

# 1 From bud formation to flowering: transcriptomic state defines the cherry developmental phases 2 of sweet cherry bud dormancy

3

4 Noémie Vimont<sup>1,2,3</sup>, Mathieu Fouché<sup>1</sup>, José Antonio Campoy<sup>4,5,6</sup>, Meixuezi Tong<sup>3</sup>, Mustapha Arkoun<sup>2</sup>,  
5 Jean-Claude Yvin<sup>2</sup>, Philip A. Wigge<sup>7</sup>, Elisabeth Dirlewanger<sup>1</sup>, Sandra Cortijo<sup>3#</sup>, Bénédicte Wenden<sup>1#</sup>

6

7 <sup>1</sup>UMR 1332 BFP, INRA, Univ. Bordeaux, 33882 Villenave d'Ornon, Cedex France; <sup>2</sup>Agro Innovation International - Centre Mondial  
8 d'Innovation - Groupe Roullier, 35400 St Malo, France; <sup>3</sup>The Sainsbury Laboratory, University of Cambridge, Cambridge CB2 1LR,  
9 United Kingdom; <sup>4</sup>Universidad Politécnica de Cartagena, Cartagena, Spain; <sup>5</sup>Universidad de Murcia, Murcia, Spain; <sup>6</sup>Current address:  
10 Department of Plant Developmental Biology, Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany; <sup>7</sup>Leibniz-  
11 Institute für Gemüse- und Zierpflanzenbau (IGZ), Plant Adaptation, Grossbeeren, Germany

12 #Corresponding authors: [sandra.cortijo@slcu.cam.ac.uk](mailto:sandra.cortijo@slcu.cam.ac.uk); [benedicte.wenden@inra.fr](mailto:benedicte.wenden@inra.fr)

13

14 [vimont.noemie@hotmail.fr](mailto:vimont.noemie@hotmail.fr); [mathieu.fouche@inra.fr](mailto:mathieu.fouche@inra.fr); [campoy@mpipz.mpg.de](mailto:campoy@mpipz.mpg.de); [meixuezi.tong@gmail.com](mailto:meixuezi.tong@gmail.com);

15 [mustapha.arkoun@roullier.com](mailto:mustapha.arkoun@roullier.com); [JeanClaude.Yvin@roullier.com](mailto:JeanClaude.Yvin@roullier.com); [wigge@igzev.de](mailto:wigge@igzev.de); [Elisabeth.dirlewanger@inra.fr](mailto:Elisabeth.dirlewanger@inra.fr);

16 [Sandra.cortijo@slcu.cam.ac.uk](mailto:Sandra.cortijo@slcu.cam.ac.uk); [benedicte.wenden@inra.fr](mailto:benedicte.wenden@inra.fr);

17

## 18 ABSTRACT

19 **Background.** Bud dormancy is a crucial stage in perennial trees and allows survival over winter to  
20 ensure optimal flowering and fruit production. Recent work highlighted physiological and molecular  
21 events occurring during bud dormancy in trees. However, they usually examined bud development or  
22 bud dormancy in isolation. In this work, we aimed to further explore the global transcriptional changes  
23 happening throughout bud development and dormancy onset, progression and release.

24 **Results.** Using next-generation sequencing and modelling, we conducted an in-depth transcriptomic  
25 analysis for all stages of flower buds in several sweet cherry (*Prunus avium* L.) cultivars that are  
26 characterized for their contrasted dates of dormancy release. We find that buds in organogenesis,  
27 paradormancy, endodormancy and ecodormancy stages are defined by the expression of genes  
28 involved in specific pathways, and these are conserved between different sweet cherry cultivars. In

29 particular, we found that *DORMANCY ASSOCIATED MADS-box (DAM)*, floral identity and  
30 organogenesis genes are up-regulated during the pre-dormancy stages while endodormancy is  
31 characterized by a complex array of signalling pathways, including cold response genes, ABA and  
32 oxidation-reduction processes. After dormancy release, genes associated with global cell activity,  
33 division and differentiation are activated during ecodormancy and growth resumption. We then went  
34 a step beyond the global transcriptomic analysis and we developed a model based on the transcriptional  
35 profiles of just seven genes to accurately predict the main bud dormancy stages.

36 **Conclusions.** Overall, this study has allowed us to better understand the transcriptional changes  
37 occurring throughout the different phases of flower bud development, from bud formation in the  
38 summer to flowering in the following spring. Our work sets the stage for the development of fast and  
39 cost effective diagnostic tools to molecularly define the dormancy stages. Such integrative approaches  
40 will therefore be extremely useful for a better comprehension of complex phenological processes in  
41 many species.

42

43 **KEY WORDS:** Transcriptomic, RNA sequencing, time course, *Prunus avium* L., prediction, seasonal  
44 timing

45

## 46 **BACKGROUND**

47 Temperate trees face a wide range of environmental conditions including highly contrasted  
48 seasonal changes. Among the strategies to enhance survival under unfavourable climatic conditions,  
49 bud dormancy is crucial for perennial plants since its progression over winter is determinant for  
50 optimal growth, flowering and fruit production during the subsequent season. Bud dormancy has long  
51 been compared to an unresponsive physiological phase, in which metabolic processes within the buds  
52 are halted by cold temperature and/or short photoperiod. However, several studies have shown that  
53 bud dormancy progression can be affected in a complex way by temperature, photoperiod or both,

54 depending on the tree species [1–5]. Bud dormancy has traditionally been separated into three main  
55 phases: (i) paradormancy, also named “summer dormancy” [6]; (ii) endodormancy, mostly triggered  
56 by internal factors; and (iii) ecodormancy, controlled by external factors [7, 8]. Progression through  
57 endodormancy requires cold accumulation whereas warmer temperatures, i.e. heat accumulation, drive  
58 the competence to resume growth over the ecodormancy phase. Dormancy is thus highly dependent  
59 on external temperatures, and changes in seasonal timing of bud break and blooming have been  
60 reported in relation with global warming. Notably, advances in bud break and blooming dates in spring  
61 have been observed for tree species, such as apple, cherry, birch, oak or Norway spruce, in the northern  
62 hemisphere, thus increasing the risk of late frost damages [9–14] while insufficient cold accumulation  
63 during winter may lead to incomplete dormancy release associated with bud break delay and low bud  
64 break rate [15, 16]. These phenological changes directly impact the production of fruit crops, leading  
65 to large potential economic losses [17]. Consequently, it becomes urgent to acquire a better  
66 understanding of bud responses to temperature stimuli in the context of climate change in order to  
67 tackle fruit losses and anticipate future production changes.

68 In the recent years, an increasing number of studies have investigated the physiological and  
69 molecular mechanisms of bud dormancy transitions in perennials using RNA sequencing technology,  
70 thereby giving a new insight into potential pathways involved in dormancy. The results suggest that  
71 the transitions between the three main bud dormancy phases (para-, endo- and eco- dormancy) are  
72 mediated by pathways related to *DORMANCY ASSOCIATED MADS-box (DAM)* genes [18],  
73 phytohormones [19–22], carbohydrates [22, 23], temperature [24, 25], photoperiod [26], reactive  
74 oxygen species [27, 28], water deprivation [26], cold acclimation and epigenetic regulation [29].  
75 Owing to these studies, a better understanding of bud dormancy has been established in different  
76 perennial species [18, 30, 31]. However, we are still missing a fine-resolution temporal understanding  
77 of transcriptomic changes happening over the entire bud development, from bud organogenesis to bud  
78 break.

79 Indeed, the small number of sampling dates in existing studies seems to be insufficient to capture all  
80 the information about changes occurring throughout the dormancy cycle as it most likely corresponds  
81 to a chain of biological events rather than an on/off mechanism. Many unresolved questions remain:  
82 What are the fine-resolution dynamics of gene expression related to dormancy? Are specific sets of  
83 genes associated with dormancy stages? Since the timing for the response to environmental cues is  
84 cultivar-dependant [32, 33], are transcriptomic profiles during dormancy different in cultivars with  
85 contrasted flowering date?

86 To explore these mechanisms, we conducted a transcriptomic analysis of sweet cherry (*Prunus*  
87 *avium* L.) flower buds from bud organogenesis until the end of bud dormancy using next-generation  
88 sequencing. Sweet cherry is a perennial species highly sensitive to temperature [34] and we focused  
89 on three sweet cherry cultivars displaying contrasted flowering dates and response to environmental  
90 conditions. We carried out a fine-resolution time-course spanning the entire bud development, from  
91 flower organogenesis in July to flowering in spring of the following year (February to April),  
92 encompassing para-, enco- and ecodormancy phases. Our results indicate that transcriptional changes  
93 happening during dormancy are conserved between different sweet cherry cultivars, opening the way  
94 to the identification of key factors involved in the progression through bud dormancy.

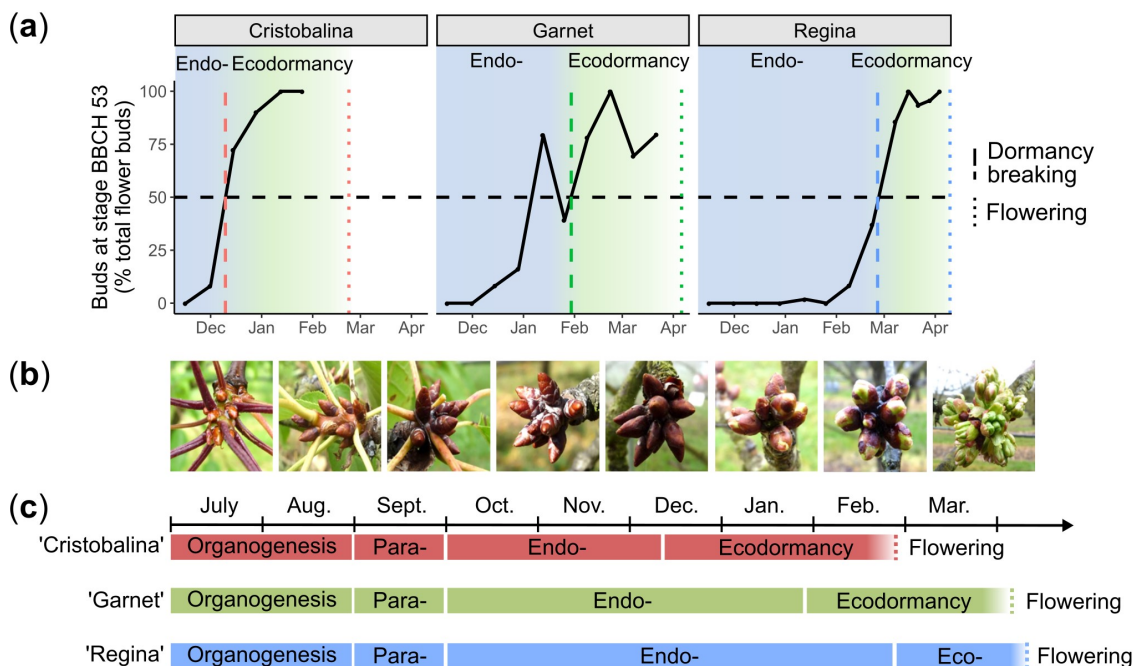
95

## 96 **RESULTS**

### 97 **Transcriptome accurately captures the dormancy state**

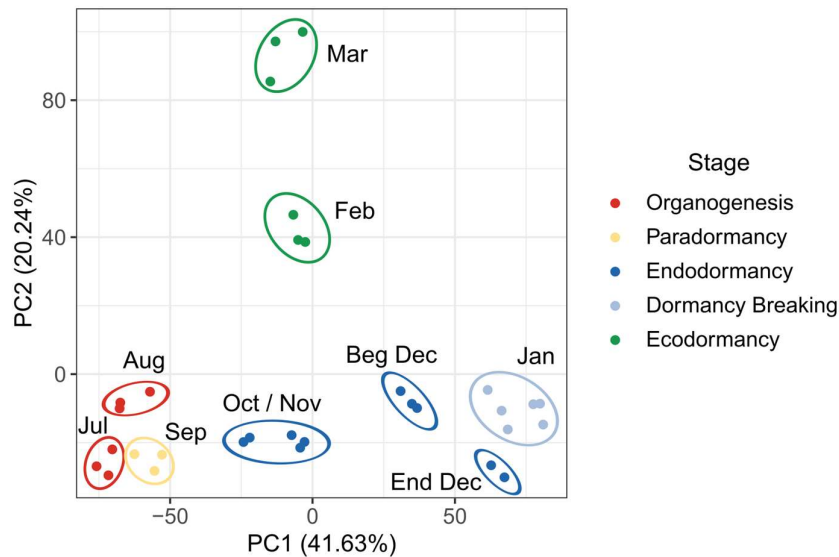
98 In order to define transcriptional changes happening over the sweet cherry flower bud  
99 development, we performed a transcriptomic-wide analysis using next-generation sequencing from  
100 bud organogenesis to flowering. According to bud break percentage (Fig. 1a), morphological  
101 observations (Fig. 1b), average temperatures (see Fig. S1 in Additional File 1) and descriptions from  
102 Lang *et al.*, (1987), we assigned five main stages to the flower buds samples (Fig. 1c): i) flower bud  
103 organogenesis occurs in July and August; ii) paradormancy corresponds to the period of growth

104 cessation in September; iii) during the endodormancy phase, initiated in October, buds are  
 105 unresponsive to forcing conditions therefore the increasing bud break percentage under forcing  
 106 conditions suggests that endodormancy was released on 9<sup>th</sup> December 2015, 29<sup>th</sup> January 2016, and  
 107 26<sup>th</sup> February 2016 for the three cultivars ‘Cristobalina’, ‘Garnet’ and ‘Regina’, respectively, thus  
 108 corresponding to iv) dormancy breaking; and v) ecodormancy starting from the estimated dormancy  
 109 release date until flowering. We harvested buds at eleven dates spanning all these bud stages for the  
 110 sweet cherry cultivars ‘Cristobalina’, ‘Garnet’ and ‘Regina’, and generated a total of 81 transcriptomes  
 111 (Table S1 in Additional File 2). First, in order to explore the transcriptomic characteristics of each bud  
 112 stage separately from the cultivar effect, we focused the analysis on the early flowering cultivar  
 113 ‘Garnet’.  
 114



**Figure 1. Dormancy status under environmental conditions and RNA-seq sampling dates**

(a) Evaluation of bud break percentage under forcing conditions was carried out for three sweet cherry cultivars displaying different flowering dates in ‘Cristobalina’, ‘Garnet’ and ‘Regina’ for the early, medium and late cultivar, respectively. The dash and dotted lines correspond to the dormancy release date, estimated at 50% of buds at BBCH stage 53 [90], and the flowering date, respectively. (b) Pictures of the sweet cherry buds corresponding to the different sampling dates. (c) Sampling time points for the transcriptomic analysis are represented by coloured stars. Red for ‘Cristobalina’, green for ‘Garnet’ and blue for ‘Regina’.

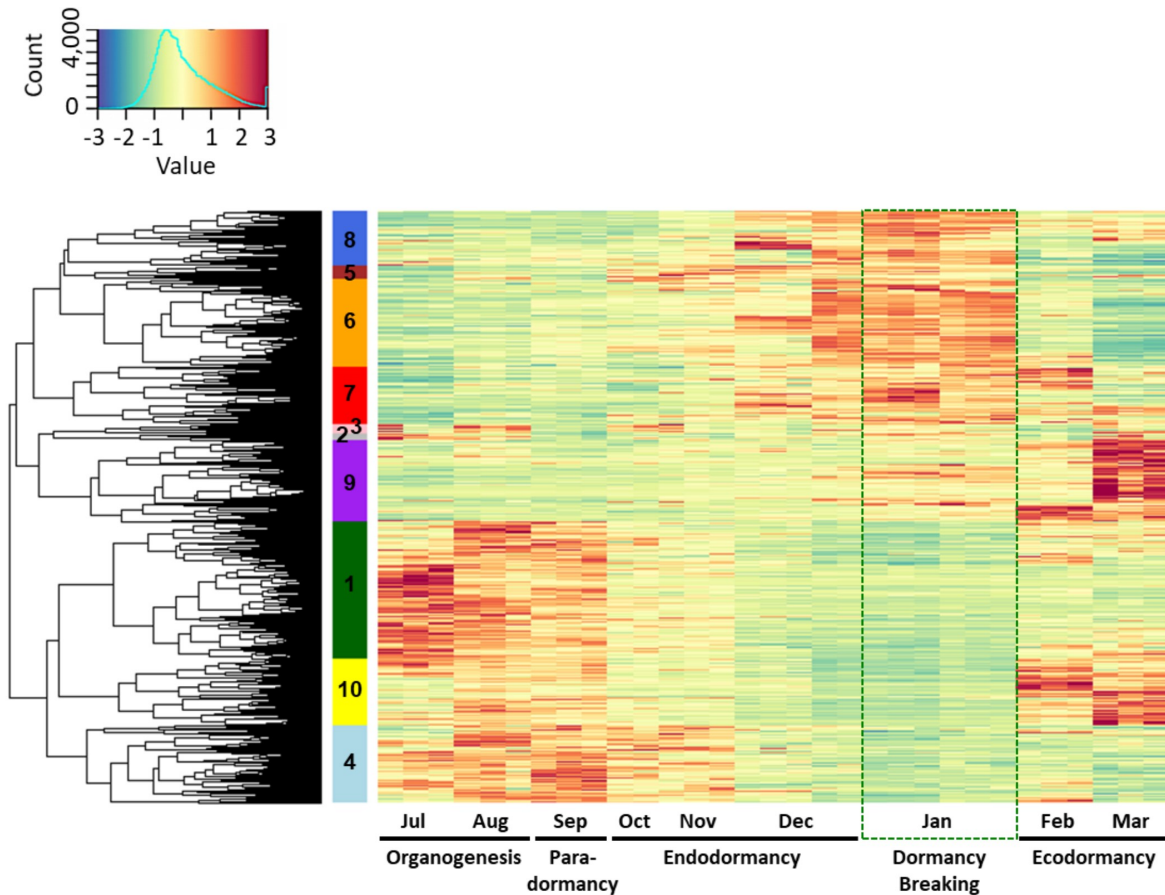


**Fig 2 Separation of samples by dormancy stage using differentially expressed genes**

The principal component analysis was conducted on the TPM (transcripts per millions reads) values for the differentially expressed genes in the cultivar ‘Garnet’ flower buds, sampled on three trees between July and March.

115 Using DESeq2 and a threshold of 0.05 on the adjusted p-value, we identified 6,683 genes that  
116 are differentially expressed (DEGs) between the defined bud stages for the sweet cherry cultivar  
117 ‘Garnet’ (Table S2 in Additional File 2). When projected into a two-dimensional space (Principal  
118 Component Analysis, PCA), data for these DEGs show that transcriptomes of samples harvested at a  
119 given date are projected together (Fig. 2), showing the high quality of the biological replicates and that  
120 different trees are in a very similar transcriptional state at the same date. Very interestingly, we also  
121 observe that flower bud states are clearly separated on the PCA, with the exception of organogenesis  
122 and paradormancy, which are projected together (Fig. 2). The first dimension of the analysis (PC1)  
123 explains 41,63% of the variance and clearly represents the strength of bud dormancy where samples  
124 on the right of the axis are in late endodormancy (Dec) or dormancy breaking stages, while samples  
125 on the left of the axis are in organogenesis and paradormancy. Samples harvested at the beginning of  
126 the endodormancy (Oct and Nov) are mid-way between samples in paradormancy and in late  
127 endodormancy (Dec) on PC1. The second dimension of the analysis (PC2) explains 20.24% of the  
128 variance and distinguishes two main phases of the bud development: before and after dormancy  
129 breaking. We obtain very similar results when performing the PCA on all genes (Fig. S2 in Additional

130 File 1). These results indicate that the transcriptional state of DEGs accurately captures the dormancy  
131 state of flower buds.



**Figure 3. Clusters of expression patterns for differentially expressed genes in the sweet cherry cultivar 'Garnet'**

Heatmap for 'Garnet' differentially expressed genes during bud development. Each column corresponds to the gene expression for flower buds from one single tree at a given date. Clusters are ordered based on the chronology of the expression peak (from earliest – July, 1-dark green cluster – to latest – March, 9 and 10). Expression values were normalized and *z-scores* are represented here.

132 **Bud stage-dependent transcriptional activation and repression are associated with different**  
133 **pathways**

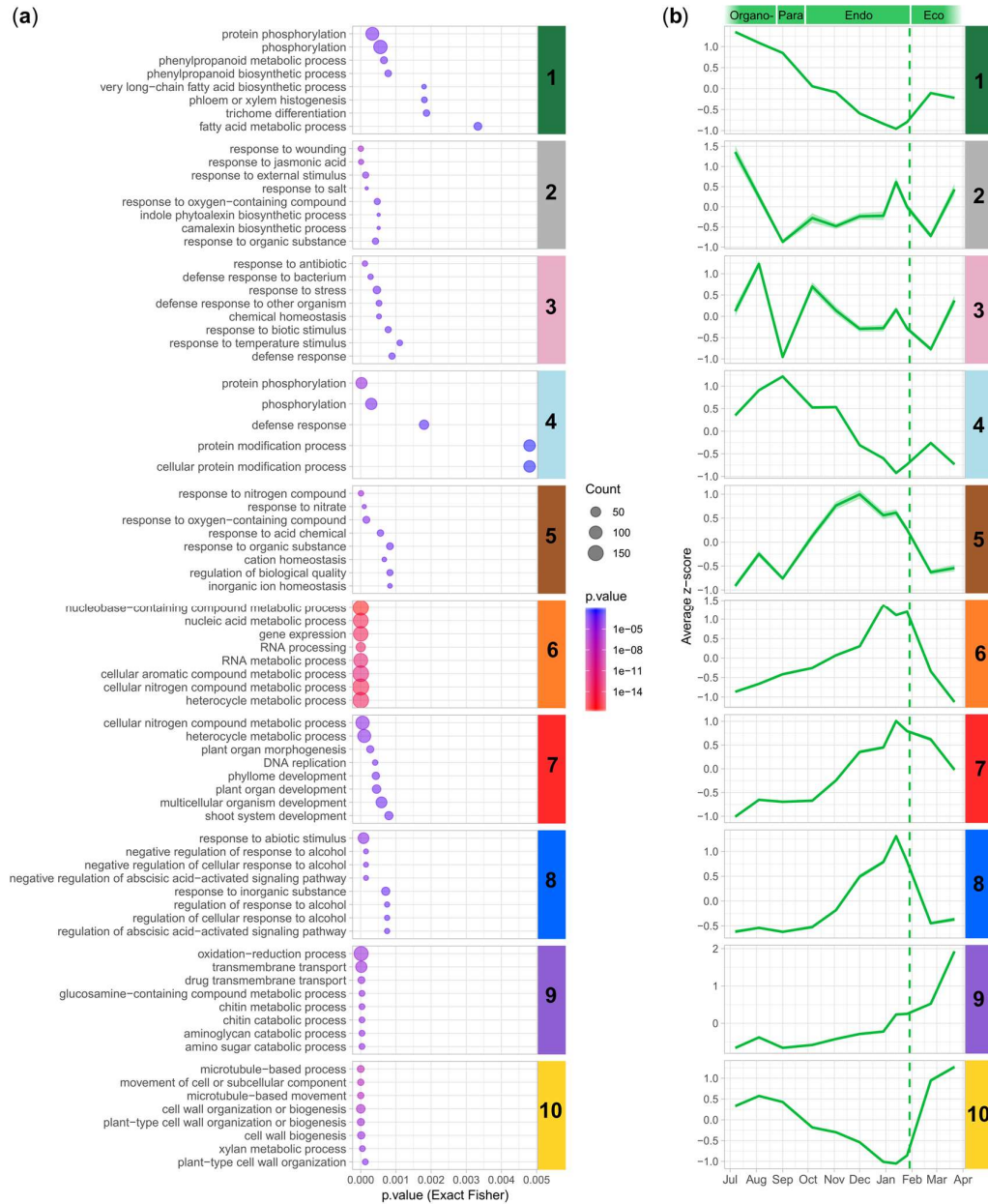
134 We further investigated whether specific genes or signalling pathways could be associated with  
135 the different flower bud stages. For this, we performed a hierarchical clustering of the DEGs based on  
136 their expression in all samples. We could group the genes in ten clusters clearly showing distinct  
137 expression profiles throughout the bud development (Fig. 3). Overall, three main types of clusters can

138 be discriminated: the ones with a maximum expression level during organogenesis and paradormancy  
139 (cluster 1: 1,549 genes; cluster 2: 70 genes; cluster 3: 113 genes; cluster 4: 884 genes and cluster 10:  
140 739 genes, Fig. 3), the clusters with a maximum expression level during endodormancy and around  
141 the time of dormancy breaking (cluster 5: 156 genes; cluster 6: 989 genes; cluster 7: 648 genes and  
142 cluster 8: 612 genes, Fig. 3), and the clusters with a maximum expression level during ecodormancy  
143 (cluster 9: 924 genes and cluster 10: 739 genes, Fig. 3). This result shows that different groups of genes  
144 are associated with these three main flower bud phases. Interestingly, we also observed that during the  
145 endodormancy phase, some genes are expressed in October and November then repressed in December  
146 (cluster 4, Fig. 3), whereas another group of genes is expressed in December (clusters 8, 5, 6 and 7,  
147 Fig. 3) therefore separating endodormancy in two periods with distinct transcriptional states, which  
148 supports the PCA observation.

149 In order to explore the functions and pathways associated with the gene clusters, we performed  
150 a GO enrichment analysis for each of the ten identified clusters (Fig. 4, Fig. S3). GO terms associated  
151 with the response to stress as well as biotic and abiotic stimuli were enriched in the clusters 2, 3 and  
152 4, with genes mainly expressed during organogenesis and paradormancy. In addition, we observed  
153 high expression of genes associated with floral identity before dormancy, including *AGAMOUS-*  
154 *LIKE20* (*PavAGL20*) and the bZIP transcription factor *PavFD* (Fig. 5). On the opposite, at the end of  
155 the endodormancy phase (cluster 6, 7 and 8), we highlighted different enrichments in GO terms linked  
156 to basic metabolisms such as nucleic acid metabolic processes or DNA replication but also to response  
157 to alcohol and abscisic acid (ABA). For example, *ABA BINDING FACTOR 2* (*PavABF2*),  
158 *ARABIDOPSIS THALIANA HOMEBOX 7* (*PavATHB7*) and ABA 8'-hydroxylase (*PavCYP707A2*),  
159 associated with the ABA pathway, as well as the stress-induced gene *PavHVA22*, were highly  
160 expressed during endodormancy (Fig. 5). During ecodormancy, genes in cluster 9 and 10 are enriched  
161 in functions associated with transport, cell wall biogenesis as well as oxidation-reduction processes  
162 (Fig. 4 and see Additional file 3). Indeed, we identified the *GLUTATHION S-TRANSFERASE8*



163 (*PavGST8*) gene and a peroxidase specifically activated during ecodormancy (Fig. 5). However,  
 164 oxidation-reduction processes are likely to occur during endodormancy as well, as suggested by the  
 165 expression patterns of *GLUTATHION PEROXIDASE 6* (*PavGPX6*) and *GLUTATHION REDUCTASE*  
 166 (*PavGR*). Interestingly, *AGAMOUS* (*PavAG*) and *APETALA3* (*PavAP3*) showed an expression peak

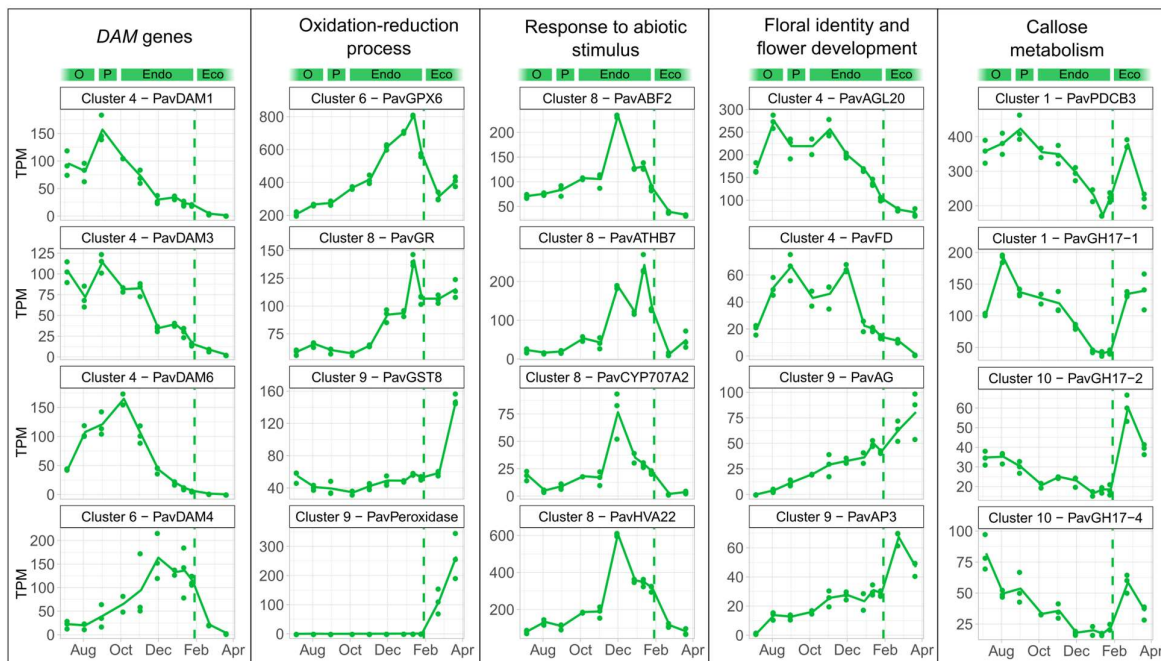


**Fig 4 Enrichments in gene ontology terms for biological processes and average expression patterns in the different clusters in the sweet cherry cultivar ‘Garnet’**

(a) Using the topGO package (Alexa & Rahnenführer, 2018), we performed an enrichment analysis on GO terms for biological processes based on a classic Fisher algorithm. Enriched GO terms with the lowest *p-value* were selected for representation. Dot size represent the number of genes belonging to the clusters associated with the GO term. (b) Average z-score values for each cluster. The coloured dotted line corresponds to the estimated date of dormancy release.

167 during ecodormancy (Fig. 5). These results show that different functions and pathways are specific to  
 168 flower bud development stages.

169 We further investigated whether dormancy-associated genes were specifically activated and  
 170 repressed during the different bud stages. Among the six annotated *DAM* genes, four were  
 171 differentially expressed in the dataset. *PavDAM1*, *PavDAM3* and *PavDAM6* were highly expressed  
 172 during paradormancy and at the beginning of endodormancy (cluster 4, Fig. 5) whereas the expression  
 173 peak for *PavDAM4* was observed at the end of endodormancy (cluster 6, Fig. 5). In addition, we found  
 174 that genes coding for 1,3- $\beta$ -glucanases from the Glycosyl hydrolase family 17 (*PavGH17*), as well as  
 175 a *PLASMODESMATA CALLOSE-BINDING PROTEIN 3* (*PavPDCB3*) gene were repressed during  
 176 dormancy (clusters 1 and 10, Fig. 5).



**Figure 5. Expression patterns of key genes involved in sweet cherry bud dormancy**  
 Expression patterns, expressed in transcripts per million reads (TPM) were analysed for the cultivar ‘Garnet’ from August to March, covering bud organogenesis (O), paradormancy (P), endodormancy (Endo), and ecodormancy (Eco). Dash lines represent the estimated date of dormancy breaking.

177

178 **Specific transcription factor target genes are expressed during the main flower bud stages**

179 To better understand the regulation of genes that are expressed at different flower bud stages,  
180 we investigated whether some transcription factors (TFs) targeted genes in specific clusters. Based on  
181 a list of predicted regulation between TFs and target genes that is available for peach in PlantTFDB  
182 [35], we identified the TFs with enriched targets in each cluster (Table 1). We further explored these  
183 target genes and their biological functions with a GO enrichment analysis (Tables S3, S4 in Additional  
184 file 2). Moreover, to have a complete overview of the TFs' targets, we also identified enriched target  
185 promoter motifs in the different gene clusters (Table S5 in Additional file 2), using motifs we  
186 discovered with Find Individual Motif occurrences (FIMO) [36] and reference motifs obtained from  
187 PlantTFDB 4.0 [35]. Results show that different pathways are activated throughout bud development.

188 Among the genes expressed during the organogenesis and paradormancy phases (clusters 1, 2,  
189 3 and 4), we observed an enrichment for motifs targeted by several MADS-box TFs such as  
190 AGAMOUS (AG), APETALA3 (AP3) and SEPALLATA3/AGAMOUS-like 9 (SEP3/AGL9), several  
191 of them potentially involved in flower organogenesis [37]. On the other hand, for the same clusters,  
192 results show an enrichment in MYB-related targets, WRKY and ethylene-responsive element (ERF)  
193 binding TFs (Table 1, Table S5 in Additional file 2). Several members of these TF families have been  
194 shown to participate in the response to abiotic factors. Similarly, we found in the cluster 4 target motifs  
195 enriched for DEHYDRATION RESPONSE ELEMENT-BINDING2 (PavDREB2C), potentially  
196 involved in the response to cold [38]. PavMYB63 and PavMYB93 transcription factors, expressed  
197 during organogenesis and paradormancy, likely activate genes involved in secondary metabolism  
198 (Table 1, Tables S3, S4 in Additional file 2).

199 During endodormancy, we found that PavMYB14 and PavMYB40 specifically target genes  
200 from cluster 10 that are involved in secondary metabolic processes and growth (Tables S3, S4 in  
201 Additional file 2). Expression profiles suggest that PavMYB14 and PavMYB40 repress expression of  
202 these target genes during endodormancy (Fig. S4 in Additional file 1). This is consistent with the  
203 functions of *Arabidopsis thaliana* MYB14 that negatively regulates the response to cold [39]. One of

204 the highlighted TFs was *PavWRKY40*, which is activated before endodormancy and preferentially  
205 regulates genes associated with oxidative stress (Fig. S4 and Table S4 in Additional files 1 and 2).

206 Interestingly, we observed a global response to cold and stress during endodormancy since we  
207 identified an enrichment of targets for PavCBF4, and of genes with motifs for several ethylene-  
208 responsive element binding TFs such as PavDREB2C in the cluster 5. We also observed an enrichment  
209 in the same cluster for genes with motifs for PavABI5 (Table S5 in Additional file 2). All these TFs  
210 are involved in the response to cold, in agreement with the fact that genes in the cluster 5 are expressed  
211 during endodormancy. Genes belonging to the clusters 6, 7 and 8 are highly expressed during deep  
212 dormancy and we found targets and target motifs for many TFs involved in the response to abiotic  
213 stresses. For example, we found motifs enriched in the cluster 7 for many TFs of the C2H2 family,  
214 which are involved in the response to a wide spectrum of stress conditions, such as extreme  
215 temperatures, salinity, drought or oxidative stress (Table S5; [40, 41]). Similarly, in the cluster 8, we  
216 also identified an enrichment in targets and motifs of many genes involved in the response to ABA  
217 and to abiotic stimulus, such as *PavABF2*, *PavAREB3*, *PavABI5*, *PavDREB2C* and *PavERF110*  
218 (Tables S3, S4 in Additional file 2) [38, 42]. Their targets include ABA-related genes *HIGHLY ABA-*  
219 *INDUCED PP2C GENE 1 (PavHAI1)*, *PavCYP707A2* that is involved in ABA catabolism, *PavPYL8*  
220 a component of ABA receptor 3 and *LATE EMBRYOGENESIS ABUNDANT PROTEIN (PavLEA)*,  
221 involved in the response to desiccation [4].

222 We also observe during endodormancy an enrichment for targets of TFs involved in the  
223 response to light and temperature, such as *PavPIL5*, *PavSPT*, *PavRVE1* and *PavPIF4* (Table 1, [5, 43–  
224 45]), and *PavRVE8* that preferentially target genes involved in cellular transport like *LIPID*  
225 *TRANSFER PROTEINI (PavLPI1)*, Table S3 in Additional file 2). Interestingly, we found that among  
226 the TFs with enriched targets in the clusters, only ten display changes in expression during flower bud  
227 development (Table 1, Fig. S4 in Additional file 1), including *PavABF2*, *PavABI5* and *PavRVE1*.  
228 Expression profiles for these three genes are very similar, and are also similar to their target genes,

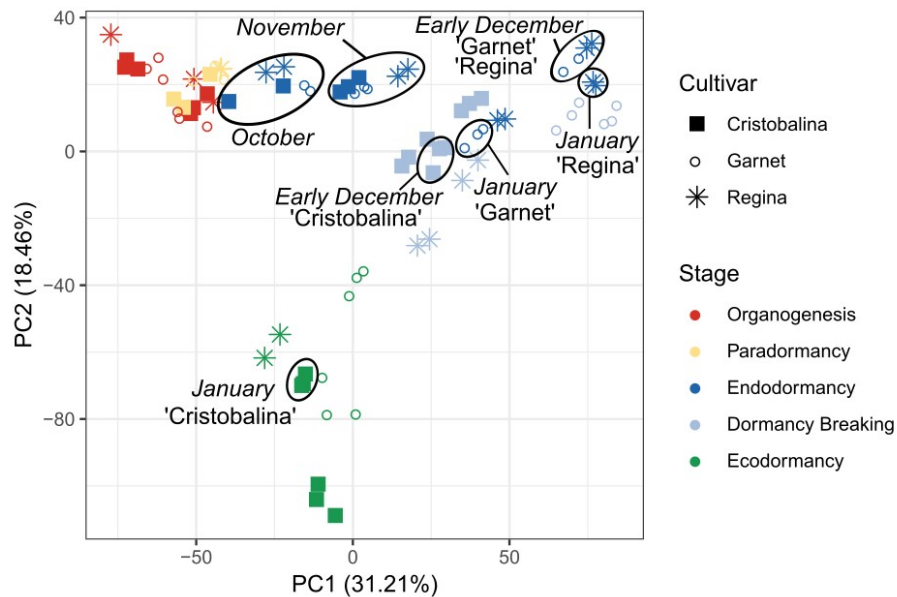
229 with a peak of expression around the estimated dormancy release date, indicating that these TFs are  
230 positively regulating their targets (see Fig. S4 in Additional file 1).

231

### 232 **Expression patterns highlight bud dormancy similarities and disparities between three cherry** 233 **tree cultivars**

234 Since temperature changes and progression through the flower bud stages are happening  
235 synchronously, it is challenging to discriminate transcriptional changes that are mainly associated with  
236 one or the other. In this context, we also analysed the transcriptome of two other sweet cherry cultivars:  
237 ‘Cristobalina’, characterized by very early flowering dates, and ‘Regina’, with a late flowering time.  
238 The span between flowering periods for the three cultivars is also found in the transition between  
239 endodormancy and ecodormancy since ten weeks separated the estimated dates of dormancy release  
240 between the cultivars: 9th December 2015 for ‘Cristobalina’, 29th January 2016 for ‘Garnet’ and 26th  
241 February 2016 for ‘Regina’ (Fig. 1a). The transition from organogenesis to paradormancy is not well  
242 documented and many studies suggest that endodormancy onset is under the strict control of  
243 environment in *Prunus* species [3]. Therefore, we considered that these two transitions occurred at the  
244 same time in all three cultivars. However, the two months and half difference in the date of transition  
245 from endodormancy to ecodormancy between the cultivars allow us to look for transcriptional changes  
246 associated with this transition independently of environmental conditions. To do so, we generated a  
247 total of 50 transcriptomes from buds harvested at ten dates for the cultivar ‘Cristobalina’, and eleven  
248 dates for the cultivar ‘Regina’, spanning all developmental stages from bud organogenesis to  
249 flowering. We then compared the expression patterns between the three contrasted cultivars throughout

250 flower bud stages for the genes we identified as differentially expressed in the cultivar ‘Garnet’ (Fig.  
251 1b).

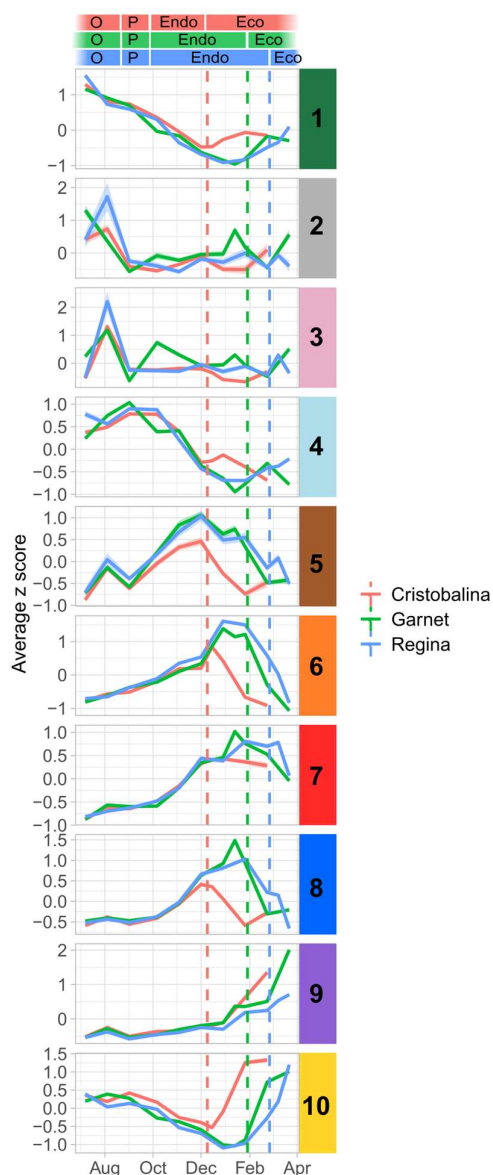


**Fig 6 Separation of samples by dormancy stage and cultivar using differentially expressed genes**

The principal component analysis was conducted on the TPM (transcripts per millions reads) values for the differentially expressed genes in the flower buds of the cultivars ‘Cristobalina’ (filled squares), ‘Garnet’ (empty circles) and ‘Regina’ (stars). Each point corresponds to one sampling time in a single tree.

252 When projected into a PCA 2-components plane, all samples harvested from buds at the same  
253 stage cluster together, whatever the cultivar (Fig. 6 and Fig. S5), suggesting that the stage of the bud  
254 has more impact on the transcriptional state than time or external conditions. Interestingly, the 100  
255 genes that contributed the most to the PCA dimensions 1 and 2 were very specifically associated with  
256 each dimension (Fig. S6, Table S6). We further investigated which clusters were over-represented in  
257 these genes (see Fig. S6b in Additional file 1) and we found that genes belonging to the clusters 6 and  
258 8, associated with endodormancy, were particularly represented in the best contributors to the  
259 dimension 1. In particular, we identified genes involved in oxidation-reduction processes like  
260 *PavGPX6*, and stress-induced genes such as *PavLEA14*, together with genes potentially involved in  
261 leaf and flower development, including *GROWTH-REGULATING FACTOR7* (*PavGRF7*) and  
262 *PavSEPI* (Table S6). In contrast, genes that best contributed to the dimension 2 strictly belonged to  
263 clusters 9 and 10, therefore characterized by high expression during ecodormancy (Fig. S6 in

264 Additional file 1). These results suggest that bud stages can mostly be separated by two criteria:  
 265 dormancy depth before dormancy release, defined by genes highly expressed during endodormancy,



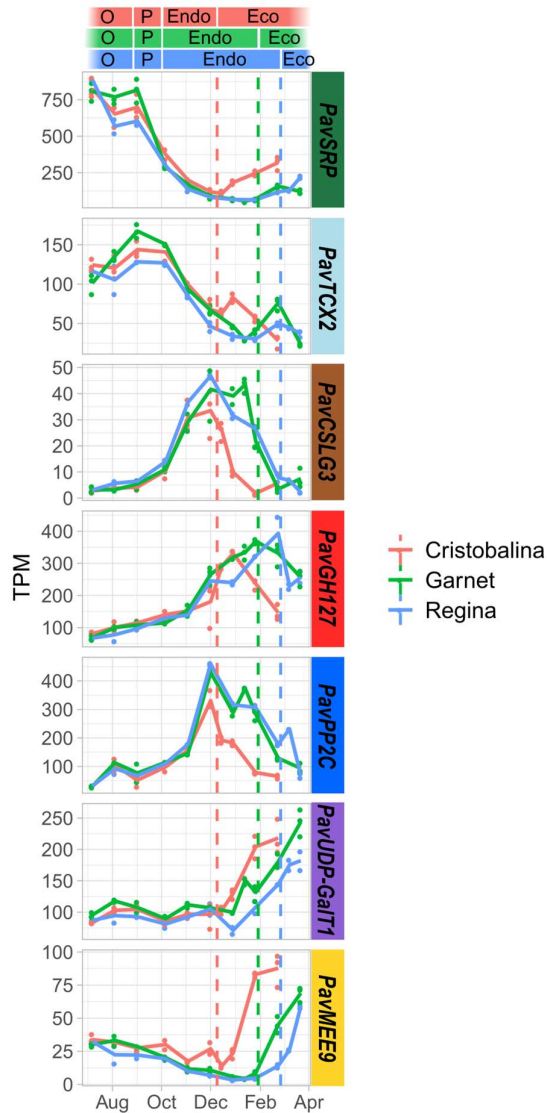
**Figure 7. Expression patterns in the ten clusters for the three cultivars**

Expression patterns were analysed from August to March, covering bud organogenesis (O), paradormancy (P), endodormancy (Endo), and ecodormancy (Eco). Dash lines represent the estimated date of dormancy breaking, in red for ‘Cristobalina’, green for ‘Garnet’ and blue for ‘Regina’. Average *z-score* patterns (line) and standard deviation (ribbon), calculated using the TPM values from the RNA-seq analysis, for the genes belonging to the ten clusters.

and the dichotomy defined by the status before/after dormancy release.

To go further, we compared transcriptional profiles throughout the time course in all cultivars. For this we analysed the expression profiles in each cultivar for the clusters previously identified for the cultivar ‘Garnet’ (Fig. 7, see also Fig. S7). In general, averaged expression profiles for all clusters are very similar in all three varieties, with the peak of expression happening at a similar period of the year. However, we can distinguish two main phases according to similarities or disparities between cultivars. First, averaged expression profiles are almost similar in all cultivars between July and November. This is especially the case for clusters 1, 4, 7, 8 and 9.

On the other hand, we can observe a temporal shift in the peak of expression between varieties from December onward for genes in clusters 1, 5, 6, 8 and 10. Indeed, in these clusters, the peak or drop in expression happens earlier in ‘Cristobalina’, and slightly later in



**Figure 8. Expression patterns for the seven marker genes in the three cultivars**

Expression patterns were analysed from August to March, covering bud organogenesis (O), paradormancy (P), endodormancy (Endo), and ecodormancy (Eco). Dash lines represent the estimated date of dormancy breaking, in red for ‘Cristobalina’, green for ‘Garnet’ and blue for ‘Regina’. TPM were obtained from the RNA-seq analysis for the seven marker genes from clusters 1, 4, 5, 7, 8, 9 and 10. Lines represent the average TPM, dots are the actual values from the biological replicates. *SRP*: *STRESS RESPONSIVE PROTEIN*; *TCX2*: *TESMIN/TSO1-like CXC 2*; *CSLG3*: *Cellulose Synthase like G3*; *GH127*: *Glycosyl Hydrolase 127*; *PP2C*: *Phosphatase 2C*; *UDP-GalT1*: *UDP-Galactose transporter 1*; *MEE9*: *maternal effect embryo arrest 9*.

‘Regina’ compared to ‘Garnet’ (Fig. 7), in correlation with their dormancy release dates. These results seem to confirm that the organogenesis and paradormancy phases occur concomitantly in the three cultivars while temporal shifts between cultivars are observed after endodormancy onset. Therefore, similarly to the PCA results (Fig. 6), the expression profile of these genes is more associated with the flower bud stage than with external environmental conditions.

### Flower bud stage can be predicted using a small set of marker genes

We have shown that flower buds in organogenesis, paradormancy, endodormancy and ecodormancy are characterised by specific transcriptional states. In theory, we could therefore use transcriptional data to infer the flower bud stage. For this, we selected a minimum number of seven marker genes, one gene for each of the clusters 1, 4, 5, 7, 8, 9 and 10 (identified in Fig 3), for which expression presented the best correlation with the average expression profiles of their cluster (Fig. 8). We

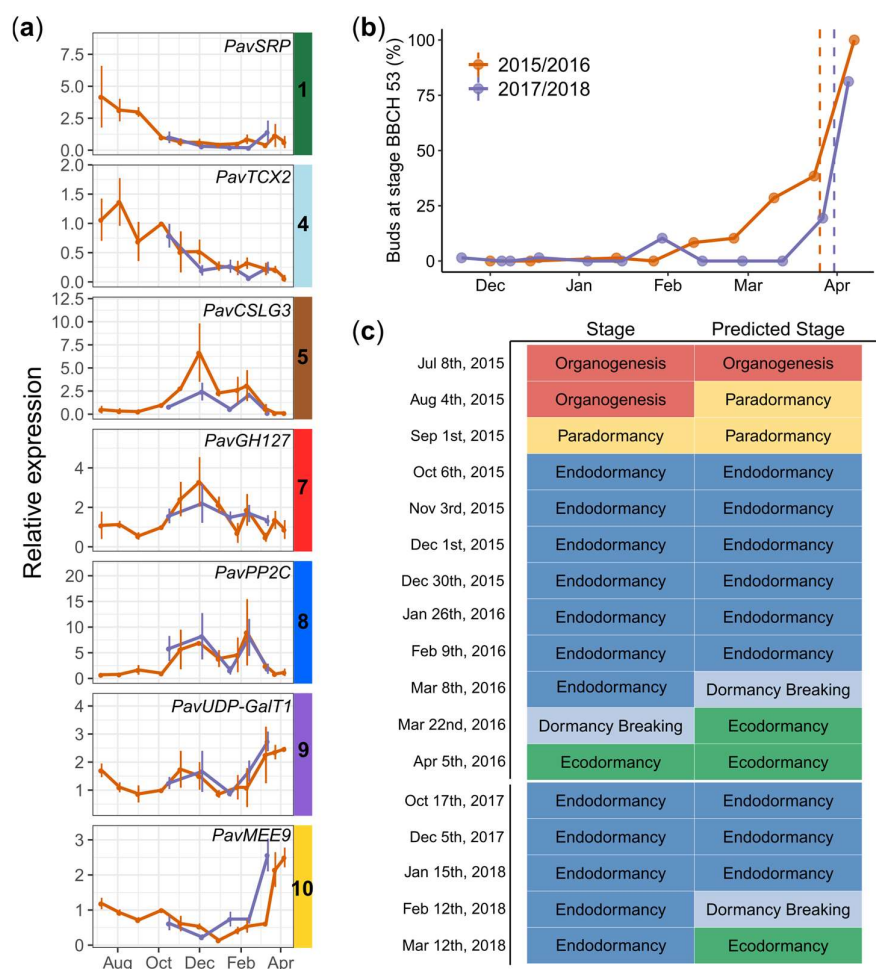


314 aimed to select the minimum number of marker genes that are sufficient to infer the flower bud stage,  
315 therefore excluding the clusters 2, 3 and 6 as they either had very small number of genes, or had  
316 expression profiles very similar to another cluster.

317 Expression for these marker genes not only recapitulates the average profile of the cluster they  
318 originate from, but also temporal shifts in the profiles between the three cultivars (Fig. 8). In order to  
319 define if these genes encompass as much information as the full transcriptome, or all DEGs, we  
320 performed a PCA of all samples harvested for all three cultivars using expression levels of these seven  
321 markers (Fig. S9). The clustering of samples along the two main axes of the PCA using these seven  
322 markers is very similar, if not almost identical, to the PCA results obtained using expression for all  
323 DEGs (Fig. 6). This indicates that the transcriptomic data can be reduced to only seven genes and still  
324 provides accurate information about the flower bud stages.

325         To test if these seven markers can be used to define the flower bud stage, we used a multinomial  
326 logistic regression modelling approach to predict the flower bud stage in our dataset based on the  
327 expression levels for these seven genes in the three cultivars ‘Garnet’, ‘Regina’ and ‘Cristobalina’  
328 (Fig. 9). For this, we trained and tested the model, on randomly picked sets, to predict the five bud  
329 stage categories, and obtained a very high model accuracy (100%; Fig. S9). These results indicate that  
330 the bud stage can be accurately predicted based on expression data by just using seven genes. In order  
331 to go further and test the model in an independent experiment, we analysed the expression for the  
332 seven marker genes by RT-qPCR on buds sampled from another sweet cherry tree cultivar ‘Fertard’  
333 for two consecutive years (Fig. 9a, b). Based on these RT-qPCR data, we predicted the flower bud  
334 developmental stage using the parameters of the model obtained from the training set on the three  
335 cultivars ‘Garnet’, ‘Regina’ and ‘Cristobalina’. We achieved a high accuracy of 71% for our model  
336 when tested on RT-qPCR data to predict the flower bud stage for the ‘Fertard’ cultivar (Fig. 9c and  
337 Fig. S9c). In particular, the chronology of bud stages was very well predicted. This result indicates that

338 these seven genes can be used as a diagnostic tool in order to infer the flower bud stage in sweet cherry  
 339 trees.



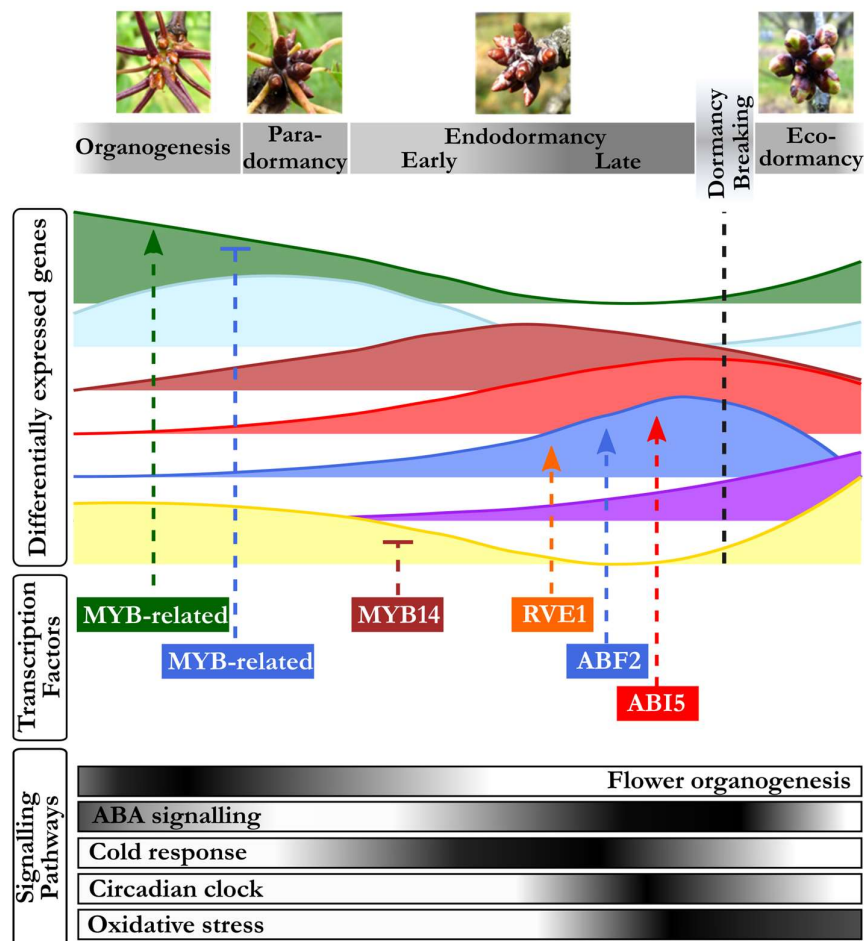
**Figure 9. Expression for the seven marker genes allows accurate prediction of the bud dormancy stages in the late flowering cultivar ‘Fertard’ during two bud dormancy cycles**

(a) Relative expressions were obtained by qRT-PCR and normalized by the expression of two reference constitutively expressed genes *PavRPII* and *PavEF1*. (b) Evaluation of the dormancy status in ‘Fertard’ flower buds during the two seasons using the percentage of open flower buds (BBCH stage 53). (c) Predicted vs experimentally estimated bud stages. *SRP*: *STRESS RESPONSIVE PROTEIN*; *TCX2*: *TESMIN/TSO1-like CXC 2*; *CSLG3*: *Cellulose Synthase like G3*; *GH127*: *Glycosyl Hydrolase 127*; *PP2C*: *Phosphatase 2C*; *UDP-GalT1*: *UDP-Galactose transporter 1*; *MEE9*: *maternal effect embryo arrest 9*.

## 340 DISCUSSION

341 In this work, we have characterised transcriptional changes at a genome-wide scale happening  
 342 throughout cherry tree flower bud dormancy, from organogenesis to the end of dormancy. To do this,  
 343 we have analysed expression in flower buds at 11 dates from July 2015 to March 2016 for three

344 cultivars displaying different dates of dormancy release, generating 82 transcriptomes in total. This  
 345 resource, with a fine time resolution, reveals key aspects of the regulation of cherry tree flower buds  
 346 during dormancy (Fig. 10). We have shown that buds in organogenesis, paradormancy, endodormancy  
 347 and ecodormancy are characterised by distinct transcriptional states (Fig. 2, 3) and we highlighted the  
 348 different pathways activated during the main cherry tree flower bud dormancy stages (Fig. 4 and Table  
 349 1). Finally, we found that just seven genes are enough to accurately predict the main cherry tree flower  
 350 bud dormancy stages (Fig. 9).



**Figure 10. From bud formation to flowering: transcriptomic regulation of flower bud dormancy**

Our results highlighted seven main expression patterns corresponding to the main dormancy stages. During organogenesis and paradormancy (July to September), signalling pathways associated with flower organogenesis and ABA signalling are upregulated. Distinct groups of genes are activated during different phases of endodormancy, including targets of transcription factors involved in ABA signalling, cold response and circadian clock. ABA: abscisic acid.

351 ***DAMs, floral identity and organogenesis genes characterize the pre-dormancy stages***

352 To our knowledge, this is the first report on the transcriptional regulation of early stages of  
353 flower bud development in temperate fruit trees. Information on dormancy onset and pre-dormancy  
354 bud stages are scarce and we arbitrarily delimited the organogenesis and paradormancy in July/August  
355 and September, respectively. However, based on transcriptional data, we could detect substantial  
356 discrepancies suggesting that the definition of the bud stages can be improved. Indeed, we observe that  
357 samples harvested from buds during phases that we defined as organogenesis and paradormancy  
358 cluster together in the PCA, but away from samples harvested during endodormancy. Moreover, most  
359 of the genes highly expressed during paradormancy are also highly expressed during organogenesis.  
360 This is further supported by the fact that paradormancy is a flower bud stage predicted with less  
361 accuracy based on expression level of the seven marker genes. In details, paradormancy is defined as  
362 a stage of growth inhibition originating from surrounding organs [7] therefore it is strongly dependent  
363 on the position of the buds within the tree and the branch. Our results suggest that defining  
364 paradormancy for multiple cherry flower buds based on transcriptomic data is difficult and even raise  
365 the question of whether paradormancy can be considered as a specific flower bud stage. Alternatively,  
366 we propose that the pre-dormancy period should rather be defined as a continuum between  
367 organogenesis, growth and/or growth cessation phases. Further physiological observations, including  
368 flower primordia developmental context [46], could provide crucial information to precisely link the  
369 transcriptomic environment to these bud stages. Nonetheless, we found very few, if not at all,  
370 differences between the three cultivars for the expression patterns during organogenesis and  
371 paradormancy, supporting the hypothesis that pre-dormancy processes are not associated with the  
372 different timing in dormancy release and flowering that we observe between these cultivars.

373 Our results showed that specific pathways were specifically activated before dormancy onset.  
374 The key role of ABA in the control of bud set and dormancy onset has been known for decades and  
375 we found that the ABA-related transcription factor *PavWRKY40* is expressed as early as during

376 organogenesis. Several studies have highlighted a role of *PavWRKY40* homolog in Arabidopsis in  
377 ABA signalling, in relation with light transduction [47, 48] and biotic stresses [49]. These results  
378 suggest that there might be an early response to ABA in flower buds. Furthermore, we uncovered the  
379 upregulation of several pathways linked to organogenesis during the summer months, including  
380 *PavMYB63* and *PavMYB93*, expressed during early organogenesis, with potential roles in the  
381 secondary wall formation [50] and root development [51]. Interestingly, *TESMIN/TSO1-like CXC 2*  
382 (*PavTCX2*), defined here as a marker gene for organogenesis and paradormancy, is the homolog of an  
383 Arabidopsis TF potentially involved in stem cell division [52]. We found that targets for *PavTCX2*  
384 may be over-represented in genes up-regulated during endodormancy, thus suggesting that *PavTCX2*  
385 acts on bud development by repressing dormancy-associated genes. In accordance with the  
386 documented timing of floral initiation and development in sweet cherry [53], several genes involved  
387 in floral identity and flower development, including *PavAGL20*, *PavFD*, as well as targets of  
388 *PavSEP3*, *PavAP3* and *PavAG*, were markedly upregulated during the early stages of flower bud  
389 development. Many studies conducted on fruit trees support the key role of *DAM* genes in the control  
390 of dormancy establishment and maintenance [18] and we found expression patterns very similar to the  
391 peach *DAM* genes with *PavDAM1* and *PavDAM3*, as well as *PavDAM6*, expressed mostly during  
392 summer [54]. The expression of these three genes was at the highest before endodormancy and seems  
393 to be inhibited by cold exposure from October onward, similarly to previous results obtained in sweet  
394 cherry [55], peach [56], Japanese apricot [57] and apple [58]. These results further suggest a major  
395 role for *PavDAM1*, *PavDAM3* and *PavDAM6* in dormancy establishment, bud onset and growth  
396 cessation in sweet cherry.

397

398 **Integration of environmental and internal signals through a complex array of signaling**  
399 **pathways during endodormancy**

400 Previous studies have proved the key role of a complex array of signaling pathways in the  
401 regulation of endodormancy onset and maintenance that subsequently lead to dormancy breaking,  
402 including genes involved in cold response, phytohormone-associated pathways and oxidation-  
403 reduction processes. Genes associated with the response to cold, notably, have been shown to be up-  
404 regulated during endodormancy such as dehydrins and *DREB* genes identified in oak, pear and leafy  
405 spurge [24, 27, 59]. We observe an enrichment for GO involved in the response to abiotic and biotic  
406 responses, as well as an enrichment for targets of many TFs involved in the response to environmental  
407 factors. In particular, our results suggest that *PavMYB14*, which has a peak of expression in November  
408 just before the cold period starts, is repressing genes that are subsequently expressed during  
409 endodormancy. This is in agreement with the fact that *AtMYB14*, the *PavMYB14* homolog in  
410 *Arabidopsis thaliana*, is involved in cold stress response regulation [39]. Although these results were  
411 not confirmed in *Populus* [60], two MYB DOMAIN PROTEIN genes (*MYB4* and *MYB14*) were also  
412 up-regulated during the induction of dormancy in grapevine [61]. Similarly, we identified an  
413 enrichment in genes highly expressed during endodormancy with target motifs of a transcription factor  
414 belonging to the CBF/DREB family. These TFs have previously been implicated in cold acclimation  
415 and endodormancy in several perennial species [59, 62]. These results are in agreement with the  
416 previous observation showing that genes responding to cold are differentially expressed during  
417 dormancy in other tree species [24]. Cold acclimation is the ability of plants to adapt to and withstand  
418 freezing temperatures and is triggered by decreasing temperatures and photoperiod. Therefore  
419 mechanisms associated with cold acclimation are usually observed concomitantly to the early stages  
420 of endodormancy. The stability of membranes and a strict control of cellular homeostasis are crucial  
421 in the bud survival under cold stress and we observe that genes associated with cell wall organization  
422 and nutrient transporters are up-regulated at the beginning of endodormancy, including the  
423 *CELLULOSE SYNTHASE-LIKE G3* (*PavCSLG3*) marker gene.

424 Similarly to seed dormancy processes, hormonal signals act in a complex way to balance  
425 dormancy maintenance and growth resumption. In particular, ABA levels have been shown to increase  
426 in response to environmental signals such as low temperatures and/or shortening photoperiod, and  
427 trigger dormancy induction [63–65]. Several studies have also shown that a subsequent drop in ABA  
428 concentration is associated with dormancy release [64, 66]. These results are supported by previous  
429 reports where genes involved in ABA signaling are differentially expressed during dormancy in  
430 various tree species (for e.g., see [19, 20, 22, 24, 67]). We find ABA-related pathways to be central in  
431 our transcriptomic analysis of sweet cherry bud dormancy, with the enrichment of GO terms related  
432 to ABA found in the genes highly expressed during endodormancy. These genes, including ABA-  
433 degradation gene *PavCYP707A2*, ABA-response factor *PavABF2*, and the Protein phosphatase 2C  
434 (*PavPP2C*) marker gene, are then inhibited after dormancy release in the three cultivars. Accordingly,  
435 we identified a key role for ABA-associated genes *PavABI5* and *PavABF2* in the regulation of  
436 dormancy progression in our dataset. These two transcription factors are mainly expressed around the  
437 time of dormancy release, like their target, and their homologs in Arabidopsis are involved in key ABA  
438 processes, especially during seed dormancy [68]. These results are consistent with records that  
439 *PmABF2* is highly expressed during endodormancy in Japanese apricot [22]. These results suggest  
440 ABA-related mechanisms similar in sweet cherry to those previously observed in other trees control  
441 bud dormancy onset and release. One of the hypotheses supports an activation of ABA-induced  
442 dormancy by *DAM* genes [64, 69] and we observed that *PavDAM4* expression pattern is very similar  
443 to ABA-related genes. We can therefore hypothesize that *PavDAM4* has a key role in dormancy onset  
444 and maintenance, potentially by regulating ABA metabolism. On the other side of the pathway,  
445 ground-breaking works have revealed that ABA signaling is crucial in triggering dormancy onset by  
446 inducing plasmodesmata closure, potentially through callose deposit [65, 70]. Accordingly, we found  
447 that *PavGHI7* genes involved in callose degradation are highly activated before and after

448 endodormancy while their expression is inhibited during endodormancy, thus suggesting that callose  
449 deposit is activated during endodormancy in sweet cherry flower buds.

450 In plants, response to environmental and developmental stimuli usually involves pathways  
451 associated with circadian clock regulation. This is also true for bud dormancy where the interplay  
452 between environmental and internal signals necessitates circadian clock genes for an optimal response  
453 [4, 71–74]. Indeed, transcriptomic analyses conducted in poplar showed that among the genes up-  
454 regulated during endodormancy, were genes with the EVENING ELEMENT (EE) motifs, that are  
455 important regulators of circadian clock and cold-responsive genes, and components of the circadian  
456 clock, including *LATE-ELONGATE HYPOCOTYL (LHY)* and *ZEITLUPE (ZTL)* [60, 67]. We  
457 identified an enrichment of targets for *PavRVE8* and *PavRVE1* among the genes expressed around the  
458 time of dormancy release. Homologs of *RVE1* are also up-regulated during dormancy in leafy spurge  
459 [43] and apple [75]. These TFs are homologs of Arabidopsis MYB transcription factors involved in  
460 the circadian clock. In particular, *AtRVE1* seems to integrate several signalling pathways including  
461 cold acclimation and auxin [76–78] while *AtRVE8* is involved in the regulation of circadian clock by  
462 modulating the pattern of H3 acetylation [79]. Our findings that genes involved in the circadian clock  
463 are expressed and potentially regulate genes at the time of dormancy release strongly support the  
464 hypothesis that environmental cues might be integrated with internal factors to control dormancy and  
465 growth in sweet cherry flower buds.

466 Consistently with observations that elevated levels of the reactive species of oxygen  $H_2O_2$  are  
467 strongly associated with dormancy release [80], oxidative stress is considered as one of the important  
468 processes involved in the transition between endodormancy and ecodormancy [30, 81, 82]. In line with  
469 these findings, we identified genes involved in oxidation-reduction processes that are up-regulated just  
470 before endodormancy release including *PavGPX6* and *PavGR*, that are involved in the detoxification  
471 systems. In their model for the control of dormancy, Ophir and colleagues [82] hypothesize that  
472 respiratory stress, ethylene and ABA pathways interact to control dormancy release and growth



473 resumption. Our results concur with this hypothesis to some extent albeit the key role of *DAM* genes  
474 should be further explored. Co-regulation analyses will be needed to investigate whether oxidative  
475 stress signalling is involved upstream to trigger dormancy release or downstream as a consequence of  
476 cellular activity following dormancy release in sweet cherry buds, leading to a better understanding of  
477 how other pathways interact or are directly controlled by oxidative cues.

478

#### 479 **Global cell activity characterizes the ecodormancy stage in sweet cherry flower buds**

480 Following the release of endodormancy, buds enter the ecodormancy stage, which is a state of inhibited  
481 growth controlled by external signals that can therefore be reversed by exposure to growth-promoting  
482 signals [7]. This transition towards the ability to grow is thought to be associated with the prolonged  
483 downregulation of *DAM* genes (see [18] for review), regulated by epigenetic mechanisms such as  
484 histone modifications [62, 83–85] and DNA methylation [55], in a similar way to *FLC* repression  
485 during vernalization in Arabidopsis. We observe that the expression of all *PavDAM* genes is inhibited  
486 before dormancy release, thus supporting the hypothesis that *DAM* genes may be involved in dormancy  
487 maintenance. In particular, the transition to ecodormancy coincides with a marked decrease in  
488 *PavDAM4* expression, which suggest that the regulation of its expression is crucial in the progression  
489 of dormancy towards growth resumption. However, other MADS-box transcription factors were found  
490 to be up-regulated during ecodormancy, including *PavAG* and *PavAP3*, similarly to previous results  
491 obtained in Chinese cherry (*Prunus pseudocerasus*) [28]. We also found that the marker gene  
492 *PavMEE9*, expressed during ecodormancy, is orthologous to the Arabidopsis gene *MATERNAL*  
493 *EFFECT EMBRYO ARREST 9 (MEE9)*, required for female gametophyte development [86], which  
494 could suggest active cell differentiation during the ecodormancy stage.

495 As mentioned before, in-depth studies conducted on poplar have led to the discovery that the regulation  
496 of the movements through the plasma membrane plays a key role not only in dormancy onset but also  
497 in dormancy release [87]. This is also true for long-distance transport with the observation that in

498 peach, for example, active sucrose import is renewed during ecodormancy [88]. In sweet cherry, our  
499 results are consistent with these processes since we show that GO terms associated with  
500 transmembrane transporter activity are enriched for genes highly expressed during ecodormancy.  
501 Transmembrane transport capacity belongs to a wide range of membrane structures modifications  
502 tightly regulated during dormancy. For example, lipid content, linoleic and linolenic acids composition  
503 and unsaturation degree of fatty acids in the membrane are modified throughout dormancy progression  
504 [30] and these changes in the membrane structure may be associated with modifications in the  
505 cytoskeleton [87]. Consistently, we find that genes involved in microtubule-based processes and cell  
506 wall organization are up-regulated during ecodormancy in sweet cherry flower buds. For example, the  
507 marker gene *PavUDP-GalT1*, orthologous to a putative UDP-galactose transmembrane transporter, is  
508 highly express after dormancy release in all three cultivars.

509 Overall, all processes triggered during ecodormancy are associated with cell activity. The  
510 trends observed here suggest that after endodormancy release, transmembrane and long distance  
511 transports are reactivated, thus allowing an active uptake of sugars, leading to increased oxidation-  
512 reduction processes and cell proliferation and differentiation.

513

#### 514 **Development of a diagnostic tool to define the flower bud dormancy stage using seven genes**

515 We find that sweet cherry flower bud stage can be accurately predicted with the expression of  
516 just seven genes. It indicates that combining expression profiles of just seven genes is enough to  
517 recapitulate all transcriptional states in our study. This is in agreement with previous work showing  
518 that transcriptomic states can be accurately predicted using a relatively low number of markers [89].  
519 Marker genes were not selected on the basis of their function and indeed, two genes are orthologous  
520 to Arabidopsis proteins of unknown function: *PavSRP* (Stress responsive A/B Barrel Domain-  
521 containing protein) and *PavGHI27* (putative glycosyl hydrolase). However, as reported above, some  
522 of the selected marker genes are involved in the main pathways regulating dormancy progression,

523 including cell wall organization during the early phase of endodormancy (*PavCSLG3*), ABA  
524 (*PavPP2C*), transmembrane transport (*PavUDP-GalT1*) and flower primordia development  
525 (*PavMEE9*).

526 Interestingly, when there are discrepancies between the predicted bud stages and the ones  
527 defined by physiological observations, the model always predicts that stages happen earlier than the  
528 actual observations. For example, the model predicts that dormancy breaking occurs instead of  
529 endodormancy, or ecodormancy instead of dormancy breaking. This could suggest that transcriptional  
530 changes happen before we can observe physiological changes. This is indeed consistent with the  
531 indirect phenotyping method currently used, based on the observation of the response to growth-  
532 inducible conditions after ten days. Using these seven genes to predict the flower bud stage would thus  
533 potentially allow to identify these important transitions when they actually happen.

534 We show that the expression level of these seven genes can be used to predict the flower bud  
535 stage in other conditions and genotypes by performing RT-qPCR. Also this independent experiment  
536 has been done on two consecutive years and shows that RT-qPCR for these seven marker genes as  
537 well as two control genes are enough to predict the flower bud stage in cherry trees. It shows that  
538 performing a full transcriptomic analysis is not necessary if the only aim is to define the dormancy  
539 stage of flower buds.

540

## 541 **CONCLUSIONS**

542 In this work, we have characterized transcriptional changes throughout all stages of sweet  
543 cherry flower bud development and dormancy. To our knowledge, no analysis had previously been  
544 conducted on this range of dates in temperate trees. Pathways involved at different stages of bud  
545 dormancy have been investigated in other species and we confirmed that genes associated with the  
546 response to cold, ABA and development processes were also identified during sweet cherry flower  
547 bud dormancy. We took advantage of the extended timeframe and we highlighted genes and pathways

548 associated with specific phases of dormancy, including early endodormancy, deep endodormancy and  
549 dormancy release. For that reason, our results suggest that commonly used definitions of bud dormancy  
550 are too restrictive and transcriptomic states might be useful to redefine the dormancy paradigm, not  
551 only for sweet cherry but also for other species that undergo overwintering. We advocate for large  
552 transcriptomic studies that take advantage of the wide range of genotypes available in forest and fruit  
553 trees, aiming at the mechanistic characterization of dormancy stages. Furthermore, we then went a step  
554 beyond the global transcriptomic analysis and we developed a model based on the transcriptional  
555 profiles of just seven genes to accurately predict the main dormancy stages. This offers an alternative  
556 approach to methods currently used such as assessing the date of dormancy release by using forcing  
557 conditions. In addition, this result sets the stage for the development of a fast and cost effective  
558 diagnostic tool to molecularly define the dormancy stages in cherry trees. This approach, from  
559 transcriptomic data to modelling, could be tested and transferred to other fruit tree species and such  
560 diagnostic tool would be very valuable for researchers working on fruit trees as well as for plant  
561 growers, notably to define the best time for the application of dormancy breaking agents, whose  
562 efficiency highly depends on the state of dormancy progression.

563

## 564 **METHODS**

### 565 *Plant material*

566 Branches and flower buds were collected from four different sweet cherry cultivars with contrasted  
567 flowering dates: ‘Cristobalina’, ‘Garnet’, ‘Regina’ and ‘Fertard’, which display extra-early, early, late  
568 and very late flowering dates, respectively. ‘Cristobalina’, ‘Garnet’, ‘Regina’ trees were grown in an  
569 orchard located at the Fruit Experimental Unit of INRA in Bourran (South West of France, 44° 19' 56''  
570 N, 0° 24' 47'' E), under the same agricultural practices. ‘Fertard’ trees were grown in an orchard at the  
571 Fruit Experimental Unit of INRA in Toulence, near Bordeaux (48° 51' 46'' N, 2° 17' 15'' E). During  
572 the first sampling season (2015/2016), ten or eleven dates spanning the entire period from flower bud

573 organogenesis (July 2015) to bud break (March 2016) were chosen for RNA sequencing (Fig. 1a and  
574 Additional file 11), while bud tissues from ‘Fertard’ were sampled in 2015/2016 (12 dates) and  
575 2017/2018 (7 dates) for validation by RT-qPCR (Additional file 11). For each date, flower buds were  
576 sampled from different trees, each tree corresponding to a biological replicate. Upon harvesting, buds  
577 were flash frozen in liquid nitrogen and stored at -80°C prior to performing RNA-seq.

578

### 579 *Measurements of bud break and estimation of the dormancy release date*

580 For the two sampling seasons, 2015/2016 and 2017/2018, three branches bearing floral buds were  
581 randomly chosen fortnightly from ‘Cristobalina’, ‘Garnet’, ‘Regina’ and ‘Fertard’ trees, between  
582 November and flowering time (March-April). Branches were incubated in water pots placed under  
583 forcing conditions in a growth chamber (25°C, 16h light/ 8h dark, 60-70% humidity). The water was  
584 replaced every 3-4 days. After ten days under forcing conditions, the total number of flower buds that  
585 reached the BBCH stage 53 [46, 90] was recorded. The date of dormancy release was estimated as the  
586 date when the percentage of buds at BBCH stage 53 was above 50% after ten days under forcing  
587 conditions (Fig. 1a).

588

### 589 *RNA extraction and library preparation*

590 Total RNA was extracted from 50-60 mg of frozen and pulverised flower buds using RNeasy Plant  
591 Mini kit (Qiagen) with minor modification: 1.5% PVP-40 was added in the extraction buffer RLT.  
592 RNA quality was evaluated using Tapestation 4200 (Agilent Genomics). Library preparation was  
593 performed on 1 µg of high quality RNA (RNA integrity number equivalent superior or equivalent to  
594 8.5) using the TruSeq Stranded mRNA Library Prep Kit High Throughput (Illumina cat. no. RS-122-  
595 2103) for ‘Cristobalina’, ‘Garnet’ and ‘Regina’ cultivars. DNA quality from libraries was evaluated  
596 using Tapestation 4200. The libraries were sequenced on a NextSeq500 (Illumina), at the Sainsbury  
597 Laboratory Cambridge University (SLCU), using paired-end sequencing of 75 bp in length.

598

### 599 *Mapping and differential expression analysis*

600 The raw reads obtained from the sequencing were analysed using several publicly available software  
601 and in-house scripts. The quality of reads was assessed using FastQC  
602 ([www.bioinformatics.babraham.ac.uk/projects/fastqc/](http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)) and possible adaptor contaminations were  
603 removed using Trimmomatic [91]. Trimmed reads were mapped to the peach (*Prunus persica* (L)  
604 Batsch) reference genome v.2 [92] (genome sequence and information can be found at the following  
605 address: [https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org\\_Ppersica](https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Ppersica)) using Tophat [93].  
606 Possible optical duplicates were removed using Picard tools (<https://github.com/broadinstitute/picard>).  
607 The total number of mapped reads of each samples are given in Additional file 15. For each gene, raw  
608 read counts and TPM (Transcripts Per Million) numbers were calculated [94].

609 We performed a differential expression analysis on data obtained from the ‘Garnet’ samples. First,  
610 data were filtered by removing lowly expressed genes (average read count < 3), genes not expressed  
611 in most samples (read counts = 0 in more than 75% of the samples); and genes presenting little change  
612 in expression between samples (coefficient of variation < 0.3). Then, differentially expressed genes  
613 (DEGs) between bud stages (organogenesis – 6 biological replicates, paradormancy – 3 biological  
614 replicates, endodormancy – 10 biological replicates, dormancy breaking – 6 biological replicates,  
615 ecodormancy – 6 biological replicates, see Additional file 11) were assessed using DEseq2 R  
616 Bioconductor package [95], in the statistical software R (R Core Team 2018), on filtered data. Genes  
617 with an adjusted *p-value* (*padj*) < 0.05, using the Benjamini-Hochberg multiple testing correction  
618 method, were assigned as DEGs (Additional file 12). To enable researchers to access this resource, we  
619 have created a graphical web interface to allow easy visualisation of transcriptional profiles throughout  
620 flower bud dormancy in the three cultivars for genes of interest ([bwenden.shinyapps.io/DorPatterns/](http://bwenden.shinyapps.io/DorPatterns/)).

621

### 622 *Principal component analyses and hierarchical clustering*

623 Distances between the DEGs expression patterns over the time course were calculated based on  
624 Pearson's correlation on 'Garnet' TPM values. We applied a hierarchical clustering analysis on the  
625 distance matrix to define ten clusters (Additional file 12). For expression patterns representation, we  
626 normalized the data using *z-score* for each gene:

$$627 \quad z \text{ score} = \frac{(TPM_{ij} - mean_i)}{Standard \ Deviation}$$

628 where  $TPM_{ij}$  is the TPM value of the gene  $i$  in the sample  $j$ ,  $mean_i$  and  $standard \ deviation_i$  are the *mean*  
629 and *standard deviation* of the TPM values for the gene  $i$  over all samples.

630 Principal component analyses (PCA) were performed on TPM values from different datasets using the  
631 *prcomp* function from R.

632 For each cluster, using data for 'Garnet', 'Regina' and 'Cristobalina', mean expression pattern was  
633 calculated as the mean *z-score* value for all genes belonging to the cluster. We then calculated the  
634 Pearson's correlation between the *z-score* values for each gene and the mean *z-score* for each cluster.  
635 We defined the marker genes as genes with the highest correlation values, i.e. genes that represent the  
636 best the average pattern of the clusters. Keeping in mind that the marker genes should be easy to  
637 handle, we then selected the optimal marker genes displaying high expression levels while not  
638 belonging to extended protein families.

639

#### 640 ***Motif and transcription factor targets enrichment analysis***

641 We performed enrichment analysis on the DEG in the different clusters for transcription factor targets  
642 genes and target motifs.

643 Motif discovery on the DEG set was performed using Find Individual Motif occurrences (FIMO) [36].

644 Motif list available for peach was obtained from PlantTFDB 4.0 [35]. To calculate the  
645 overrepresentation of motifs, DEGs were grouped by motif (grouping several genes and transcripts in  
646 which the motif was found). Overrepresentation of motifs was performed using hypergeometric tests  
647 using Hypergeometric {stats} available in R. Comparison was performed for the number of

648 appearances of a motif in one cluster against the number of appearances on the overall set of DEG. As  
649 multiple testing implies the increment of false positives, *p-values* obtained were corrected using False  
650 Discovery Rate [96] correction method using `p.adjust{stats}` function available in R.

651 A list of predicted regulation between transcription factors and target genes is available for peach in  
652 PlantTFDB [35]. We collected the list and used it to analyse the overrepresentation of genes targeted  
653 by TF, using Hypergeometric {stats} available in R, comparing the number of appearances of a gene  
654 controlled by one TF in one cluster against the number of appearances on the overall set of DEG. *p-*  
655 *values* obtained were corrected using a false discovery rate as described above. Predicted gene  
656 homology to *Arabidopsis thaliana* and functions were retrieved from the data files available for *Prunus*  
657 *persica* (GDR, [https://www.rosaceae.org/species/prunus\\_persica/genome\\_v2.0.a1](https://www.rosaceae.org/species/prunus_persica/genome_v2.0.a1)).

658

### 659 ***GO enrichment analysis***

660 The list for the gene ontology (GO) terms was retrieved from the database resource PlantRegMap [35].  
661 Using the topGO package [97], we performed an enrichment analysis on GO terms for biological  
662 processes, cellular components and molecular functions based on a classic Fisher algorithm. Enriched  
663 GO terms were filtered with a *p-value* < 0.005 and the ten GO terms with the lowest *p-value* were  
664 selected for representation.

665

### 666 ***Marker genes selection and RT-qPCR analyses***

667 The seven marker genes were selected based on the following criteria:

- 668 • Their expression presented the best correlation with the average expression profiles of their  
669 cluster.
- 670 • They were not members of large families (in order to reduce issues caused by redundancy).
- 671 • We only kept genes for which we could design high efficiency primers for RT-qPCR.

672 Marker genes were not selected based on modelling fit, nor based on their function.



673 cDNA was synthesised from 1µg of total RNA using the iScript Reverse Transcriptase Kit (Bio-rad  
674 Cat no 1708891) in 20 µl of final volume. 2 µL of cDNA diluted to a third was used to perform the  
675 qPCR in a 20 µL total reaction volume. qPCRs were performed using a Roche LightCycler 480. Three  
676 biological replicates for each sample were performed. Primers used in this study for qPCR are:  
677 *PavCSLG3* F:CCAACCAACAAAGTTGACGA, R:CAACTCCCCCAAAAAGATGA; *PavMEE9*:  
678 F:CTGCAGCTGAACTGGAACAG, R:ACTCATCCATGGCACTCTCC; *PavSRP*:  
679 F:ACAGGATCTGGAAAGCCAAG, R:AGGGTGGCTCTGAAACACAG; *PavTCX2*:  
680 F:CTTCCCACAACGCCTTTACG, R:GGCTATGTCTCTCAAACCTGGA; *PavGHI27*:  
681 F:GCCATTGGTTGTAGGGTTTG, R:ATCCATTCAGCATTGTTTC; *PavUDP-GALT1*  
682 F:CAATGTTGCTGGAAACCTCA, R:GTTATTCCACATCCGACAGC; *PavPP2C*  
683 F:CTGTGCCTGAAGTGACACAGA, R:CTGCACTGCTTCTTGATTG; *PavRPII*  
684 F:TGAAGCATAACCTATGATGATGAAG, R:CTTTGACAGCACCAGTAGATTCC; *PavEF1*  
685 F:CCCTTCGACTTCCACTTCAG, R:CACAAGCATAACCAGGCTTCA. Primers were tested for non-specific  
686 products by separation on 1.5% agarose gel electrophoresis and by sequencing each amplicon.  
687 Realtime data were analyzed using custom R scripts. Expression was estimated for each gene in each  
688 sample using the relative standard curve method based on cDNA diluted standards. For the  
689 visualization of the marker genes' relative expression, we normalized the RT-qPCR results for each  
690 marker gene by the average RT-qPCR data for the reference genes *PavRPII* and *PavEF1*.

691

### 692 ***Bud stage predictive modelling***

693 In order to predict the bud stage based on the marker genes transcriptomic data, we used TPM values  
694 for the marker genes to train and test several models. First, all samples were projected into a 2-  
695 dimensional space using PCA, to transform potentially correlated data to an orthogonal space. The  
696 new coordinates were used to train and test the models to predict the five bud stage categories. In  
697 addition, we tested the model on RT-qPCR data for samples harvested from the 'Fertard' cultivar. For

698 the modelling purposes, expression data for the seven marker genes were normalized by the expression  
699 corresponding to the October sample. We chose the date of October as the reference because it  
700 corresponds to the beginning of dormancy and it was available for all cultivars. For each date, the  
701 October-normalized expression values of the seven marker genes were projected in the PCA 2-  
702 dimension plan calculated for the RNA-seq data and they were tested against the model trained on  
703 ‘Cristobalina’, ‘Garnet’ and ‘Regina’ RNA-seq data.

704 We tested five different models (Multinomial logistic regression, Random forest classifier, k-nearest  
705 neighbour classifier, multi-layer perceptron and support vector classification) for 500 different  
706 combination of training/testing RNA-seq datasets, all implemented using the scikit-learn Python  
707 package [98]. The models were 5-fold cross-validated to ensure the robustness of the coefficients and  
708 to reduce overfitting. The models accuracies were calculated as the percentage of correct predicted  
709 stages in the RNA-seq testing set and the RT-qPCR dataset. Results presented in Fig. S10 (Additional  
710 file 1) show that the highest model accuracies were obtained for the support vector classification and  
711 the multinomial logistic regression models. We selected the logistic regression model for this study  
712 because the coefficients are biologically relevant.

713

#### 714 **LIST OF ABBREVIATIONS**

715 **ABA:** abscisic acid

716 **ABF2:** ABSCISIC ACID RESPONSE ELEMENT-BINDING FACTOR 2

717 **ABI5:** ABSCISIC ACID INSENSITIVE 5

718 **AG:** AGAMOUS

719 **AGL9:** AGAMOUS-like 9

720 **AGL20:** AGAMOUS-like 20

721 **AP3:** APETALA3

722 **AREB3:** ABSCISIC ACID RESPONSE ELEMENT-BINDING PROTEIN 3

723 **ATHB7:** ARABIDOPSIS THALIANA HOMEBOX 7

724 **CBF/DREB:** C-REPEAT/DRE BINDING FACTOR 2/DEHYDRATION RESPONSE ELEMENT-  
725 BINDING PROTEIN

726 **CSLG3**: Cellulose Synthase like G3  
727 **DAM**: DORMANCY ASSOCIATED MADS-box  
728 **DEG**: differentially expressed gene  
729 **DNA**: deoxyribonucleic acid  
730 **EE**: Evening element motif  
731 **EF1**: Elongation factor 1  
732 **ERF**: ethylene-responsive element  
733 **FD**: FLOWERING LOCUS D  
734 **FIMO**: Find Individual Motif occurrences  
735 **FLC**: FLOWERING LOCUS C  
736 **GH127**: Glycosyl Hydrolase 127  
737 **GPX6**: GLUTATHION PEROXIDASE 6  
738 **GR**: GLUTATHION REDUCTASE  
739 **GRF7**: GROWTH-REGULATING FACTOR7  
740 **GST8**: GLUTATHION S-TRANSFERASE8  
741 **GO**: gene ontology  
742 **H3**: Histone 3  
743 **LEA**: LATE EMBRYOGENESIS ABUNDANT PROTEIN  
744 **LHY**: LATE-ELONGATE HYPOCOTYL  
745 **LP1**: LIPID TRANSFER PROTEIN1  
746 **MEE9**: maternal effect embryo arrest 9  
747 **Padj**: adjusted p-value  
748 **Pav**: *Prunus avium*  
749 **PC**: principal component  
750 **PCA**: principal component analysis  
751 **PDCB3**: PLASMODESMATA CALLOSE-BINDING PROTEIN 3  
752 **PIF4**: PHYTOCHROME INTERACING FACTOR 4  
753 **PIL5**: PHYTOCHROME INTERACING FACTOR 3 LIKE 5  
754 **PP2C**: Phosphatase 2C  
755 **RNA**: ribonucleic acid  
756 **RPII**: ribonucleic acid polymerase II  
757 **RT-qPCR**: quantitative reverse transcriptase polymerase chain reaction  
758 **RVE1/8**: REVEILLE1/8  
759 **SEP3**: SEPALLATA3

760 **SPT:** SPATULA

761 **SRP:** STRESS RESPONSIVE PROTEIN

762 **TCX2:** TESMIN/TSO1-like CXC 2

763 **TF:** transcription factor

764 **TPM:** transcripts per million reads

765 **UDP-GalT1:** UDP-Galactose transporter 1

766 **ZTL:** ZEITLUPE

767

768 **DECLARATIONS**

769 **Ethics approval and consent to participate**

770 Not applicable.

771

772 **Consent for Publication**

773 Not applicable.

774

775 **Availability of data and materials**

776 RNA-seq data that support the findings of this study have been deposited in the NCBI Gene Expression

777 Omnibus under the accession code GSE130426.

778 The graphical web interface DorPatterns is available at the address:

779 [bwenden.shinyapps.io/DorPatterns](http://bwenden.shinyapps.io/DorPatterns).

780 Scripts and codes for data analysis and modelling will be available on github upon acceptance of the

781 manuscript.

782

783 **Competing interests**

784 The authors declare that they have no competing interests.

785

786 **Funding**

787 The PhD of Noemie Vimont was supported by a CIFRE grant funded by Agro Innovation International  
788 - Centre Mondial d'Innovation - Groupe Roullier (St Malo-France) and ANRT (France).

789

#### 790 **Author contributions**

791 SC, BW, ED and PAW designed the original research. MA and JCY participated to the project design.  
792 NV performed the RNA-seq and analysed the RNA-seq with SC and BW. MF performed the RT-  
793 qPCR. JAC performed the TF and motifs enrichment analysis. MT developed the model. NV, SC and  
794 BW wrote the article with the assistance of all the authors. All authors have read and approved the  
795 manuscript.

796

#### 797 **Acknowledgments**

798 We thank the Fruit Experimental Unit of INRA (Bordeaux-France) for growing and managing the  
799 trees, and Teresa Barreneche, Lydie Fouilhaux, Jacques Joly, H el ene Christman and R emi Beauvieux  
800 for the help during the harvest and for the pictures. Many thanks to Dr Varodom Charoensawan  
801 (Mahidol University, Thailand) for providing scripts for mapping and gene expression count  
802 extraction.

803

#### 804 **REFERENCES**

- 805 1. Heide OM, Prestrud AK. Low temperature, but not photoperiod, controls growth cessation and  
806 dormancy induction and release in apple and pear. *Tree Physiol.* 2005;25:109–14.  
807 <http://www.ncbi.nlm.nih.gov/pubmed/15519992>.
- 808 2. Allona I, Ramos A, Ib a nez C, Contreras A, Casado R, Aragoncillo C. Review. Molecular control  
809 of winter dormancy establishment in trees. *Spanish J Agric Res.* 2008;6:201–10.
- 810 3. Cooke JEK, Eriksson ME, Junttila O. The dynamic nature of bud dormancy in trees:  
811 environmental control and molecular mechanisms. *Plant Cell Env.* 2012;35:1707–28.

- 812 doi:10.1111/j.1365-3040.2012.02552.x.
- 813 4. Maurya JP, Triozzi PM, Bhalerao RP, Perales M. Environmentally Sensitive Molecular Switches  
814 Drive Poplar Phenology. *Front Plant Sci.* 2018;9 December:1–8.
- 815 5. Olsen JE. Light and temperature sensing and signaling in induction of bud dormancy in woody  
816 plants. *Plant Mol Biol.* 2010;73:37–47.
- 817 6. Cline MG, Deppong DO. The role of apical dominance in paradormancy of temperate woody  
818 plants: A reappraisal. *J Plant Physiol.* 1999;155:350–6. doi:10.1016/S0176-1617(99)80116-3.
- 819 7. Lang G, Early J, Martin G, Darnell R. Endo-, para-, and ecodormancy: physiological terminology  
820 and classification for dormancy research. *Hort Sci.* 1987;22:371–7.
- 821 8. Considine MJ, Considine JA. On the language and physiology of dormancy and quiescence in  
822 plants. *J Exp Bot.* 2016;67:3189–203.
- 823 9. Badeck FW, Bondeau A, Böttcher K, Doktor D, Lucht W, Schaber JJ, et al. Responses of spring  
824 phenology to climate change. *New Phytol.* 2004;162:295–309. doi:10.1111/j.1469-  
825 8137.2004.01059.x.
- 826 10. Menzel A, Sparks TH, Estrella N, Koch E, Aasa A, Ahas R, et al. European phenological  
827 response to climate change matches the warming pattern. *Glob Chang Biol.* 2006;12:1969–76.  
828 doi:10.1111/j.1365-2486.2006.01193.x.
- 829 11. Vitasse Y, Lenz A, Körner C. The interaction between freezing tolerance and phenology in  
830 temperate deciduous trees. *Front Plant Sci.* 2014;5:541.
- 831 12. Bigler C, Bugmann H. Climate-induced shifts in leaf unfolding and frost risk of European trees  
832 and shrubs. *Sci Rep.* 2018;8:1–10. doi:10.1038/s41598-018-27893-1.
- 833 13. Fu YH, Zhao H, Piao S, Peaucelle M, Peng S, Zhou G, et al. Declining global warming effects on  
834 the phenology of spring leaf unfolding. *Nature.* 2015;526:104–7.
- 835 14. Legave J-M, Guédon Y, Malagi G, El Yaacoubi A, Bonhomme M. Differentiated Responses of  
836 Apple Tree Floral Phenology to Global Warming in Contrasting Climatic Regions. *Front Plant Sci.*

- 837 2015;6 December. doi:10.3389/fpls.2015.01054.
- 838 15. Erez A. Bud Dormancy; Phenomenon, Problems and Solutions in the Tropics and Subtropics. In:  
839 Temperate Fruit Crops in Warm Climates. 2000. p. 17–48.
- 840 16. Atkinson CJ, Brennan RM, Jones HG. Declining chilling and its impact on temperate perennial  
841 crops. *Environ Exp Bot.* 2013;91:48–62. doi:10.1016/j.envexpbot.2013.02.004.
- 842 17. Snyder RL, de Melo-abreu JP. *Frost Protection : fundamentals , practice and economics.* Rome;  
843 2005.
- 844 18. Falavigna V da S, Guitton B, Costes E, Andrés F. I Want to (Bud) Break Free: The Potential  
845 Role of DAM and SVP-Like Genes in Regulating Dormancy Cycle in Temperate Fruit Trees. *Front*  
846 *Plant Sci.* 2019;9 January:1–17.
- 847 19. Zhong W, Gao Z, Zhuang W, Shi T, Zhang Z, Ni Z. Genome-wide expression profiles of  
848 seasonal bud dormancy at four critical stages in Japanese apricot. *Plant Mol Biol.* 2013;83:247–64.  
849 doi:10.1007/s11103-013-0086-4.
- 850 20. Khalil-Ur-Rehman M, Sun L, Li CX, Faheem M, Wang W, Tao JM. Comparative RNA-seq  
851 based transcriptomic analysis of bud dormancy in grape. *BMC Plant Biol.* 2017;17:1–11.  
852 doi:10.1186/s12870-016-0960-8.
- 853 21. Chao WS, Dođramacı M, Horvath DP, Anderson J V., Foley ME. Comparison of phytohormone  
854 levels and transcript profiles during seasonal dormancy transitions in underground adventitious buds  
855 of leafy spurge. *Plant Mol Biol.* 2017;94:281–302.
- 856 22. Zhang Z, Zhuo X, Zhao K, Zheng T, Han Y, Yuan C, et al. Transcriptome Profiles Reveal the  
857 Crucial Roles of Hormone and Sugar in the Bud Dormancy of *Prunus mume*. *Sci Rep.* 2018;8:1–15.  
858 doi:10.1038/s41598-018-23108-9.
- 859 23. Min Z, Zhao X, Li R, Yang B, Liu M, Fang Y. Comparative transcriptome analysis provides  
860 insight into differentially expressed genes related to bud dormancy in grapevine (*Vitis vinifera*). *Sci*  
861 *Hortic (Amsterdam).* 2017;225 March:213–20. doi:10.1016/j.scienta.2017.06.033.

- 862 24. Ueno S, Klopp C, Leplé JC, Derory J, Noirot C, Léger V, et al. Transcriptional profiling of bud  
863 dormancy induction and release in oak by next-generation sequencing. *BMC Genomics*.  
864 2013;14:236.
- 865 25. Paul A, Jha A, Bhardwaj S, Singh S, Shankar R, Kumar S. RNA-seq-mediated transcriptome  
866 analysis of actively growing and winter dormant shoots identifies non-deciduous habit of evergreen  
867 tree tea during winters. *Sci Rep*. 2014;4:1–9.
- 868 26. Lesur I, Le Provost G, Bento P, Da Silva C, Leplé JC, Murat F, et al. The oak gene expression  
869 atlas: Insights into Fagaceae genome evolution and the discovery of genes regulated during bud  
870 dormancy release. *BMC Genomics*. 2015;16:112.
- 871 27. Takemura Y, Kuroki K, Shida Y, Araki S, Takeuchi Y, Tanaka K, et al. Comparative  
872 transcriptome analysis of the less-dormant taiwanese pear and the dormant Japanese pear during  
873 winter season. *PLoS One*. 2015;10.
- 874 28. Zhu Y, Li Y, Xin D, Chen W, Shao X, Wang Y, et al. RNA-Seq-based transcriptome analysis of  
875 dormant flower buds of Chinese cherry (*Prunus pseudocerasus*). *Gene*. 2015;555:362–76.  
876 doi:10.1016/j.gene.2014.11.032.
- 877 29. Kumar G, Rattan UK, Singh AK. Chilling-mediated DNA methylation changes during dormancy  
878 and its release reveal the importance of epigenetic regulation during winter dormancy in Apple  
879 (*Malus x domestica* Borkh.). *PLoS One*. 2016;11:1–25. doi:10.1371/journal.pone.0149934.
- 880 30. Beauvieux R, Wenden B, Dirlewanger E. Bud Dormancy in Perennial Fruit Tree Species : A  
881 Pivotal Role for Oxidative Cues. *Front Plant Sci*. 2018;9 May:1–13.
- 882 31. Lloret A, Badenes ML, Ríos G. Modulation of Dormancy and Growth Responses in  
883 Reproductive Buds of Temperate Trees. *Front Plant Sci*. 2018;9:1–12.
- 884 32. Campoy JA, Ruiz D, Egea J. Dormancy in temperate fruit trees in a global warming context: A  
885 review. *Sci Hortic (Amsterdam)*. 2011;130:357–72. doi:10.1016/j.scienta.2011.07.011.
- 886 33. Wenden B, Campoy JA, Jensen M, López-Ortega G. Climatic Limiting Factors: Temperature. In:



- 887 Quero-García J, Iezzoni A, Pulawska J, Lang G, editors. Cherries: Botany, Production and Uses.  
888 CABI Publishing; 2017. p. 166–88.
- 889 34. Heide OM. Interaction of photoperiod and temperature in the control of growth and dormancy of  
890 Prunus species. *Sci Hortic (Amsterdam)*. 2008;115:309–14. doi:10.1016/j.scienta.2007.10.005.
- 891 35. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, et al. PlantTFDB 4.0: Toward a central hub  
892 for transcription factors and regulatory interactions in plants. *Nucleic Acids Res*. 2017;45:D1040–5.
- 893 36. Grant CE, Bailey TL, Noble WS. FIMO: Scanning for occurrences of a given motif.  
894 *Bioinformatics*. 2011;27:1017–8.
- 895 37. Causier B, Schwarz-Sommer Z, Davies B. Floral organ identity: 20 years of ABCs. *Semin Cell*  
896 *Dev Biol*. 2010;21:73–9. doi:10.1016/j.semcdb.2009.10.005.
- 897 38. Lee S-J, Kang J-Y, Park H-J, Kim MD, Bae MS, Choi H, et al. DREB2C Interacts with ABF2, a  
898 bZIP Protein Regulating Abscisic Acid-Responsive Gene Expression, and Its Overexpression Affects  
899 Abscisic Acid Sensitivity. *Plant Physiol*. 2010;153:716–27.
- 900 39. Chen Y, Chen Z, Kang J, Kang D, Gu H, Qin G. AtMYB14 Regulates Cold Tolerance in  
901 Arabidopsis. *Plant Mol Biol Report*. 2013;31:87–97.
- 902 40. Liu Q, Wang Z, Xu X, Zhang H, Li C. Genome-wide analysis of C2H2 zinc-finger family  
903 transcription factors and their responses to abiotic stresses in poplar (*Populus trichocarpa*). *PLoS*  
904 *One*. 2015;10:1–25.
- 905 41. Kiełbowicz-Matuk A. Involvement of plant C2H2-type zinc finger transcription factors in stress  
906 responses. *Plant Sci*. 2012;185–186:78–85.
- 907 42. Koornneef M, Léon-Kloosterziel KM, Schwartz SH, Zeevaart JAD. The genetic and molecular  
908 dissection of abscisic acid biosynthesis and signal transduction in Arabidopsis. *Plant Physiol*  
909 *Biochem*. 1998;36:83–9.
- 910 43. Dođramacı M, Horvath DP, Anderson J V. Dehydration-induced endodormancy in crown buds of  
911 leafy spurge highlights involvement of MAF3- and RVE1-like homologs, and hormone signaling

- 912 cross-talk. *Plant Mol Biol.* 2014;86:409–24.
- 913 44. Franklin KA, Lee SH, Patel D, Kumar SV, Spartz AK, Gu C, et al. PHYTOCHROME-  
914 INTERACTING FACTOR 4 (PIF4) regulates auxin biosynthesis at high temperature. *Proc Natl*  
915 *Acad Sci.* 2011;108:20231–5.
- 916 45. Penfield S, Josse E-M, Kannangara R, Gilday AD, Halliday KJ, Graham IA. Cold and light  
917 control seed germination through the bHLH transcription factor SPATULA. *Curr Biol.*  
918 2005;15:1998–2006.
- 919 46. Fadón E, Herrero M, Rodrigo J. Flower development in sweet cherry framed in the BBCH scale.  
920 *Sci Hortic (Amsterdam).* 2015;192:141–7. doi:10.1016/j.scienta.2015.05.027.
- 921 47. Geilen K, Böhmer M. Dynamic subnuclear relocalization of WRKY40, a potential new  
922 mechanism of ABA-dependent transcription factor regulation. *Plant Signal Behav.*  
923 2015;10:e1106659.
- 924 48. Liu R, Xu Y-H, Jiang S-C, Lu K, Lu Y-F, Feng X-J, et al. Light-harvesting chlorophyll a/b-  
925 binding proteins, positively involved in abscisic acid signalling, require a transcription repressor,  
926 WRKY40, to balance their function. *J Exp Bot.* 2013;64:5443–56.
- 927 49. Pandey SP, Roccaro M, Schön M, Logemann E, Somssich IE. Transcriptional reprogramming  
928 regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis. *Plant J.*  
929 2010;64:912–23.
- 930 50. Zhou J, Lee C, Zhong R, Ye Z-H. MYB58 and MYB63 Are Transcriptional Activators of the  
931 Lignin Biosynthetic Pathway during Secondary Cell Wall Formation in Arabidopsis. *Plant Cell*  
932 *Online.* 2009;21:248–66.
- 933 51. Gibbs DJ, Voß U, Harding SA, Fannon J, Moody LA, Yamada E, et al. AtMYB93 is a novel  
934 negative regulator of lateral root development in Arabidopsis. *New Phytol.* 2014;203:1194–207.
- 935 52. Simmons AR, Davies KA, Wang W, Liu Z, Bergmann DC. SOL1 and SOL2 regulate fate  
936 transition and cell divisions in the arabidopsis stomatal lineage. *Dev.* 2019;146.

- 937 53. Engin H, Ünal A. Examination of Flower Bud Initiation and Differentiation in Sweet Cherry and  
938 Peach by Scanning Electron Microscope. *Turk J Agric.* 2007;31:373–9.
- 939 54. Li Z, Reighard GL, Abbott AG, Bielenberg DG. Dormancy-associated MADS genes from the  
940 EVG locus of peach [*Prunus persica* (L.) Batsch] have distinct seasonal and photoperiodic expression  
941 patterns. *J Exp Bot.* 2009;60:3521–30. doi:10.1093/jxb/erp195.
- 942 55. Rothkegel K, Sánchez E, Montes C, Greve M, Tapia S, Bravo S, et al. DNA methylation and  
943 small interference RNAs participate in the regulation of MADS-box genes involved in dormancy in  
944 sweet cherry (*Prunus avium* L.). *Tree Physiol.* 2017; October:1–13.
- 945 56. Jiménez S, Reighard GL, Bielenberg DG. Gene expression of DAM5 and DAM6 is suppressed  
946 by chilling temperatures and inversely correlated with bud break rate. *Plant Mol Biol.* 2010;73:157–  
947 67. doi:10.1007/s11103-010-9608-5.
- 948 57. Zhao K, Zhou Y, Ahmad S, Xu Z, Li Y, Yang W, et al. Comprehensive Cloning of *Prunus mume*  
949 Dormancy Associated MADS-Box Genes and Their Response in Flower Bud Development and  
950 Dormancy. *Front Plant Sci.* 2018;9 February:1–12. doi:10.3389/fpls.2018.00017.
- 951 58. Mimida N, Saito T, Moriguchi T, Suzuki A, Komori S, Wada M. Expression of DORMANCY-  
952 ASSOCIATED MADS-BOX (DAM)-like genes in apple. *Biol Plant.* 2015;59:237–44.
- 953 59. Dođramaci M, Horvath DP, Chao WS, Foley ME, Christoffers MJ, Anderson J V. Low  
954 temperatures impact dormancy status, flowering competence, and transcript profiles in crown buds of  
955 leafy spurge. *Plant Mol Biol.* 2010;73:207–26.
- 956 60. Howe GT, Horvath DP, Dharmawardhana P, Priest HD, Mockler TC, Strauss SH. Extensive  
957 Transcriptome Changes During Natural Onset and Release of Vegetative Bud Dormancy in *Populus*.  
958 *Front Plant Sci.* 2015;6.
- 959 61. Fennell AY, Schlauch KA, Gouthu S, Deluc LG, Khadka V, Sreekantan L, et al. Short day  
960 transcriptomic programming during induction of dormancy in grapevine. *Front Plant Sci.*  
961 2015;6:834.

- 962 62. Leida C, Conesa A, Llácer G, Badenes ML, Ríos G. Histone modifications and expression of  
963 DAM6 gene in peach are modulated during bud dormancy release in a cultivar-dependent manner.  
964 *New Phytol.* 2012;193:67–80. doi:10.1111/j.1469-8137.2011.03863.x.
- 965 63. Götz K-P, Chmielewski FM, Homann T, Huschek G, Matzneller P, Rawel HM. Seasonal  
966 changes of physiological parameters in sweet cherry (*Prunus avium* L.) buds. *Sci Hortic*  
967 (Amsterdam). 2014;172:183–90. doi:10.1016/j.scienta.2014.04.012.
- 968 64. Tuan PA, Bai S, Saito T, Ito A, Moriguchi T. Dormancy-Associated MADS-Box ( DAM ) and  
969 the Abscisic Acid Pathway Regulate Pear Endodormancy Through a Feedback Mechanism. *Plant*  
970 *Cell Physiol.* 2017;58:1378–90. doi:10.1093/pcp/pcx074.
- 971 65. Tylewicz S, Petterle A, Marttila S, Miskolczi P, Azeez A, Singh RK, et al. Photoperiodic control  
972 of seasonal growth is mediated by ABA acting on cell-cell communication. *Science* (80- ).  
973 2018;360:212–5.
- 974 66. Leida C, Conejero A, Arbona V, Gómez-Cadenas A, Llácer G, Badenes ML, et al. Chilling-  
975 dependent release of seed and bud dormancy in peach associates to common changes in gene  
976 expression. *PLoS One.* 2012;7:e35777. doi:10.1371/journal.pone.0035777.
- 977 67. Ruttink T, Arend M, Morreel K, Storme V, Rombauts S, Fromm J, et al. A molecular timetable  
978 for apical bud formation and dormancy induction in poplar. *Plant Cell.* 2007;19:2370–90.  
979 doi:10.1105/tpc.107.052811.
- 980 68. Lopez-Molina L, Mongrand S, McLachlin DT, Chait BT, Chua NH. ABI5 acts downstream of  
981 ABI3 to execute an ABA-dependent growth arrest during germination. *Plant J.* 2002;32:317–28.
- 982 69. Yamane H, Wada M, Honda C, Matsuura T, Ikeda Y, Hirayama T, et al. Overexpression of  
983 *Prunus* DAM6 inhibits growth, represses bud break competency of dormant buds and delays bud  
984 outgrowth in apple plants. *PLoS One.* 2019;14:1–24.
- 985 70. Singh RK, Miskolczi P, Maurya JP, Bhalerao RP. A Tree Ortholog of SHORT VEGETATIVE  
986 PHASE Floral Repressor Mediates Photoperiodic Control of Bud Dormancy. *Curr Biol.*

- 987 2019;29:128-133.e2. doi:10.1016/j.cub.2018.11.006.
- 988 71. Ibáñez C, Kozarewa I, Johansson M, Ogren E, Rohde A, Eriksson ME. Circadian clock  
989 components regulate entry and affect exit of seasonal dormancy as well as winter hardiness in  
990 Populus trees. *Plant Physiol.* 2010;153:1823–33. doi:10.1104/pp.110.158220.
- 991 72. Kozarewa I, Ibáñez C, Johansson M, Ogren E, Mozley D, Nylander E, et al. Alteration of PHYA  
992 expression change circadian rhythms and timing of bud set in Populus. *Plant Mol Biol.* 2010;73:143–  
993 56. doi:10.1007/s11103-010-9619-2.
- 994 73. Ding J, Böhlenius H, Rühl MG, Chen P, Sane S, Zambrano JA, et al. GIGANTEA-like genes  
995 control seasonal growth cessation in Populus. *New Phytol.* 2018;218:1491–503.
- 996 74. Johansson M, Ramos-sánchez JM, Conde D, Ibáñez C, Takata N, Allona I, et al. Role of the  
997 Circadian Clock in Cold Acclimation and Winter Dormancy in Perennial Plants. In: Anderson J,  
998 editor. *Advances in Plant Dormancy.* Springer, Cham; 2015. p. 51–74.
- 999 75. Denardi Porto D, Bruneau M, Perini P, Anzanello R, Renou JJ-P, Santos HPD, et al.  
1000 Transcription profiling of the chilling requirement for bud break in apples: a putative role for FLC-  
1001 like genes. *J Exp Bot.* 2015;66:2659–72. doi:10.1093/jxb/erv061.
- 1002 76. Meissner M, Orsini E, Ruschhaupt M, Melchinger AE, Hinch DK, Heyer AG. Mapping  
1003 quantitative trait loci for freezing tolerance in a recombinant inbred line population of Arabidopsis  
1004 thaliana accessions Tenela and C24 reveals REVEILLE1 as negative regulator of cold acclimation.  
1005 *Plant, Cell Environ.* 2013;36:1256–67.
- 1006 77. Jiang Z, Xu G, Jing Y, Tang W, Lin R. Phytochrome B and REVEILLE1/2-mediated signalling  
1007 controls seed dormancy and germination in Arabidopsis. *Nat Commun.* 2016;7:1–10.  
1008 doi:10.1038/ncomms12377.
- 1009 78. Rawat R, Schwartz J, Jones MA, Sairanen I, Cheng Y, Andersson CR, et al. REVEILLE1, a  
1010 Myb-like transcription factor, integrates the circadian clock and auxin pathways. *Proc Natl Acad Sci.*  
1011 2009;106:16883–8. doi:10.1073/pnas.0813035106.

- 1012 79. Farinas B, Mas P. Histone acetylation and the circadian clock: A role for the MYB transcription  
1013 factor RVE8/LCL5. *Plant Signal Behav.* 2011;6:541–3.
- 1014 80. Pérez FJ, Vergara R, Rubio S. H<sub>2</sub>O<sub>2</sub> is involved in the dormancy-breaking effect of hydrogen  
1015 cyanamide in grapevine buds. *Plant Growth Regul.* 2008;55:149–55.
- 1016 81. Vergara R, Rubio S, Pérez FJ. Hypoxia and hydrogen cyanamide induce bud-break and up-  
1017 regulate hypoxic responsive genes (HRG) and VvFT in grapevine-buds. *Plant Mol Biol.*  
1018 2012;79:171–8.
- 1019 82. Ophir R, Pang X, Halaly T, Venkateswari J, Lavee S, Galbraith D, et al. Gene-expression  
1020 profiling of grape bud response to two alternative dormancy-release stimuli expose possible links  
1021 between impaired mitochondrial activity, hypoxia, ethylene-ABA interplay and cell enlargement.  
1022 *Plant Mol Biol.* 2009;71:403–23.
- 1023 83. Horvath DP, Sung S, Kim D-H, Chao WS, Anderson J. Characterization, expression and function  
1024 of DORMANCY ASSOCIATED MADS-BOX genes from leafy spurge. *Plant Mol Biol.*  
1025 2010;73:169–79.
- 1026 84. de la Fuente L, Conesa A, Lloret A, Badenes ML, Ríos G. Genome-wide changes in histone H3  
1027 lysine 27 trimethylation associated with bud dormancy release in peach. *Tree Genet Genomes.*  
1028 2015;11:45. doi:10.1007/s11295-015-0869-7.
- 1029 85. Saito T, Bai S, Imai T, Ito A, Nakajima I, Moriguchi T. Histone modification and signalling  
1030 cascade of the dormancy-associated MADS-box gene, PpMADS13-1, in Japanese pear (*Pyrus*  
1031 *pyrifolia*) during endodormancy. *Plant Cell Environ.* 2015;38:1157–66. doi:10.1111/pce.12469.
- 1032 86. Pagnussat GC, Yu HJ, Ngo QA, Rajani S, Mayalagu S, Johnson CS, et al. Genetic and molecular  
1033 identification of genes required for female gametophyte development and function in *Arabidopsis*.  
1034 *Development.* 2005;132:603–14.
- 1035 87. Rinne PL, van der Schoot C. Plasmodesmata at the crossroads between development, dormancy,  
1036 and defense. *Can J Bot.* 2003;81:1182–97.

- 1037 88. Marquat C, Vandamme M, Gendraud M, Pétel G. Dormancy in vegetative buds of peach:  
1038 Relation between carbohydrate absorption potentials and carbohydrate concentration in the bud  
1039 during dormancy and its release. *Sci Hortic (Amsterdam)*. 1999;79:151–62.
- 1040 89. Biswas S, Kerner K, Teixeira PJPL, Dangl JL, Jojic V, Wigge PA. Tradict enables accurate  
1041 prediction of eukaryotic transcriptional states from 100 marker genes. *Nat Commun*. 2017;8:15309.
- 1042 90. Meier U. Growth stages of mono-and dicotyledonous plants *BBCH Monograph*. 2001.  
1043 <http://pub.jki.bund.de/index.php/BBCH/article/view/515/464>.
- 1044 91. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data.  
1045 *Bioinformatics*. 2014;30:2114–20.
- 1046 92. Verde I, Jenkins J, Dondini L, Micali S, Pagliarani G, Vendramin E, et al. The Peach v2.0  
1047 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale  
1048 assembly and contiguity. *BMC Genomics*. 2017;18:225.
- 1049 93. Trapnell C, Pachter L, Salzberg SL. TopHat: Discovering splice junctions with RNA-Seq.  
1050 *Bioinformatics*. 2009;25:1105–11.
- 1051 94. Wagner D. Chromatin regulation of plant development. *Curr Opin Plant Biol*. 2003;6:20–8.
- 1052 95. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq  
1053 data with DESeq2. *Genome Biol*. 2014;15:1–21.
- 1054 96. Benjamini Y, Hochberg J. Controlling the false discovery rate: a practical and powerful approach  
1055 to multiple testing. *J R Stat Soc Ser B*. 1995;57:289–300.
- 1056 97. Alexa A, Rahnenführer J. topGO: Enrichment Analysis for Gene Ontology. R Packag version  
1057 2340. 2018. [www.bioconductor.org](http://www.bioconductor.org).
- 1058 98. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn:  
1059 Machine Learning in Python. *J Mach Learn Res*. 2011;12:2825–30. doi:10.1007/s13398-014-0173-  
1060 7.2.
- 1061

1062 **Table 1. Transcription factors with over-represented targets in the different clusters**

Targets cluster	TF Name	Peach genome (v2) gene id	TF Cluster	Predicted TF family	TF Arabidopsis homologous	TF Predicted function	Enrichment p value	Enrichment adjusted p value
1 - dark green	PavMYB63	Prupe.4G136300	1 - dark green	MYB	AT1G79180	Myb-related protein	2.1E-05	6.7E-03 (**)
	PavMYB93	Prupe.6G188300	1 - dark green	MYB	AT1G34670	Myb-related protein	9.0E-04	3.2E-02 (*)
	PavMYB40	Prupe.3G299000	8 - royal blue	MYB	AT5G14340	Myb-related protein	2.7E-04	1.7E-02 (*)
	PavMYB17	Prupe.2G164300	-	MYB	AT3G61250	Myb-related protein	6.8E-05	7.2E-03 (**)
	PavMYB94	Prupe.5G193200	-	MYB	AT3G47600	Myb-related protein	9.0E-05	7.2E-03 (**)
	PavMYB60	Prupe.7G018400	-	MYB	AT1G08810	Myb-related protein	7.0E-05	7.2E-03 (**)
	PavMYB61	Prupe.6G303300	-	MYB	AT1G09540	Myb-related protein	4.0E-04	2.1E-02 (*)
	PavMYB3	Prupe.1G551400	-	MYB	AT1G22640	Myb-related protein	6.0E-04	2.8E-02 (*)
2 - grey	PavMYB67	Prupe.4G126900	-	MYB	AT3G12720	Myb-related protein	7.8E-04	3.1E-02 (*)
3 - pink	PavWRKY40	Prupe.3G098100	3 - pink	WRKY	AT1G80840	WRKY transcription factor	8.4E-05	1.2E-02 (*)
	PavWRKY11	Prupe.1G122800	-	CAMTA	AT4G16150	Calmodulin-binding transcription activator	4.9E-09	1.4E-06 (***)
	PavWRKY11	Prupe.1G459100	-	WRKY	AT4G31550	WRKY transcription factor	4.7E-04	4.5E-02 (*)
5 - brown	PavCBF4	Prupe.2G289500	-	ERF	AT5G51990	Dehydration-responsive element-binding protein	2.0E-04	5.7E-02
6 - orange	PavERF110	Prupe.6G165700	8 - royal blue	ERF	AT5G50080	Ethylene-responsive transcription factor	3.1E-04	5.2E-02
	PavRVE8	Prupe.6G242700	8 - royal blue	MYB	AT3G09600	Homeodomain-like superfamily protein RVE8	4.3E-04	5.2E-02
	PavRAP2.12	Prupe.3G032300	-	ERF	AT1G53910	Ethylene-responsive transcription factor	4.9E-04	5.2E-02
8 - royal blue	PavRVE1	Prupe.3G014900	6 - orange	MYB	AT5G17300	Homeodomain-like superfamily protein RVE1	1.0E-03	3.6E-02 (*)
	PavABI5	Prupe.7G112200	7 - red	bZIP	AT2G36270	ABSCISIC ACID-INSENSITIVE 5	6.6E-05	7.0E-03 (**)
	PavABF2	Prupe.1G434500	8 - royal blue	bZIP	AT1G45249	abscisic acid responsive elements-binding factor	2.4E-06	7.5E-04 (***)
	PavAREB3	Prupe.2G056800	-	bZIP	AT3G56850	ABA-responsive element binding protein	1.4E-05	2.2E-03 (**)
	PavPIL5	Prupe.8G209100	-	bHLH	AT2G20180	phytochrome interacting factor 3-like 5	2.3E-04	1.9E-02 (*)
	PavbZIP16	Prupe.5G027000	-	bZIP	AT2G35530	basic region/leucine zipper transcription factor	4.3E-04	2.7E-02 (*)
	PavSPT	Prupe.7G131400	-	bHLH	AT4G36930	Transcription factor SPATULA	5.6E-04	3.0E-02 (*)
	PavBPE	Prupe.1G263800	-	bHLH	AT1G59640	Transcription factor BPE	1.0E-03	3.6E-02 (*)
	PavPIF4	Prupe.3G179800	-	bHLH	AT2G43010	phytochrome interacting factor 4	9.5E-04	3.6E-02 (*)
9 - purple	PavGBF3	Prupe.2G182800	-	bZIP	AT2G46270	G-box binding factor 3	1.1E-03	3.6E-02 (*)
	PavWRKY50	Prupe.1G407500	-	WRKY	AT5G26170	WRKY transcription factor	1.1E-04	1.8E-02 (*)
	PavWRKY1	Prupe.3G202000	-	WRKY	AT2G04880	WRKY transcription factor	5.8E-05	1.8E-02 (*)
10 - yellow	PavMYB14	Prupe.1G039200	5 - brown	MYB	AT2G31180	Myb-related protein	1.6E-04	3.9E-02 (*)
	PavNAC70	Prupe.8G002500	-	NAC	AT4G10350	NAC domain containing protein	2.4E-04	3.9E-02 (*)

1063 We investigated whether some transcription factors specifically targeted genes in specific clusters. Based on the gene regulation information  
1064 available for peach in PlantTFDB [35], overrepresentation of genes targeted by transcription factors was performed using hypergeometric tests. *p*-  
1065 values obtained were corrected using a false discovery rate: (\*\*\*) : adj. p-value < 0.001; (\*\*) : adj. p-value < 0.01; (\*) : adj. p-value < 0.05.



1066 **ADDITIONAL FILES**

1067 **Additional file 1:**

1068 **Figure S1** Field temperature during the sampling season

1069 **Figure S2** Separation of samples by dormancy stage using read counts for all genes

1070 **Figure S3** Enrichments in gene ontology terms in the ten clusters

1071 **Figure S4** Expression patterns for the transcription factors and their targets

1072 **Figure S5** Separation of samples by dormancy stage and cultivar using all genes

1073 **Figure S6** Analysis of the 100 genes that best contribute to the PCA dimensions 1 and 2

1074 **Figure S7** Clusters of expression patterns for differentially expressed genes in the sweet cherry

1075 cultivars 'Regina', 'Cristobalina' and 'Garnet'

1076 **Figure S8** Separation of samples by dormancy stage and cultivar using the seven marker genes

1077 **Figure S9** Multinomial logistic regression model details

1078 **Figure S10** Comparison of the accuracy for the five tested models

1079

1080 **Additional file 2:**

1081 **Table S1** Description of the flower bud samples used for RNA-seq and RT-qPCR

1082 **Table S2** 'Garnet' differentially expressed genes and their assigned clusters

1083 **Table S3** Transcription factors with enriched targets in the different clusters

1084 **Table S4** Enrichments in gene ontology terms in the transcription factors targets

1085 **Table S5** Transcription factors with targeted motif enrichment in the clusters

1086 **Table S6** Contribution of differentially expressed genes to the PCA dimensions 1 and 2

1087 **Table S7** RNA-seq mapped reads and gene count information

1088