = 0.6557$ for PNP, respectively. For PN and PNP, 10,205 and 16,226 genes,
172 respectively, expression values were significantly correlated (p-value < 0.05) between the
173 years 2014 and 2017. A list of the correlation strength of the eight time points with the same
174 BBCH stage is provided in Additional file 1: Table S4.
175 To visualize global trends and similarity of the gene expression values obtained from all
176 subsamples, a Principal Component Analysis (PCA) of both years was performed with
177 keeping the subsamples separate. The first component PC1 explains 57% of the variance,
178 whilst the second component PC2 explains 17% (Figure 2). Almost all data points of the
179 subsamples (triplicates within a sample) from both years cluster near to each other. The data
180 follow a track of time in a nearly consecutive and continuous way. Main actors, which
181 influence most of the variance in the data, were genes related to cell wall modification,
182 secondary metabolism, wounding-response and hormone signaling. The top 500 genes
183 describing most of the variance in PC1 and PC2 are listed in Additional file 1: Table S5 and
184 Table S6, together with functional information for each listed gene from MapMan/Mercator
185 and RefSeq.

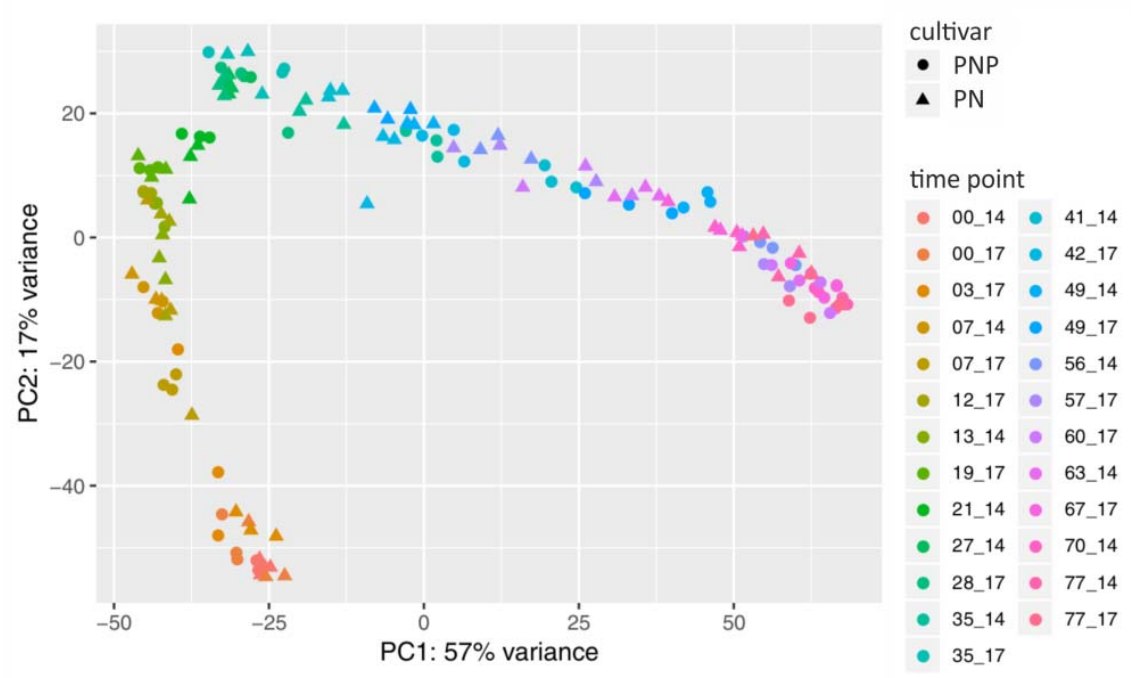


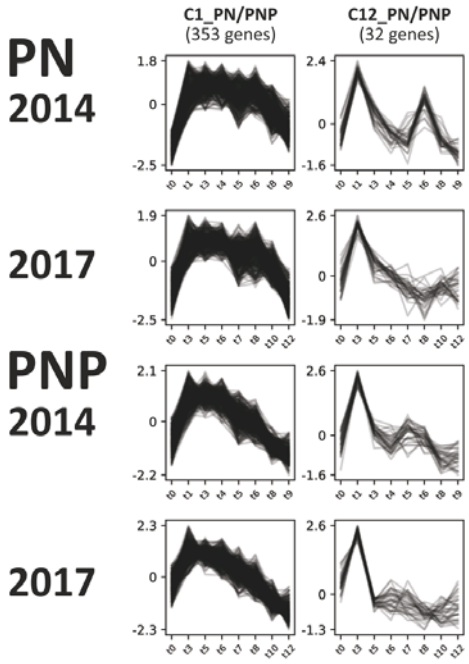
Figure 2: Principle component analysis of gene expression values from all subsamples. Each data point represents a single subsample of the triplicates for each time point of both years (2014 and 2017 as indicated by [DAF]_14 and [DAF]_17 with the color code) and for both cultivars (PN as triangles, PNP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

Cluster analysis for identification of co-expressed genes

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

202 between the clusters, which was in part the result of the restricted the number of clusters
203 that clust extracts. Manual inspection of the clusters revealed little deviation of individual
204 gene expression profiles within each individual cluster for a given year or cultivar. All cluster
205 gene memberships, also those for the additional cultivar-specific cluster analyses (see
206 below), are available in Additional file 1: Table S7-S9.
207 The PN/PNP clusters C2, C5, C6 and C12 (C12_PN/PNP selected as example, see Figure 3)
208 reveal a small but detectable difference in the gene expression profile between both
209 sampled years, but are almost identical for both cultivars. Thus, the genes in these clusters
210 may display dependence on environmental factors in their expression patterns, potentially
211 due to differences in the weather conditions between the two years studied. The PN/PNP
212 clusters C1, C7 and C11 (C1_PN/PNP selected as example, see Figure 3) show similar
213 expression profiles over the two years, but stand out by shifted expression peaks that
214 distinguish PN and PNP.
215 To characterize the clusters with respect to potential functions of the co-expressed genes
216 included in a given cluster, GO term enrichment for biological processes was calculated. The
217 full list of enriched GO terms for all clusters is listed in Additional file 1: Table S10-S12. Two
218 examples for GO terms showing up with highly significant incidence were 'response to
219 oxidative stress' in cluster C11_PN/PNP (term GO:0051276) and 'regulation of defense
220 response' in cluster C5_PN/PNP (term GO:0031347).
221 Two additional cluster analyses were performed, one for the PN data from both years
222 (Additional file 2: Figure S1B) and one for the PNP data from both years (Additional file 2:
223 Figure S1C). These analyses revealed a high abundance of genes from the expansin gene
224 family showing a similar expression profile in clusters C0_PNP and C6_PN. Moreover, the
225 cluster C16_PN showed a highly significant enrichment for 'vegetative to reproductive phase

226 transition' (GO:0010228). Cluster gene memberships for the cultivar-specific clustering are
227 available in Additional file 1: Table S8-9, and the corresponding GO term enrichment is
228 summarized in Additional file 1: Table S11-12.
229



230

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

240 **Analyses of differentially expressed genes**

241 The gene expression time series throughout berry development were analyzed for
242 differentially expressed genes (DEGs) between the two cultivars PN and PNP with DESeq2.

243 Significantly differentially expressed genes were selected by using the filters adjusted p-
244 value (PADJ) < 0.05 and log2fold change (LFC) > 2. The results are summarized in (Table 2)
245 and are detailed at the gene level per time point compared in Additional file 1: Table S13.
246

247 **Table 2:** Filtering steps applied for selecting DEGs, and the number of DEGs that were carried
248 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	

249

250 In total, 8,206 and 4,419 DEGs were identified with PADJ better than 0.05 for 2014 and 2017,
251 respectively. Almost twice as many DEGs were initially detected for 2014 compared to 2017.
252 By applying the filter for an at least 2-fold difference in expression level (LFC > 2), the
253 number of significant DEGs decreased, mainly for the PN/PNP time series from 2014. In the
254 beginning of both time series, that is during flowering (BBCH61 to 69), only few DEGs
255 between PN and PNP were observed (see Figure 4). Within berry formation (BBCH71 to 79),
256 the number of DEGs detected increased towards véraison (BBCH81) as both genotypes

257 increasingly vary for their physiological stage. The highest number of DEGs was observed in
258 parallel to the time-shifted véraison of PNP relative to PN that phenotypically is the most
259 prominent difference between the two cultivars. A set of véraison-specific genes was
260 defined by selecting the DEGs from time points DAF35 and DAF41 from 2014 that show also
261 up at DAF42 and DAF49 from 2017. These criteria identified 388 véraison-specific DEGs. This
262 set of véraison-specific genes was compared to results from similar studies and matches very
263 well (e.g. 81% [25] and 52% [26]; IDs of the 388 genes, the genes that match results from the
264 other studies and their functional annotation are included in Additional file 1: Table S14).
265 During the subsequent phase of berry ripening (BBCH81 to 89), the number of DEGs
266 detected decreases.
267 We developed a visualization for the numbers of DEGs detected and the changes with
268 respect to which genes are newly appearing as differentially expressed at a given time point
269 (sample pair PN/PNP) in the time series. Figure 4 shows this visualization. Groups of newly
270 appearing DEGs relative to an earlier time point were indicated by a new color shade in the
271 column (bar) for each time point. For members of a given group of DEGs, the attributed
272 color shade was kept for the subsequent columns. This implies that DEGs are initially
273 counted for a pair of PN/PNP samples for each time point individually. If DEGs appearing in
274 several time points are counted only once, 3,342 and 2,745 unique DEGs (different genes)
275 are detected from 2014 and 2017, respectively (compare Table 2).
276

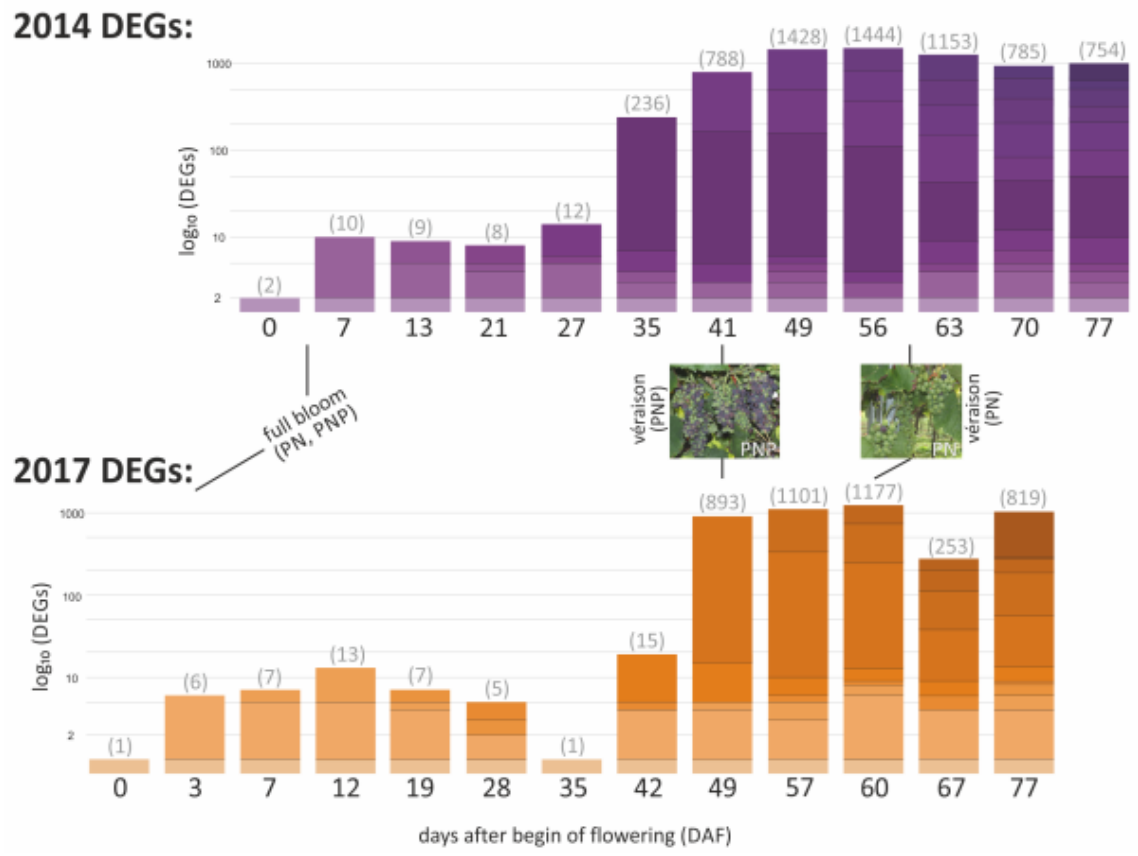


Figure 4: Visualization of the number of DEGs detected between PN and PNP in a logarithmic scale (log_{10}). Results for 2014 are shown in purple, those for 2017 in orange. The time series from the two years were aligned at véraison of PNP; the timeline is given as days after begin of flowering (DAF). 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

292 these putative ripening-regulatory DEGs are supposed to be relevant before the set of
293 véraison-specific genes implements the phenotypic changes at véraison. The full list of DEGs,
294 their identity and annotation information as well as their fit to the selection criteria on the
295 way from all (raw) DEGs to ripening-regulatory DEGs is detailed in Additional file 1: Table
296 S13. IDs of the 12 putative ripening-regulatory genes, the genes that match results from
297 related studies (7 DEGs [25], 4 DEGs [33] and 3 DEGs [26]) and their functional annotation
298 are included in Additional file 1: Table S14.
299
300 **Functional classification of DEGs**
301 To complement the gene lists with functional information from grapevine that might
302 potentially be informative for berry development, the 1,923 intersecting DEGs were
303 analyzed with respect to enrichment of genes that have been assigned to biological
304 pathways already established for grapevine (see Methods). For 46 of the 247 defined
305 grapevine pathways, significant enrichment (permuted p-value <0.1) was detected. The most
306 reliable predictions (permuted p-value <0.001) for pathways that might be relevant were
307 photosynthesis antenna proteins (vv10196; 9 DEGs); nitrogen metabolism (vv10910, 19
308 DEGs); phenylpropanoid biosynthesis (vv10940, 66 DEGs), tyrosine metabolism (vv10350, 33
309 DEGs); transport electron carriers (vv50105, 18 DEGs); phenylalanine metabolism (vv10360,
310 33 DEGs); brassinosteroid biosynthesis (vv10905, 8 DEGs) and flavonoid biosynthesis
311 (vv10941, 30 DEGs). The enrichment results are provided in (Additional file 1: Table S15). The
312 same analysis was also carried out for the 12 putative ripening-regulatory (Additional file 1:
313 Table S16) and the 388 véraison-specific DEGs (Additional file 1: Table S17).
314 A check of the 1,923 intersecting DEGs revealed that 141 TF genes are included. Of these, 48
315 DEGs were clearly up- and 93 down-regulated at their first appearance in the time series.

316 The full list of TF encoding genes that were higher expressed in PNP (up-regulated), or lower
317 expressed in PNP (down-regulated), compared to PN, is shown in Additional file 1: Table S18.
318 For a more detailed view on the expression patterns of selected TF encoding genes, we
319 generated for the TF gene family with the highest abundance among the 141 TF genes,
320 namely the R2R3-MYB-type TFs with 22 cases in the MapMan functional assignment, an
321 expression heatmap (Additional file 2: Figure S2). As a result, *VviMYB24*
322 (VIT_214s0066g01090), which is related to At3g27810/*AtMYB21*, At5g40350/*AtMYB24* and
323 At3g01530/*AtMYB57* according to TAIR/PhyloGenes, was identified as an early appearing
324 DEG that showed its highest expression level at flowering (BBCH61). Prominent *R2R3-MYB*
325 genes known to be relevant for anthocyanin accumulation like *VviMYBA1*
326 (VIT_202s0033g00410), *VviMYBA2* (VIT_202s0033g00390), *VviMYBA3*
327 (VIT_202s0033g00450) and *VviMYBA8* (VIT_202s0033g00380) were detected as expressed
328 starting from véraison (BBCH81) in both cultivars and with a time shift towards earlier
329 expression in PNP. An additional *R2R3-MYB* gene with a similar expression pattern is
330 *VviMYB15* (VIT_205s0049g01020). Other *R2R3-MYB* genes are expressed early during berry
331 formation, these include *VviMYBF1* (VIT_207s0005g01210, related to
332 At2g47460/*AtMYB12*/*AtPFG1*) as well as *VviMYBPA5* (VIT_209s0002g01400) and *VviMYBPA7*
333 (VIT_204s0008g01800, both related to At5g35550/*AtMYB123*/*AtTT2*). According to their
334 related expression patterns visualized in the heatmaps (Additional file 2: Figure S2), the
335 *R2R3-MYB* genes fall into three groups that roughly fit to the three phases marked in Figure
336 1B, namely flowering, berry formation, and berry ripening (see discussion).
337
338 **Putative candidates for ripening-regulatory genes**

339 As mentioned above, DEGs detected in both years at time points before véraison of PNP
340 were selected and considered as putative ripening-regulatory genes (Table 2, see full list in
341 Additional file 1: Table S13). The VitisNet enrichment analyses performed for these 12
342 candidates resulted in 2 pathways that showed significant (permuted p-value < 0.05)
343 enrichment with two genes in the pathway: auxin signaling (vv30003 with *VviEXPA5*
344 (VIT_206s0004g00070) and *VviEXPA14* (VIT_213s0067g02930)) and cell wall (vv40006 with
345 *VviPL1* (VIT_205s0051g00590) and *VviGRIP28* (VIT_216s0022g00960)); see Additional file 1:
346 Table S16).
347 A detailed check of the data presented in Figure 4, together with results for the putatively
348 ripening-regulatory DEGs, resulted in the identification of two DEGs that stand out from the
349 whole list of DEGs. Both genes almost completely lack expression in the early ripening
350 cultivar PNP while there is clear expression in PN. Therefore, these two genes were detected
351 as DEGs throughout the whole time series in both years. The first of the two, designated
352 *VviRTIC1* for "Ripening TIme Control" (VIT_210s0071g01145, encoding a protein similar to
353 "protein of unknown function DUF789"), is expressed during flowering (BBCH61 - 65) and is
354 more or less continuously down-regulated over time in PN (Figure 5A). The second of the
355 two, designated *VviRTIC2* (VIT_200s0366g00020, encoding a protein similar to "cysteine-rich
356 receptor-like protein kinase"), displays expression in PN during berry formation as well as
357 during berry ripening with a peak before véraison in 2017 (Figure 5B). The expression
358 patterns of the 10 remaining DEGs of the putative ripening-regulatory gene set are shown in
359 Additional file 2: Figure S3.
360

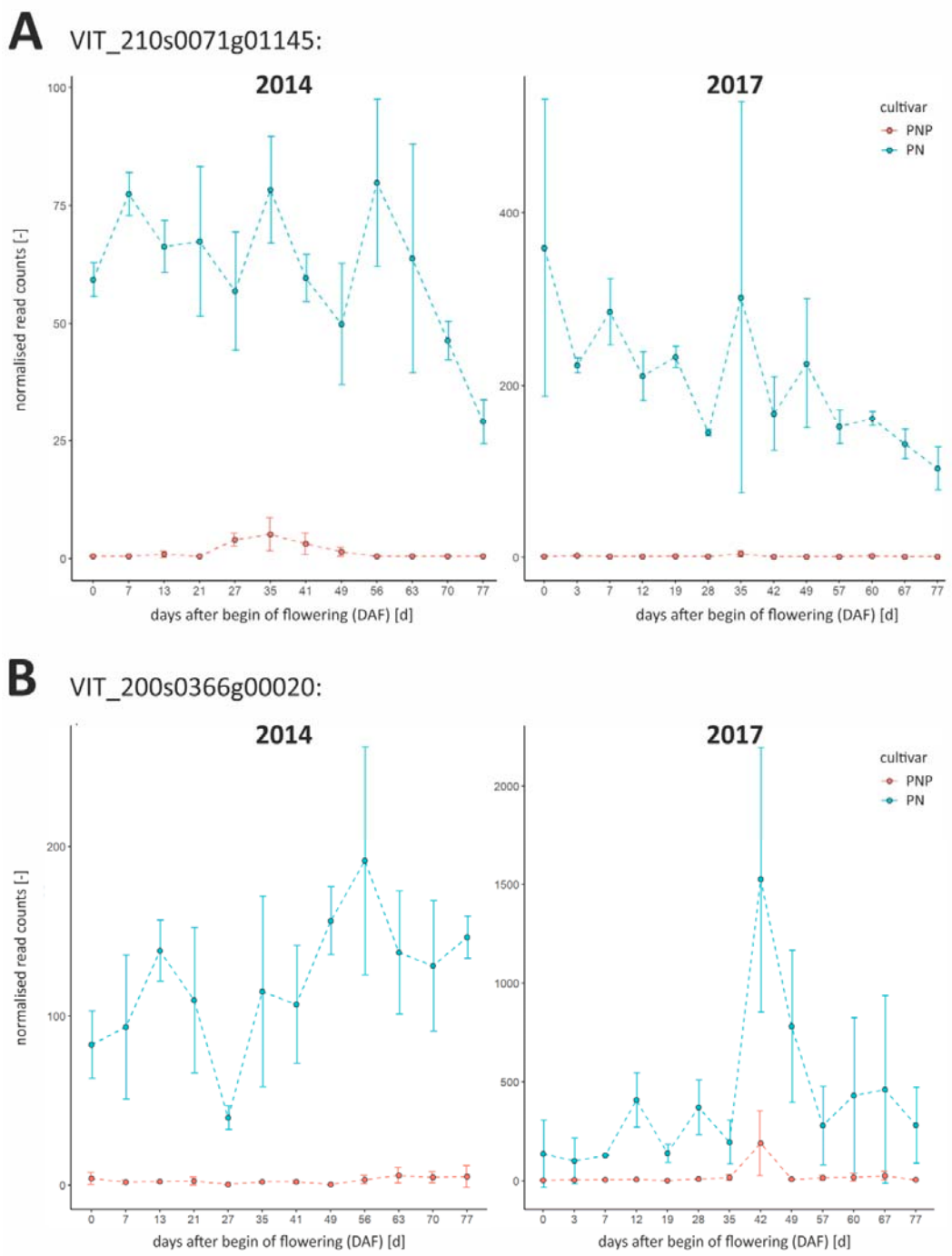


Figure 5: Expression patterns of *VviRTIC1* (VIT_210s0071g01145) in (A) and *VviRTIC2* (VIT_200s0366g00020) in (B) from RNA-Seq data of PN (blue) and PNP (red). Error bars display the standard deviation of triplicates. Left, expression profile from 2014. Right, expression profile from 2017. The y-axis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

368 **Discussion**

369 One of the first detectable mentions of the cultivar 'Pinot Precoce' in connection with the
370 synonym '(German) Früh Burgunder Traube' (PNP) is in the French book "Ampelographie
371 retrospective" [36]. We have confirmed the clonal relationship of PN and PNP by 24
372 genomically well distributed SSR markers. Although this does not prove that PN is the
373 ancestor, it is very likely that PNP was derived from the cultivar PN by somatic mutation [35].
374 We used these two isogenic cultivars, that are distinguished by a clear phenotype regarding
375 the duration of berry formation, to analyze changes in gene expression throughout berry
376 development. The aim was to identify candidates for genes controlling the speed of berry
377 development and the timing of véraison. Samples from inflorescences as well as from
378 forming and ripening berries were collected from the begin of flowering until after véraison
379 of PN and PNP in 2014 and 2017. These samples were subjected to RNA-Seq analyses in two
380 dense time series.

381

382 **Phenotypic differences between the cultivars PN and PNP**

383 The data from 4 years of careful assessment of the BBCH developmental stages of PN and
384 PNP at the same location validate earlier observations from viticulture [35] that lead to
385 establishment of PNP as a distinct grapevine cultivar in north European wine growing
386 countries. Berry formation lasts about two weeks less in PNP, is clearly accelerated
387 compared to PN and results in PNP entering véraison approximately two weeks earlier than
388 PN (Figure 1A, Table 1). It is reasonable to assume that this acceleration affects berry
389 formation throughout, i.e. from immediately after fruit set until véraison. Functionally, this
390 hypothesis implies that the genes that are responsible for the control of timing of berry
391 development and for the establishment of the phenotypic difference between PN and PNP

392 should be acting already very early in berry development, starting at least shortly after
393 flowering and at or even before BBCH61 to BBCH79. At the end of ripening (harvest), the
394 berries of PNP reach high sugar content earlier within the season when compared to PN.
395
396 **General validation of the RNA-Seq dataset**
397 To estimate overall data quality, the expression profiles obtained from PN and PNP were
398 correlated for the two sampled years 2014 and 2017. Pearson correlation was moderate, but
399 this is expected considering the conditions of the free field environment. Exposure of the
400 vines to external factors like biotic or abiotic stressors, including weather conditions that
401 differ significantly between the years, also affect the transcriptome which reduces the level
402 of correlation [37]. In a PCA, almost all datapoints lie on the same intended track, and
403 biological replicates (subsamples) from both years are located close to each other. The main
404 actors, which predominantly influence the variance in the dataset, are genes related to cell
405 wall modification, secondary metabolism, wounding and hormone signaling. These gene
406 categories fit the expectation since berry development is known to be controlled by
407 hormones, requires new cell walls, and is accompanied by accumulation of
408 secondary/specialized metabolites [22, 24]. These initial results validated the quality of the
409 dataset and indicated clearly that sampling of biologically closely related material for the
410 subsamples/triplicates was successful.
411
412 **Co-expression analysis shows similar gene expression clusters between cultivars and years**
413 To further validate the data with respect to comparability as well as reproducibility between
414 the two years, related gene expression profiles were identified among all genes by clustering
415 the data from the four different time series. Generally, clusters of the same genes with

416 similar expression patters over time were observed for both cultivars and both years. Also,
417 the cluster analyses for gene expression patterns among the years 2014 & 2017 in only PN
418 and in only PNP, confirmed comparability of the gene expression patterns obtained in these
419 two years. Detailed inspection revealed clusters representing expression profiles (and
420 clusters of genes) with and without an environmental influence. Especially the cultivar-
421 specific clusters C1_PN/PNP, C7_PN/PNP and C11_PN/PNP stand out. Comparison of the
422 expression profiles for PN to those of PNP in these clusters identified a similar pattern that is
423 moved to a different time in PNP. These findings are congruent with the shifted ripening
424 time phenology of the two cultivars, namely accelerated berry formation in PNP by about
425 two weeks. In contrast, the clusters C2_PN/PNP, C5_PN/PNP, C6_PN/PNP and C12_PN/PNP
426 display more pattern similarity among the two years than among the two cultivars which
427 indicates environmental control of the expression of the genes in these clusters. Strong
428 environmental effects on gene expression patterns have also been described for grapevine
429 berry development at 11 different environments (vineyards) from northern Italy [38]. By
430 approximation, highly similar expression profiles are summarized in the PN/PNP clusters C3,
431 C4, C7, C8, C9 and C10. The genes included in these clusters are probably less affected by
432 environmental factors and/or the genotypic difference between PN and PNP. We conclude
433 that our RNA-Seq results and expression level comparisons between two years are based on
434 valid data.
435 When the genomic location of the DEGs is analyzed, a genome region on chromosome 16
436 comes into focus. In this region, 54 of the DEGs from the set of 1,923 intersecting DEGs
437 (Table 2) are located. Of these, 28 encode stilbene synthases [39] that are all up-regulated
438 after véraison of PNP (BBCH83). This genome region fits to a major QTL (Ver1) for "timing of
439 the onset of véraison" on linkage group 16 [40]. It remains to be determined if this

440 aggregation of DEGs is by chance. Potentially, the observation is biased by co-regulation of a
441 large number of closely linked stilbene synthase genes.
442
443 **Differentially expressed genes throughout berry development and identification of**
444 **véraison-specific genes**
445 Differential gene expression analysis and subsequent filtering revealed 1,923 DEGs between
446 PNP and PN. DEG detection was based on a comparison of samples taken from the two
447 cultivars at very similar DAF. As expected for the characterized phenotype, PNP reaches
448 véraison when PN is still in the phase of berry formation. Consequently, the strong increase
449 in the number of detected DEGs shortly before and at véraison of PNP results from the
450 different developmental stage of PNP compared to the lagging PN. Subsequently, when also
451 PN enters véraison, the number of DEGs declines (note that Figure 4 uses a logarithmic
452 scale). A list of 388 genes that show up in both years with a véraison-specific expression
453 pattern was extracted and compared to published results. Interestingly, about 81.5% of the
454 388 PN/PNP véraison-specific genes were also described in the 4,351 differentially expressed
455 genes between the table grape cultivar '8612-CTR' (wild type) and its early ripening bud
456 mutation '8612-ERN' [25]. Also, analyses of berries from the cultivars 'Cabernet Sauvignon'
457 and 'Pinot Noir' by RNA-Seq identified a gene set of 5,404 genes marking the onset of berry
458 ripening [26]. This set covers 51.5% of the 388 PN/PNP véraison-specific genes (Additional
459 file 1: Table S14). Further, several "switch genes" which are supposed to encode key
460 regulators of the developmental transition at véraison [33, 41] are included in the 388
461 véraison-specific gene set (Additional file 1: Table S14). We conclude that the PN/PNP
462 véraison-specific set of 388 genes represents a core set of genes that are relevant for
463 executing the switch from berry formation to berry ripening. The fact that a relatively small

464 gene set was detected that still displays high overlap to studies that addressed a similar
465 biological question indicates that the specific experimental setup implemented filters acting
466 well against unrelated genes. The comparison of "wildtype to mutant" results in isogenic
467 background, growth in the same vineyard/location reduced environmentally controlled
468 transcriptome differences, and dense sampling together with high RNA-Seq read coverage
469 allowed good resolution power.
470 In order to check for potentially co-expressed genes within the véraison-specific gene set,
471 the memberships for these genes in the PN/PNP cluster analysis were investigated. A total of
472 48 véraison-specific DEGs were detected in cluster C6_PN/PNP (contains 914 genes). These
473 48 genes include several prominent ripening-related genes like *VviGRIP61*
474 (VIT_201s0011g05110), *VviMYBA8*, *VviMRIP1* (VIT_205s0049g00760, [42]), *VviGRIP4*
475 (VIT_205s0049g00520) and *VviGRIP28*. The about 20 *VviGRIP* genes were initially detected
476 by differential cDNA screening as ripening-induced genes in grape [43]. Another relevant
477 cluster is C5_PN/PNP (contains 263 genes) which includes 37 of the 388 véraison-specific
478 DEGs. Among these are *VviMYBA1*, *VviMYBA2*, *VviMYB15* and *VviGRIP22*
479 (VIT_206s0004g02560). The two clusters C5_PN/PNP and C6_PN/PNP show quite similar
480 patterns (Additional file 2: Figure S1A). At first glance it is not obvious which difference has
481 forced clust to put a given gene in either C5_PN/PNP or C6_PN/PNP. A comparison of the
482 expression patterns of *VviMYBA2* (in C5_PN/PNP) and *VviMYBA8* (in C6_PN/PNP; see
483 Additional file 2: Figure S2 for a heatmap) shows that there are borderline cases regarding
484 assignment to either C5_PN/PNP or C6_PN/PNP.
485 In total, 22 genes encoding R2R3-MYB TFs were found among the 1,923 intersecting DEGs.
486 Based on the timing of expression in PN and PNP, the 22 *R2R3-MYB* genes can be classified
487 into three groups (Additional file 2: Figure S2). The first group is represented by *VviMYB24*

488 which is expressed during early flowering (BBCH61) but switched off already at the transition
489 from flowering to berry formation (BBCH71). *VviMYB24* is potentially orthologous to a group
490 of three *A. thaliana R2R3-MYB* genes (*AtMYB21/24/57*) that are expressed in flowers and
491 function redundantly to regulate stamen development in the context of jasmonate action
492 [44]. It is tempting to speculate that *VviMYB24* has a similar function in grape.
493 The second group covers about 15 *R2R3-MYB* genes that are expressed during berry
494 formation and pre-véraison (BBCH71 to 77). This group includes *VviMYBF1* which regulates
495 flavonol biosynthesis [45], and *VviMYBPA5* as well as *VviMYBPA7* which belong to the clade
496 of *AtTT2*-related genes that control proanthocyanidin (PA, flavan-3-ol) biosynthesis [7, 46,
497 47]. The other *R2R3-MYB* genes in this group are less well characterized although there are
498 functions described for some of them, e.g. *VviMYBC2-L3* (VIT_214s0006g01620) as repressor
499 of specific branches of the phenylpropanoid pathway [48].
500 The third group of *R2R3-MYB* genes is active starting at véraison (after BBCH81) and covers
501 about six genes. Among them are the anthocyanin accumulation controlling genes,
502 *VviMYBA1*, *VviMYBA2, VviMYBA3* and *VviMYBA8* for which there is good evidence that they
503 trigger anthocyanin biosynthesis [49]. Since PN and PNP are red berry cultivars, activity of
504 the TF genes that direct anthocyanin accumulation is expected. In addition, this group
505 includes *VviMYB14* and *VviMYB15* that are supposed to regulate the stilbene biosynthetic
506 pathway [50]. With regard to the heatmaps (Additional file 2: Figure S2) and the analyses of
507 the DEGs in this study in general, it should be noted that while the resolution within the
508 developmental program and time is quite good, our data do not resolve the exact location of
509 gene expression. Therefore, it remains to be determined if the expression detected is
510 derived from berry skin, flesh, the seed or other tissues/cells.
511

512 **Putative ripening-regulatory genes acting early in berry development**

513 To focus on genes that are contributing to the acceleration of berry formation in PNP, and/or
514 to the control of timing of véraison, we selected DEGs detected at time points prior to
515 véraison of PNP (Table 2, Figure 4). This resulted in a set of only 12 genes that are potentially
516 involved in the regulation of ripening time. According to our hypothesis that the genes
517 relevant for acceleration of berry formation in PNP, which cause the earlier begin of ripening
518 of PNP, should be acting from at least shortly after flowering, we designated this set of genes
519 as "ripening-regulatory". However, genes that encode components of the respective
520 regulatory networks and target genes of regulators including secondarily affected DEGs are
521 surely included as well [51]. The 12 ripening-regulatory genes, i.e. the DEGs detected before
522 véraison of PNP, encode proteins related to auxin action, pectin processing enzymes related
523 to cell wall modification, TFs from the HD-Zip as well as NF-Y/LEC families, a cysteine-rich
524 receptor-like protein kinase, an oleosin, and proteins with domains of unknown function.
525 The two genes assigned to auxin signaling by VitisNet (vv30003) encode expansins (*VviEXPA5*
526 and *VviEXPA14*, [52]). Expansins are known to be involved in fruit ripening through cell wall
527 expansion and cell enlargement [53]. Auxin can delay the onset of véraison and ripening
528 processes in grapevine [17-19]. Since reduced expression of genes from the auxin signaling
529 pathway may indicate reduced auxin action due to lower auxin levels, the accelerated entry
530 of PNP into véraison might be initiated by reduced auxin levels. Additionally, the genes
531 *VviPL1* (pectate lyase 1 [54]), *VviPME10* (pectin methylesterase 10, VIT_206s0009g02560)
532 *VviGRIP28* (encoding a pectin methylesterase inhibitor precursor-like protein) are also
533 related to cell wall processes, indicating that cell wall modification is an important target
534 process also prior to véraison [53]. The gene *VviGRIP28* was also detected within a véraison-

535 specific meta-QTL designated ver/ph16.1 [28]. It remains to be determined if this correlation
536 has a functional basis.
537 The two genes in the set of 12 that encode TFs are *VviHDZ28* (VIT_216s0098g01170, [55])
538 and *VviLEC1* (VIT_200s0956g00020, [56]). The *V. vinifera* gene VIT_216s0098g01170 that has
539 been designated *VviHDZ28* has also been considered as a homolog of *AtHB12* (At3g61890),
540 but it seems that *VviHDZ07* (VIT_202s0025g02590) and *VviHDZ27* (VIT_215s0048g02870) are
541 more similar to *AtHB12*. In these cases, which lack clearly assignable homologs, transfer of
542 functional information reaches its limits and might be restricted to concluding that *VviHDZ28*
543 is important for organ development in *Vitis*. The gene *VviLEC1* is one of three genes in *V.*
544 *vinifera* which are homologs of *AtLEC1* (At1g21970, NF-YB9) and *AtL1L* (LEC1-like,
545 At1g21970, NF-YB6). LEC1 and L1L are central regulators of embryo and endosperm
546 development. They are controlling, among other processes, embryo morphogenesis and
547 accumulation of storage reserve [57]. It is tempting to speculate that the reason for the
548 detection of *VviLEC1* among the 12 ripening-regulatory genes is that also seed development
549 needs to be accelerated in PNP compared to PN. This would explain earlier and higher
550 expression of *VviLEC1* in PNP compared to PN as observed (Additional file 2: Figure S3G).
551 Consequently, *VviOLE5* (VIT_216s0013g00880, encoding an oleosin involved in oil body
552 formation [58]) would fit into the picture as relevant for lipid storage during seed
553 development. According to the proposed enzyme function as alcohol acyltransferase of the
554 protein encoded by VIT_209s0018g01490 involved in volatile ester formation [59], this gene
555 could fit in a similar way. For the gene VIT_205s0077g01980 no functional annotation is
556 available (uncharacterized protein), although homologs exist throughout the Magnoliophyta.
557
558 **Candidates for causal genes explaining the difference between PN and PNP**

559 Among the 12 putative ripening-regulatory genes, of which 10 are discussed above, two are
560 especially interesting. Detailed analyses of the full set of DEGs, visualized in Figure 4,
561 resulted in the identification of *VviRTIC1* and *VviRTIC2*, that could possibly be centrally
562 involved in the accelerated berry development and earlier beginning of ripening in PNP
563 compared to PN. The special feature of the expression patters of the two genes (Figure 5) is
564 that both are differentially expressed already at the first time point analyzed which was
565 selected to hit the BBCH stage 61 (flowering before full bloom). Also, both genes are only
566 barely expressed in PNP in both years studied, while expression in PN is high at almost all
567 time points. *VviRTIC1* is annotated to encode a protein containing a domain of unknown
568 function (DUF789), while *VviRTIC2* is annotated to encode a "cysteine-rich receptor-like
569 protein kinase". The best blastP hit to *A. thaliana* protein sequences indicates that it is
570 related to At4g23180/AtCRK10, but a closer inspection shows that similarity to
571 At4g05200/AtCRK25, At4g23160/AtCRK8 and At4g23140/AtCRK6 is almost as high. This
572 ambiguity, and also the fact that the *V. vinifera* genome contains several genes related to
573 *VviRTIC2* (e.g. VIT_210s0071g01200, VIT_202s0087g01020 or VIT_203s0017g01550 as listed
574 by PhyloGenes), complicates transfer of functional information. For *AtCRK10*, relevance in
575 cytokinin signaling has been provided by a systematic analysis of the *A. thaliana* CRKs [60].
576 As pointed out above, it is well possible that the genes we have identified are part of a
577 genetic pathway that realizes control of timing of berry development in *V. vinifera*, but that
578 we have not reached the start of this pathway. The relevance of these two candidate genes
579 in the causal genetic difference between PN and PNP remains to be determined. Phase-
580 separated genome sequences of the cultivars will be required to resolve the genome
581 structure of both alleles of *VviRTIC1 and 2* the genes in PN and PNP for an informative

582 comparison. In future studies, we will address this question, for example by long read DNA
583 sequencing.
584
585 **Conclusions**
586 This study detected 1,923 DEGs between the Pinot cultivars PN and PNP. The two clonal
587 cultivars display a phenotypic difference, namely acceleration of berry development from
588 full bloom to véraison in PNP. We defined 388 DEGs as véraison-specific and 12 DEGs as
589 ripening-regulatory. The relatively small number of véraison-specific genes displays a very
590 high overlap with results published for similar studies (see Additional file 1: Table S14) and
591 could be used for studying a phytohormone network that is similarly realized in PN and PNP,
592 but accelerated by about two weeks in PNP. Additionally, the ripening-regulatory gene set
593 might offer access to a set of genes putatively important for triggering or delaying the start
594 of berry ripening within PNP or PN, respectively. Further investigations are needed on the
595 DNA sequence level to elucidate structural differences in the genomes, the function of the
596 observed DEGs and their role in shifting ripening time in grapevine.
597

598 **Material and Methods**

599 **Plant material and analysis of clonal relation**

600 The grapevine cultivar PNP (Pinot Precoce Noir, VIVC No. 9280) is early ripening and has
601 been described to be related to the cultivar PN (Pinot Noir, VIVC No. 9279) [34] that ripens
602 later than PNP. To prove the clonal relation, DNA from both cultivars was genotyped utilizing
603 24 polymorphic SSR markers (VVS2, VVMD7, VVMD5, VVMD32, VVMD28, VVMD27,
604 VVMD25, VVMD24, VVMD21, VVIV67, VVIV37, VVIQ52, VVIP60, VVIP31, VVIN73, VVIN16,
605 VVIH54, VVIB01, VrZAG83, VrZAG79, VRZAG67, VrZAG62, VMC4F3.1, VMC1B11) as described
606 [61].

607

608 **Phenotypical characterization and sampling of plant material**

609 Plant material was harvested from PN and PNP grapevines trained in trellis. The plants are
610 growing at the vineyards of JKI Geilweilerhof located at Siebeldingen, Germany (N
611 49°21.747, E 8°04.678). The grapevine plants were planted with an interrow distance of 2.0
612 m and spacing of 1.0 m in north-south direction. Inflorescences, developing and ripening
613 berry samples of PNP and PN for RNA extraction were collected in two years with three
614 independent biological replicates (subsamples) each. Sampling took place at systematic time
615 points (12 time points in 2014, 13 time points in 2017), and at approx. 8 a.m. each day. In
616 2014, harvesting took place regularly every 7 days with only two exceptions (one day
617 deviation, DAF 13 and DAF 27). In 2017, harvesting was adapted to BBCH stages (Figure 1A).
618 The timeline in both years is described as days after begin of flowering (DAF), with begin of
619 flowering defined as the day at which 10% of the individual flowers have lost their caps [3].
620 For each subsample within the triplicates, material from two neighboring grapevines was
621 selected. Grapevine plants were weekly phenotyped according to BBCH stage [3, 4].

622 Phenotyping was performed repeatedly to ensure sampling from vines of the same
623 development stage (e.g. percentage of open flowers during flowering, or berry development
624 stage) to reach uniform subsamples. The phenotypical observations were summarized in
625 Additional file 1 - Table S2 and S3. From these, the durations of flowering, berry formation
626 and berry ripening as well as the resulting shifts between the cultivars were calculated
627 (Table 1). Furthermore, images from berry developmental stages of both cultivars were
628 taken in 2014 for 35, 41, 49 and 56 DAF. The sampled material was directly frozen in liquid
629 nitrogen and stored at -70°C until RNA extraction.
630
631 **RNA extraction and cDNA library construction**
632 Biological replicates, i.e. the subsamples, were ground separately under liquid nitrogen.
633 Total RNA was extracted using an RNA Isolation Kit (Sigma-Aldrich Spectrum™ Plant Total
634 RNA) according to suppliers' instructions. For RNA-Seq, 500 ng total RNA per subsample
635 were used to prepare sequencing libraries according to the Illumina TruSeq RNA Sample
636 Preparation v2 Guide. For subsamples from 2014 and 2017, 72 and 78 libraries were
637 constructed and sequenced, respectively. Enrichment of poly-A containing mRNA was
638 performed twice, using poly-T oligos attached to magnetic beads included in the Illumina kit.
639 During the second elution of the poly-A+ RNA, the RNA was fragmented and primed for
640 cDNA synthesis. After cDNA synthesis, the fragments were end-repaired and A-tailing was
641 performed. Multiple indexing adapters were ligated to the ends of the cDNA fragments and
642 the adapter ligated fragments were enriched by 10 cycles of PCR. After quality check and
643 quantification, the libraries were pooled equimolarly.
644
645 **RNA-Seq**

646 Single end (SE) sequencing of the pooled barcoded libraries from 2014 was performed on an
647 Illumina HiSeq1500 in HighOutput mode generating 100 nt reads. For samples from 2017,
648 sequencing was done using an Illumina NextSeq500 generating 83 nt SE reads; two runs
649 were performed with the same pool of barcoded libraries from 2017.
650
651 **Processing of RNA-Seq read data**
652 Raw reads were trimmed with Trimmomatic (version 0.36) [62]. For raw reads from the year
653 2014, the following settings were used: LEADING:10 TRAILING:10 SLIDINGWINDOW:4:15
654 MINLEN:50. In addition, a collection of all available Illumina adapter sequences was supplied
655 to remove matches within the parameter 2:30:10. For raw reads from the year 2017,
656 trimming settings were set to LEADING:6 TRAILING:6 SLIDING WINDOW:4:15 MINLEN:36. All
657 trimmed reads were quality-checked via FastQC (version 0.11.8) [63]. Thus, possible adapter
658 sequences and low-quality bases were removed. All trimmed reads passing QC were mapped
659 to the reference genome sequence PN40024 (version 12Xv2) [64] using the graph-based
660 alignment tool HISAT2 (version 2.1.0) [65, 66] with no additional soft clipping. Afterwards, all
661 tagged genes (structural gene annotation: CRIBI v2.1) were counted with Featurecounts
662 (Bioconductor package Rsubread version 3.8 [67]). To estimate transcript abundance as a
663 measure for gene expression, counts for Transcripts Per Kilobase Million (TPM, [68]) were
664 determined.
665
666 **Basic gene expression analyses**
667 TPM counts from the various samples were used for manual gene expression inspection, for
668 determination of the number of expressed and not expressed genes, and to calculate the
669 correlation between gene expression values from both years. Genes with a TPM value > 0

670 added up over all samples from one year were classified as expressed, conversely genes with
671 a TPM value = 0 added up over all samples as not expressed. A custom python script was
672 applied utilizing the function pearsonr from SciPy python package (v. 1.2.3) [69], which
673 calculates the Pearson correlation coefficient and the p-value for all year-to-year
674 comparisons. Expression data pairs for TPM counts per gene from both sampled years,
675 averaged over the three subsamples of each sample, were used. To test correlations and
676 relationships between expression values from the two years, where samples were harvested
677 with slightly different sampling patterns (see Figure 1A), eight equivalent time points with
678 the same BBCH stages between the cultivars of each year were selected (see Additional file
679 1: Table S4).
680
681 **Principal component analysis**
682 To explore data similarity, a Principal Component Analysis (PCA) was calculated over all gene
683 expression values from both years and cultivars for all subsamples. All data points were
684 normalized using variance stabilizing transformation function 'vst' from the R package
685 DESeq2 (v. 1.12.4) [70]. Subsequently, the principal components were generated using
686 'prcomp' from the R package 'stats' (v. 3.5.2) [71]. The resulting PCA object, displaying the
687 main components PC1 and PC2, was plotted and exported. Additionally, genes with the
688 highest variance contribution to PC1 and PC2 were extracted separately.
689
690 **Functional annotation of genes**
691 Transfer of annotation information from other plant species, mainly *A. thaliana,* was
692 calculated using MapMan's sequence annotation tool Mercator (v. 3.6) [72, 73]. Additionally,
693 all open reading frame (CDS from *V. vinifera*/grapevine genes) sequences were aligned to

694 the non-redundant protein sequence data base RefSeq [74] with the basic local alignment
695 tool for proteins BLASTp [75] (e-value ≤ 0.001). Short descriptions of gene functions were
696 extracted and added to the gene lists in Additional file 1: Table S5, S6, S13, S14.
697 GO term enrichment for biological processes was calculated via the R package 'topGO' (v.
698 2.38.1) [76]. Subsequently, statistical reliability was calculated using Fishers exact test. All
699 Gene IDs and their corresponding GO terms were extracted from the CRIBI database
700 (http://genomes.cribi.unipd.it/DATA/V2/annotation/bl2go.annot_with_GO_description.txt).
701 All results of the GO term enrichment are deposited in Additional file 1 - Table S10-12.
702
703 **Cluster analysis**
704 To reveal co-expressed genes over all four datasets, the tool clust (v. 1.10.8) was used with
705 default parameters [77]. As input, raw read counts from eight time points were used. These
706 time points were selected to cover the same BBCH stages of PN and PNP from the years
707 2014 and 2017 (Additional file 1: Table S4). First, all data were pre-processed as described in
708 the clust manual. Values from corresponding subsamples (triplicates) were combined and
709 averaged. To filter out uninformative (very low) gene expression values, an additional filter
710 was applied: genes not reaching a sample expression value > 1 in at least three conditions
711 and in at least one cultivar from one year were discarded (-fil-v 1 -fil-c 3 -fil-d 1). Afterwards,
712 the data were quantile normalized according to the RNA-Seq defaults of clust. Genes
713 showing a flat expression profile were filtered out by applying the default settings [77].
714
715 **Differential gene expression analyses**
716 For analyses of differentially expressed genes, DESeq2 (v. 1.12.4; R Bioconductor) was
717 employed. To test if gene expression differs significantly between two samples, the

718 likelihood ratio test nbinomLRT, included in the DESeq2 package, was used. Normalization
719 factors and dispersion estimates were used as described [70]. The output table contained all
720 differentially expressed genes (DEGs) and the corresponding values for baseMean,
721 log2FoldChange (LFC), lfcSE (LFC standard error), stat (difference in deviation between the
722 reduced model and the full model), p-value and PADJ (adjusted p-value). To focus on
723 significantly differentially expressed genes from the DESeq2 analyses, cut-off filters PADJ ≤
724 0.05 and LFC > 2 were applied.
725
726 **Selection of gene sets potentially relevant for ripening and comparison with literature data**
727 In order to identify gene sets from the DEGs relevant for control and realisation of ripening,
728 an intersection between the DEGs detected at all time points between both years was built.
729 To determine a subset of putatively ripening-regulatory genes, the intersection between
730 both years covering the development stages BBCH61 (begin of flowering) to BBCH79 (one
731 developmental BBCH stage before véraison) was used (time points 2014: DAF0-35; 2017:
732 DAF0-42). Furthermore, a set of véraison specific genes was defined from the the DEGs
733 detected at the intersection of development stages BBCH79 (one developmental BBCH stage
734 before véraison) to BBCH81 (begin of ripening / véraison; time points 2014: DAF35-41; 2017:
735 DAF42-49). To test for biological relevance of the subsets, all DEGs were screened to their
736 occurrence in similar relevant studies [25, 26, 28, 29, 33, 41, 78, 79].
737
738 **Visualization of gene numbers newly appearing as differentially expressed**
739 To visualize appearance of DEGs over time, a stacked bar plot script was set up using the R
740 package 'plotly' (v. 4.9.2.1) [80]. Each bar represents the amount of DEGs of a given time
741 point or condition. In order to track groups of DEGs newly appearing at a given time point

742 throughout the following time points, the colorshade representing the group of DEGs
743 remains the same.
744
745 **Pathway enrichment analysis**
746 To search for possible targets in known pathways of grapevine, a pathway enrichment
747 analysis using the tool VitisPathways [81] was performed. To achieve a reliable enrichment,
748 1000 permutations, a Fisher's exact test of $p < 0.05$ and permuted p-value < 0.1 were set.
749 Thus, all significant enriched pathway genes and their relations can be displayed in VitisNet
750 [82], a specific molecular network for grapevine.
751
752 **Heatmaps**
753 As an extension to assignment of genes to biosynthesis pathways, the genes were also
754 filtered for annotation as coding for transcription factors (TFs). This filter was based on the
755 annotation information transferred from Mercator and RefSeq (see above). To look at the
756 entire family of *R2R3-MYB* TF genes, the list of *MYB* genes identified via MapMan was
757 extended by additional grapevine *R2R3-MYB* gene family members that have been
758 characterized [6, 7]. The *R2R3-MYB* genes detected among the intersecting DEGs were
759 displayed in heatmaps addressing the four individual time series (2 cultivars, 2 years) using
760 the R package 'pheatmap' (v. 2.1.3) [83]. Predictions for phylogenetic relationships were
761 deduced from PhyloGenes v. 2.2 [84].
762
763 **Abbreviations**
764 ABA abscisic acid
765 DAF days after begin of flowering

766 DEGs differentially expressed genes

767 PNP Pinot Noir Precoce

768 PN Pinot Noir

769 LFC log2fold change

770 PADJ adjusted p-value

771 TF transcription factor

772 TPM transcripts per kilobase million

773

774 **Declarations**

775 **Ethics approval and consent to participate**

776 Not applicable

777

778 **Consent for publication**

779 Not applicable

780

781 **Availability of data and material**

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

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

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

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

786 this study are available on GitHub [https://github.com/bpucker;

787 https://github.com/jenthein].

788

789 **Competing interests**

The authors declare that they have no competing interests

Funding

This project was funded by BMBF/PtJ (FKZ 031A3496).

Authors' contributions

Sampling and phenotyping were done at the JKI (Julius Kühn-Institute, Institute for Grapevine Breeding Geilweilerhof, Siebeldingen, Germany). Sequencing and data analysis were performed at Bielefeld University, Center for Biotechnology (CeBiTec). Organization, supervision at Bielefeld University and parts of sampling were done by DH. KH, AK and FS coordinate weekly sampling, phenotyping, and image capture for documentation. Logistic work was done by KH. The principle design of the experiment was designed by FS, DH and KH. RNA isolation and cDNA synthesis were carried out by DH and PV. PV accomplished library preparation and sequencing. RT and KH supervised the work at JKI Geilweilerhof. LH performed the SSR analysis. BW supervised the work at Bielefeld University. FS, KH, RT and BW acquired project funding and wrote the project proposal. All bioinformatic data analyses, creation of figures, tables and writing of the manuscript were performed by JT with the help of DH.

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

814 to the Bioinformatics Resource Facility support team of the CeBiTec for providing computing
815 infrastructure and excellent technical support. This publication is based in part upon
816 contributions from COST Action CA 17111 INTEGRAPE, supported by COST (European
817 Cooperation in Science and Technology); the authors thank BMBF/PtJ for funding in the
818 context of IPAS (project acronym novisys). The authors from UniBi also wish to thank the
819 other members of the Chair of Genetics and Genomics of Plants for their support. We would
820 also like to thank the Open Access Publication Funds of Bielefeld University for funding the
821 publication charge.

822

823 **Supplementary material**

824 **Additional file 1: Table S1**: Clonal relation of PN and PNP confirmed by 24 SSR markers
825 **Additional file 1: Table S2** - Sampling and phenotypical observations (BBCH) for 2014
826 **Additional file 1: Table S3** - Sampling and phenotypical observations (BBCH) for 2017
827 **Additional file 1: Table S4** - Pearsonr correlations of gene expression values over time
828 between both years
829 **Additional file 1: Table S5** - Top 500 genes influencing principal component PC1
830 **Additional file 1: Table S6** - Top 500 genes influencing principal component PC2
831 **Additional file 1: Table S7** - Cluster memberships of clusters obtained from the
832 clusteranalysis PN/PNP
833 **Additional file 1: Table S8** - Cluster memberships of clusters obtained from the cluster
834 analysis PN
835 **Additional file 1: Table S9** - Cluster memberships of clusters obtained from the cluster
836 analysis PNP
837 **Additional file 1: Table S10** - GO term enrichment for the clusters from the cluster analysis
838 of PN/PNP
839 **Additional file 1: Table S11** - GO term enrichment for the clusters from the cluster analysis
840 of PN
841 **Additional file 1: Table S12** - GO term enrichment for the clusters from the cluster analysis
842 of PNP

843 **Additional file 1: Table S13** - Differentially expressed genes between PN and PNP in both
844 years and filtering steps
845 **Additional file 1: Table S14** - Overlap of 12 ripening regulatory DEGs (time points: DAF0-35
846 from 2014 intersecting with DAF0-42 from 2017) and 388 véraison-specific DEGs (time
847 points: DAF 35+41 from 2014 intersecting with DAF42+49 from 2017) with previous
848 transcriptomic grapevine studies.
849 **Additional file 1: Table S15** - VitisNet enrichment for intersecting DEGs
850 **Additional file 1: Table S16** - VitisNet enrichment for regulatory DEGs
851 **Additional file 1: Table S17** - VitisNet enrichment for véraison specific DEGs
852 **Additional file 1: Table S18** - Up- & Down-regulated TFs among the DEGs
853 **Additional file 1: Table S19** - ENA sample identifier and metadata linked to the study for PNP
854 RNA-Seq reads from 2014 (PRJEB39262)
855 **Additional file 1: Table S20** - ENA sample identifier and metadata linked to the study for PN
856 RNA-Seq reads from 2014 (PRJEB39261)
857 **Additional file 1: Table S21** - ENA sample identifier and metadata linked to the study for PNP
858 RNA-Seq reads from 2017 (PRJEB39264)
859 **Additional file 1: Table S22** - ENA sample identifier and metadata linked to the study for PN
860 RNA-Seq reads from 2017 (PRJEB39263)
861
862 **Additional file 2: Figure S1** - Cluster analyses of PN/PNP in both years (A), PN in both years
863 (B) and PNP in both years (C)
864 **Additional file 2: Figure S2** - Heatmap of all MYB-TFs gene expression among the DEGs for
865 PN and PNP in both years
866 **Additional file 2: Figure S3** - Expression patterns of genes from the ripening-regulatory gene
867 set
868
869

870 **References**

871

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

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

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

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

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

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

888 7. Wong DCJ, Schlechter R, Vannozzi A, Holl J, Hmmam I, Bogs J, Tornielli GB, Castellarin SD,
889 Matus JT: **A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family**
890 **uncovers new insights into the regulation of stilbene accumulation**. *DNA Research* 2016,
891 **23**(5):451-466.

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

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

898 10. Osorio S, Scossa F, Fernie AR: **Molecular regulation of fruit ripening**. *Frontiers in Plant Science*
899 2013, **4**:198.

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

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

905 13. Zhu Z, Li G, Yan C, Liu L, Zhang Q, Han Z, Li B: **DRL1, Encoding A NAC Transcription Factor, Is**
906 **Involved in Leaf Senescence in Grapevine.** *International Journal of Molecular Sciences* 2019,
907 **20**(11):2678.

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

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

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

918 17. Böttcher C, Burbidge CA, Boss PK, Davies C: **Interactions between ethylene and auxin are**
919 **crucial to the control of grape (Vitis vinifera L.) berry ripening**. *BMC Plant Biology* 2013,
920 **13**:222.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

962 33. Massonnet M, Fasoli M, Tornielli GB, Altieri M, Sandri M, Zuccolotto P, Paci P, Gardiman M,
963 Zenoni S, Pezzotti M: **Ripening Transcriptomic Program in Red and White Grapevine Varieties**
964 **Correlates with Berry Skin Anthocyanin Accumulation.** *Plant Physiology* 2017, **174**(4):2376-
965 2396.

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

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

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

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

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

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

981 40. Zyprian E, Ochßner I, Schwander F, Šimon S, Hausmann L, Bonow-Rex M, Moreno-Sanz P,
982 Grando MS, Wiedemann-Merdinoglu S, Merdinoglu D *et al*: **Quantitative trait loci affecting**
983 **pathogen resistance and ripening of grapevines.** *Molecular Genetics and Genomics* 2016,
984 **291**(4):1573-1594.

985 41. Palumbo MC, Zenoni S, Fasoli M, Massonnet M, Farina L, Castiglione F, Pezzotti M, Paci P:
986 **Integrated network analysis identifies fight-club nodes as a class of hubs encompassing key**
987 **putative switch genes that induce major transcriptome reprogramming during grapevine**
988 **development.** *The Plant Cell* 2014, **26**(12):4617-4635.

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

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

995 44. Qi T, Huang H, Song S, Xie D: **Regulation of Jasmonate-Mediated Stamen Development and**
996 **Seed Production by a bHLH-MYB Complex in Arabidopsis.** *The Plant Cell* 2015, **27**(6):1620-
997 1633.

998 45. Czemmel S, Stracke R, Weisshaar B, Cordon N, Harris NN, Walker AR, Robinson SP, Bogs J: **The**
999 **grapevine R2R3-MYB transcription factor VvMYBF1 regulates flavonol synthesis in**
1000 **developing grape berries.** *Plant Physiology* 2009, **151**(3):1513-1530.

1001 46. Bogs J, Jaffe FW, Takos AM, Walker AR, Robinson SP: **The grapevine transcription factor**
1002 **VvMYBPA1 regulates proanthocyanidin synthesis during fruit development**. *Plant Physiology*
1003 2007, **143**(3):1347-1361.

1004 47. Koyama K, Numata M, Nakajima I, Goto-Yamamoto N, Matsumura H, Tanaka N: **Functional**
1005 **characterization of a new grapevine MYB transcription factor and regulation of**
1006 **proanthocyanidin biosynthesis in grapes**. *Journal of Experimental Botany* 2014, **65**(15):4433-
1007 4449.

1008 48. Cavallini E, Matus JT, Finezzo L, Zenoni S, Loyola R, Guzzo F, Schlechter R, Ageorges A, Arce-
1009 Johnson P, Tornielli GB: **The phenylpropanoid pathway is controlled at different branches by**
1010 **a set of R2R3-MYB C2 repressors in grapevine.** *Plant Physiology* 2015, **167**(4):1448-1470.

1011 49. Matus JT, Cavallini E, Loyola R, Holl J, Finezzo L, Dal Santo S, Vialet S, Commisso M, Roman F,
1012 Schubert A *et al*: **A group of grapevine MYBA transcription factors located in chromosome 14**
1013 **control anthocyanin synthesis in vegetative organs with different specificities compared with**
1014 **the berry color locus**. *The Plant Journal* 2017, **91**(2):220-236.

1015 50. Vannozzi A, Wong DCJ, Holl J, Hmmam I, Matus JT, Bogs J, Ziegler T, Dry I, Barcaccia G, Lucchin
1016 M: **Combinatorial Regulation of Stilbene Synthase Genes by WRKY and MYB Transcription**
1017 **Factors in Grapevine (Vitis vinifera L.)**. *Plant and Cell Physiology* 2018, **59**(5):1043-1059.

1018 51. Forlani S, Masiero S, Mizzotti C: **Fruit ripening: the role of hormones, cell wall modifications,**
1019 **and their relationship with pathogens**. *Journal of Experimental Botany* 2019, **70**(11):2993-
1020 3006.

1021 52. Dal Santo S, Vannozzi A, Tornielli GB, Fasoli M, Venturini L, Pezzotti M, Zenoni S: **Genome-wide**
1022 **analysis of the expansin gene superfamily reveals grapevine-specific structural and**
1023 **functional characteristics**. *PLoS One* 2013, **8**(4):e62206.

1024 53. Du M, Spalding EP, Gray WM: **Rapid Auxin-Mediated Cell Expansion**. *Annual Review of Plant*
1025 *Biology* 2020, **71**:379-402.

1026 54. Guillaumie S, Fouquet R, Kappel C, Camps C, Terrier N, Moncomble D, Dunlevy JD, Davies C,
1027 Boss PK, Delrot S: **Transcriptional analysis of late ripening stages of grapevine berry**. *BMC*
1028 *Plant Biology* 2011, **11**:165.

1029 55. Li Z, Zhang C, Guo Y, Niu W, Wang Y, Xu Y: **Evolution and expression analysis reveal the**
1030 **potential role of the HD-Zip gene family in regulation of embryo abortion in grapes (Vitis**
1031 **vinifera L.)**. *BMC Genomics* 2017, **18**(1):744.

1032 56. Cui M, Wang W, Guo F, Fan X, Guan L, Zheng T, Zhu X, Jia H, Fang J, Wang C *et al*:
1033 **Characterization and temporal–spatial expression analysis of LEC1 gene in the development**
1034 **of seedless berries in grape induced by gibberellin**. *Plant Growth Regulation* 2020, **90**(3):585-
1035 596.

1036 57. Jo L, Pelletier JM, Harada JJ: **Central role of the LEAFY COTYLEDON1 transcription factor in**
1037 **seed development**. *Journal of Integrative Plant Biology* 2019, **61**(5):564-580.

1038 58. Liu Q, Sun Y, Su W, Yang J, Liu X, Wang Y, Wang F, Li H, Li X: **Species-specific size expansion and**
1039 **molecular evolution of the oleosins in angiosperms**. *Gene* 2012, **509**(2):247-257.

1040 59. Qian X, Liu Y, Zhang G, Yan A, Wang H, Wang X, Pan Q, Xu H, Sun L, Zhu B: **Alcohol**
1041 **acyltransferase gene and ester precursors differentiate composition of volatile esters in**
1042 **three interspecific hybrids of Vitis labruscaxV. Vinifera during berry development period**.
1043 *Food Chemistry* 2019, **295**:234-246.

1044 60. Bourdais G, Burdiak P, Gauthier A, Nitsch L, Salojarvi J, Rayapuram C, Idanheimo N, Hunter K,
1045 Kimura S, Merilo E *et al*: **Large-Scale Phenomics Identifies Primary and Fine-Tuning Roles for**
1046 **CRKs in Responses Related to Oxidative Stress**. *PLoS Genetics* 2015, **11**(7):e1005373.

1047 61. Dallakyan M, Zyprian E, Maul E, Yesayan A, Hovhannisyan N: **Genetic characterization of rape**
1048 **varieties in Armenia**. *Vitis* 2015, **54**:23-26.

1049 62. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data.**
1050 *Bioinformatics* 2014, **30**(15):2114-2120.

1051 63. **FastQC, a quality control tool for high throughput sequence data.**
1052 [http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]

1053 64. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N,
1054 Jubin C *et al*: **The grapevine genome sequence suggests ancestral hexaploidization in major**
1055 **angiosperm phyla**. *Nature* 2007, **449**(7161):463-467.

1056 65. Kim D, Langmead B, Salzberg SL: **HISAT: a fast spliced aligner with low memory requirements**.
1057 *Nature Methods* 2015, **12**(4):357-360.

1058 66. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL: **Graph-based genome alignment and**
1059 **genotyping with HISAT2 and HISAT-genotype**. *Nature Biotechnology* 2019, **37**(8):907-915.

1060 67. Liao Y, Smyth GK, Shi W: **featureCounts: an efficient general purpose program for assigning**
1061 **sequence reads to genomic features.** *Bioinformatics* 2014, **30**(7):923-930.

1062 68. Zhao S, Ye Z, Stanton R: **Misuse of RPKM or TPM normalization when comparing across**
1063 **samples and sequencing protocols**. *RNA* 2020, **26**(8):903-909.

1064 69. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E,
1065 Peterson P, Weckesser W, Bright J *et al*: **SciPy 1.0: fundamental algorithms for scientific**
1066 **computing in Python**. *Nature Methods* 2020, **17**(3):261-272.

1067 70. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-**
1068 **seq data with DESeq2.** *Genome Biology* 2014, **15**(12):550.

1069 71. **R: A language and environment for statistical computing** [https://www.R-project.org/]

1070 72. Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M:
1071 **MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic**
1072 **pathways and other biological processes.** *The Plant Journal* 2004, **37**(6):914-939.

1073 73. Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie AR, Stitt M, Usadel
1074 B: **Mercator: a fast and simple web server for genome scale functional annotation of plant**
1075 **sequence data**. *Plant, Cell & Environment* 2014, **37**(5):1250-1258.

1076 74. Pruitt KD, Tatusova T, Maglott DR: **NCBI reference sequences (RefSeq): a curated non-**
1077 **redundant sequence database of genomes, transcripts and proteins.** *Nucleic Acids Research*
1078 2007, **35**(Database issue):D61-65.

1079 75. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *Journal*
1080 *of Molecular Biology* 1990, **215**(3):403-410.

1081 76. Alexa A, Rahnenfuhrer J, Lengauer T: **Improved scoring of functional groups from gene**
1082 **expression data by decorrelating GO graph structure**. *Bioinformatics* 2006, **22**(13):1600-1607.

1083 77. Abu-Jamous B, Kelly S: **Clust: automatic extraction of optimal co-expressed gene clusters**
1084 **from gene expression data.** *Genome Biology* 2018, **19**:172.

1085 78. Zenoni S, D'Incà E, Tornielli GB: **Genetic dissection of grape berry ripening control: defining a**
1086 **role for NAC transcription factors**. *Acta Horticulturae* 2019(1248):387-402.

1087 79. Wei L, Cao Y, Cheng J, Xiang J, Shen B, Wu J: **Comparative transcriptome analyses of a table**
1088 **grape 'Summer Black' and its early-ripening mutant 'Tiangong Moyu' identify candidate**
1089 **genes potentially involved in berry development and ripening**. *Journal of Plant Interactions*
1090 2020, **15**(1):213-222.

1091 80. Sievert C: **Interactive Web-Based Data Visualization with R, plotly, and shiny**: CRC Press;
1092 2020.

1093 81. Osier MV: **VitisPathway: gene pathway analysis for V. vinifera**. *Vitis* 2016, **55**:129-133.

1094 82. Grimplet J, Van Hemert J, Carbonell-Bejerano P, Díaz-Riquelme J, Dickerson J, Fennell A,
1095 Pezzotti M, Martínez-Zapater JM: **Comparative analysis of grapevine whole-genome gene**
1096 **predictions, functional annotation, categorization and integration of the predicted gene**
1097 **sequences.** *BMC Research Notes* 2012, **5**:213.

1098 83. **pheatmap: Pretty Heatmaps. R package version 1.0.12** [https://CRAN.R-
1099 project.org/package=pheatmap]

1100 84. Zhang P, Berardini TZ, Ebert D, Li Q, Mi H, Muruganujan A, Prithvi T, Reiser L, Sawant S, Thomas
1101 PD *et al*: **PhyloGenes: An online phylogenetics and functional genomics resource for plant**
1102 **gene function inference**. *Plant Direct* 2020, **4**(12):e00293.

1103