Transcriptome Dynamics of Floral Organs Approaching Blooming in the Flowering Cherry (*Cerasus × yedoensis*) Cultivar 'Somei-Yoshino'

1 **Kenta Shirasawa[1*], Tomoya Esumi[2], Akihiro Itai[3], Sachiko Isobe[1]**

2 [1]Laboratory of Plant Genetics and Genomics, Department of Frontier Research and Development,
3 Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818, Japan

4 [2]Laboratory of Pomology & Viticulture, Department of Agricultural and Life Sciences, Shimane
5 University, Matsue, Shimane 690-8504, Japan

6 [3]Laboratory of Plant Resource Science, Department of Agricultural and Life Science, Kyoto
7 Prefectural University, Sakyo, Kyoto 606-8522, Japan

8 *** Correspondence:**
9 Kenta Shirasawa
10 shirasaw@kazusa.or.jp

11 Word count: 2796
12 Number of figures: 3
13 Number of tables: 1
14 Number of supplementary figures: 4
15 Number of supplementary tables: 3

16 **Keywords: cherry blossom forecast, flowering cherry, RNA sequencing, time-course analysis,**
17 **transcriptomics**

18 **Abstract**

19 To gain insights into the genetic mechanisms underlying blooming and petal movement in flowering
20 cherry (*Cerasus × yedoensis*), we performed time-course RNA-seq analysis of the floral buds and
21 open-flowers of the most popular flowering cherry cultivar, 'Somei-Yoshino'. Independent biological
22 duplicate samples of floral buds and open-flowers were collected from 'Somei-Yoshino' trees grown
23 at three different locations in Japan. RNA-seq reads obtained from floral bud and open-flower
24 samples collected in the current study (in 2019) and in a previous study (in 2017) were aligned
25 against the genome sequence of 'Somei-Yoshino' to quantify gene transcript levels. Clustering
26 analysis of RNA-seq reads revealed dynamic changes in the transcriptome, with genes in seven
27 modules predominantly expressed at specific time points, ranging from 5 weeks before flowering to 2
28 weeks after flowering. Based on the identified gene modules and Gene Ontology (GO) terms
29 enriched at different floral stages, we speculate that the genetic mechanisms underlying petal
30 movement and flower opening in cherry involve the processes of development, cell wall
31 organization, reproduction, and metabolism, which are executed by genes encoding transcription
32 factors, phytohormones, transporters, and polysaccharide metabolic enzymes. Furthermore, we
33 propose a method for cherry bloom forecasting, based on gene expression levels at different time
34 points before flowering as RNA markers.

35 **1   Introduction**

36 Flowering cherry, also known as sakura, typically blooms in the spring and is valued as a popular
37 ornamental flower across the world. 'Somei-Yoshino' (*Cerasus × yedoensis*), which is presumed to
38 be an interspecific hybrid between *C. spachiana* and *C. speciosa* (Takenaka, 1963; Innan et al., 1995;
39 Nakamura et al., 2015), is the most popular cultivar of flowering cherry in Japan. Given its genomic
40 heterozygosity and self-incompatibility, 'Somei-Yoshino' is propagated by grafting (Iketani et al.,
41 2007). Because of its clonal nature, 'Somei-Yoshino' trees planted within a specific location bloom
42 at the same time; however, across the Japanese archipelago, the blooming front progresses from
43 south to north because of differences in environmental conditions. Since the blooming date is
44 important for the tourism industry in the spring season, forecasting methods based on cumulative
45 temperature have been developed to predict the flowering date of cherry blossoms (Aono and
46 Murakami, 2017).

47 Flowering involves two main processes: floral bud initiation and flower opening. The molecular
48 mechanisms of floral bud initiation have been well studied in *Arabidopsis thaliana* and rice (*Oryza*
49 *sativa*) as flowering models of long- and short-day plants, respectively, revealing that *FLOWERING*
50 *LOCUS T* (*FT*) is the key gene involved in floral bud differentiation (Izawa et al., 2003). This
51 mechanism is widely conserved across plant species (Higuchi, 2018). In woody plants, following
52 floral bud initiation, buds enter a period of dormancy, which is followed by dormancy release and
53 bud break. In the Rosaceae family members, *DORMANCY-ASSOCIATED MADS-box* (*DAM*) genes
54 have been reported to regulate bud dormancy (Yamane, 2014). However, while the physiological
55 aspect of the flower opening mechanism has been thoroughly investigated, only a few studies have
56 been conducted to explore the genetic basis of this mechanism (van Doorn and Van Meeteren, 2003).

57 The genome sequence of 'Somei-Yoshino' has been determined at the chromosome level, and genes
58 involved in the regulation of dormancy and flowering time have been identified in this cultivar
59 through time-course transcriptome analysis (Shirasawa et al., 2019). Since expression levels of key
60 genes transmit the environmental conditions, such as day-length and temperature, to biological
61 processes such as blooming, the identification of these genes might help predict the blooming date in
62 flowering cherry. Because the transcriptome is affected by environmental conditions (e.g., changes in
63 weather and habitat), biological replications of flowering cherry trees over multiple years and
64 locations would be required for the accurate identification of genes affecting the blooming date.

65 In this study, we aimed to identify genes uniquely expressed before and after flowering, and to obtain
66 insights into the molecular mechanisms underlying blooming in flowering cherry. We collected floral
67 buds and open-flowers from 'Somei-Yoshino' trees planted at three locations in Japan (Chiba, Kyoto,
68 and Shimane) in 2019, and analyzed their transcriptome using the RNA sequencing (RNA-seq)
69 technology. The time-course transcriptome data generated in the current study and in our previous
70 study (Shirasawa et al., 2019) was used to characterize gene expression patterns in 'Somei-Yoshino'
71 floral buds, some of which could be used to forecast the blooming time and to gain insights into the
72 molecular mechanisms controlling blooming in flowering cherry.

73 **2   Materials and Methods**

74 **2.1   Plant Materials**

75 Seven clonal trees of 'Somei-Yoshino' were used in this study. One tree was planted at the Kazusa
76 DNA Research Institute (KDRI; Kisarazu, Chiba, Japan), and three trees each were planted at two
77 different locations: Shimane University (SU; Matsue, Shimane, Japan; trees 1–3) and Kyoto
78 Prefectural University (KPU; Sakyo, Kyoto, Japan; trees A–C). Floral buds and open-flowers were

collected in 2019 over an extended time period, ranging from 36 days before flowering (DBF) to 11 days after flowering (DAF) (Figure 1), corresponding to 6 weeks before flowering (WBF) to 2 weeks after flowering (WAF) (Supplementary Table S1).

2.2 RNA-seq Analysis

Library preparation and sequencing analysis were performed as described in Shirasawa et al. (2019). In short, total RNA was extracted from the buds and open-flowers with RNeasy Plant Min Kit (Qiagen, Hilden, Germany). The RNA was treated with RNase-free DNase (Promega, Madison, WI, USA) and used for library construction with TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA, USA). The library was sequenced on NextSeq 500 (Illumina) in paired-end, 76 bp mode.

2.3 RNA Sequence Data Processing

RNA-seq data was analyzed as described previously (Shirasawa et al., 2019). High-quality reads were selected by trimming the adapter sequences using fastx_clipper (parameter, -a AGATCGGAAGAGC) in the FASTX-Toolkit v.0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit), and by deleting low-quality bases using PRINSEQ v.0.20.4 (Schmieder and Edwards, 2011). High-quality reads were mapped to the CYE_r3.1_pseudomolecule sequence (Shirasawa et al., 2019) using HISAT2 v.2.1.0 (Kim et al., 2015), and reads mapped to each gene model were quantified and normalized to determine the number of fragments per kilobase of transcript per million mapped reads (FPKM) using StringTie v.1.3.5 (Pertea et al., 2015) and Ballgown v.2.14.1 (Frazee et al., 2015), as described previously (Pertea et al., 2016). Genes showing a variance of ≥1 in expression levels among samples were used for further analysis.

2.4 Gene Ontology (GO) Enrichment Analysis of Gene Modules

The weighted gene correlation network analysis (WGCNA; v.1.66) package of R (Langfelder and Horvath, 2008) was used for module detection. The floral bud and open-flower samples were grouped in accordance with gene expression levels using the hierarchical clustering algorithm. Then, highly co-expressed gene clusters were constructed as modules. Genes included in the modules were functionally annotated by performing a BLAST sequence similarity search (Altschul et al., 1997) against the non-redundant (nr) protein database. GO terms were assigned to genes, and the statistical analysis of the GO term enrichment in each module was performed using Fisher's exact test implemented in OmicsBox (BioBam, Valencia, Spain). GO maps were drawn with QuickGO (Huntley et al., 2009).

3 Results

3.1 RNA-seq Analysis and Quantification of Gene Expression Levels

A total of 94 samples were collected from flowering cherry trees at KDRI, SU, and KPU over a period of 47 days (i.e., 36 days before anthesis to 11 days post-anthesis) in 2019 (Figure 1, Supplementary Table S1). Total RNA was extracted from the samples and subjected to RNA-seq analysis, which returned approximately 4.7 million reads per sample. Nucleotide sequence data of the RNA-seq were deposited in the DDBJ Sequence Read Archive (accession numbers DRA012801, DRA012802, and DRA012803). We also utilized the RNA-seq data generated previously (Shirasawa et al. 2019; DRA accession number DRA008100) from samples collected at KDRI in 2017. The RNA-seq reads from both datasets were mapped on to the genome sequence of 'Somei-Yoshino' (Shirasawa et al., 2019), and expression levels were normalized in all samples based on the total

120 reads to obtain FPKM values for each gene across the samples. Of the 95,076 genes predicted in the
121 'Somei-Yoshino' genome (Shirasawa et al., 2019), a total of 29,712 genes (31.3%) were expressed
122 across the different samples, with a variance of ≥1.

123 3.2 Gene Module Detection

124 On the basis of the expression patterns of 29,712 genes, one floral bud sample collected at 27 days
125 before anthesis from the KDRI location was identified as an outlier and therefore was excluded from
126 further analyses (Supplementary Figure S1). The remaining 93 samples were roughly clustered into
127 53 highly co-expressed gene modules, based on the expression patterns of 29,712 genes
128 (Supplementary Figure S2), which were further classed into four groups, based on eigengenes
129 (Supplementary Figure S3). Genes in 7 out of 53 modules were prominently expressed at specific
130 days before or after anthesis (Figure 2): dark-red module (110 genes expressed at 4–5 weeks before
131 flowering [WBF]), tan module (267 genes, 4 WBF), pink module (332 genes, 3 WBF), royal blue
132 module (127 genes, 2–3 WBF), midnight blue module (176 genes, 1–2 WBF), black module (457
133 genes, at flowering days), and sky-blue module (72 genes, 1–2 weeks after flowering [WAF]).

134 3.3 GO Enrichment Analysis

135 To identify the GO terms enriched in each module, the ratios of GO terms in each gene module were
136 compared with those of the remaining gene set (Table 1, Figure 2, Supplementary Table S2,
137 Supplementary Figure S4).

138 At 4–5 WBF (dark-red module), the following GO terms were enriched: biological process (BP)
139 category: 'developmental process' related GO terms including 'aging', 'embryo development',
140 'developmental maturation', 'developmental process', and 'anatomical structure development';
141 molecular function (MF) category: 'oxidoreductase activity'; and cellular component (CC) category:
142 'plasma membrane' and 'cell periphery'.

143 At 4 WBF (tan module), GO terms enriched in the BP category were not only related to
144 'developmental process' ('anatomical structure formation involved in morphogenesis' and
145 'anatomical structure development') but also 'cellular process' ('sulfur compound metabolic
146 process', 'biosynthetic process', 'developmental process', and 'cell wall organization or biogenesis').
147 GO terms enriched in the MF category included 'DNA binding', 'DNA-binding transcription factor
148 activity', 'catalytic activity', 'oxidoreductase activity', 'transferase activity, transferring acyl groups',
149 'hydrolase activity, acting on glycosyl bonds', 'lyase activity', and 'transcription regulator activity',
150 while those enriched in the CC category included 'extracellular region'.

151 At 3 WBF (pink module), the following GO terms were enriched: BP category: 'lipid metabolic
152 process', 'autophagy', 'small molecule metabolic process', and 'process utilizing autophagic
153 mechanism'; MF category: 'catalytic activity', 'transporter activity', 'lipid binding', 'oxidoreductase
154 activity', 'transferase activity, transferring acyl groups', and 'transmembrane transporter activity';
155 and CC category: 'extracellular region', 'extracellular space', 'vacuole', 'plasma membrane',
156 'membrane', and 'cell periphery'.

157 At 2–3 WBF (royal blue module), the following GO terms were enriched: BP category: 'mitotic cell
158 cycle', 'carbohydrate metabolic process', 'catabolic process', 'secondary metabolic process', 'cell
159 division', and 'cell wall organization or biogenesis'; and CC category: 'extracellular region', 'cell
160 wall', 'Golgi apparatus', 'plasma membrane', 'endomembrane system', 'membrane', 'external
161 encapsulating structure', 'cell periphery', and 'cellular anatomical entity'.

162 At 1–2 WBF (midnight blue module), the enriched GO terms included 'pollen development' and
163 'gametophyte development' (BP category); and 'transporter activity' and 'transmembrane transporter
164 activity' (MF category).

165 At the flowering stage (black module), the following GO terms were enriched: BP category:
166 'carbohydrate metabolic process', 'cytoskeleton organization', 'catabolic process', and 'cell wall
167 organization or biogenesis'; MF category: 'lipid binding', 'lyase activity', 'enzyme regulator
168 activity', and 'molecular function regulator'; and CC category: 'extracellular region' and
169 'extracellular space'.

170 At 1–2 WAF (sky-blue module), the following GO terms were enriched: BP category: 'carbohydrate
171 metabolic process', 'cellular protein modification process', 'response to stress', 'cell
172 communication', 'signal transduction', 'metabolic process', 'catabolic process', 'signaling', 'protein
173 modification process', 'macromolecule modification', 'primary metabolic process', 'regulation of
174 biological process', 'regulation of cellular process', 'response to stimulus', 'cellular response to
175 stimulus', 'biological regulation', 'cell wall organization or biogenesis', and 'organic substance
176 metabolic process'; MF category: 'lipid binding', 'hydrolase activity, acting on glycosyl bonds',
177 'enzyme regulator activity', and 'molecular function regulator'; and CC category: 'nucleus'. Thus,
178 the GO terms enriched at this stage were drastically different from those enriched at the other stages.

179 **3.4 Genetic Mechanisms Underlying Blooming in Flowering Cherry**

180 To gain insights into the genetic mechanisms regulating blooming in flowering cherry, we focused on
181 genes categorized in four functional categories (Figure 2, Supplementary Table S3): 1) transcription
182 factor genes; 2) phytohormone-related genes; 3) transporter and aquaporin genes; and 4) cell wall-
183 related genes.

184 **3.4.1 Transcription factor genes**

185 Genes encoding three types of transcription factors that trigger blooming were predominant during
186 the flowering period. The *MYB* transcription factor genes were overrepresented from 4 to 3 WBF,
187 while the ethylene-responsive transcription factor (*ERF*) genes and *NAC* transcription factor genes
188 were expressed at 4 and 3 WBF, respectively.

189 **3.4.2 Phytohormone-related genes**

190 Genes involved in biosynthesis and signal transduction pathways of gibberellin, ethylene, and
191 cytokinin were enriched at 3–5 WBF, 3–4 WBF, and 3 WBF, respectively. Auxin-related genes were,
192 on the other hand, expressed at 1 WBF and at the flowering days.

193 **3.4.3 Transporter and aquaporin genes**

194 Genes encoding sugar and inorganic transporters and aquaporins that affect the turgor and osmotic
195 pressure of cells and vacuoles were constitutively expressed from 5 WBF to the day for flowering.
196 Among these, genes encoding 10 types of transporters (ABC transporter G family members,
197 bidirectional sugar transporters, cationic amino acid transporters, lysine histidine transporters,
198 organic cation/carnitine transporters, phosphate transporters, potassium transporters, UDP-galactose
199 transporters, vacuolar amino acid transporters, and zinc transporters) were overrepresented at 3 WBF.
200 Aquaporin genes were overrepresented at 1–2 WAF.

3.4.4 Cell wall-related genes

Cell wall organization followed by cell expansions actuate petal movement, leading to flower opening. Genes encoding polysaccharide biosynthesis enzymes including UDP-glycosyltransferases and cellulose synthases were expressed from 5 WBF to the day for flowering. Genes encoding fasciclin-like arabinogalactan proteins were mainly expressed at the initial stage of the flowering period, while those encoding arabinogalactan peptides were expressed at the flowering days. Pectinesterase genes were overrepresented at the flowering days.

4   Discussion

Time-course RNA-seq analysis revealed the dynamics of the transcriptome of floral buds and flowers of the flowering cherry cultivar 'Somei-Yoshino'. Subsequent WGCNA of the RNA-seq data indicated that 1,541 genes belonging to seven modules were involved in blooming at stages from 4–5 WBF to 1–2 WAF (Figure 2). In accordance with the functional annotations of these genes, those involved in floral bud and open-flower development as well as blooming were identified in flowering cherry (Figure 3). Floral bud development potentially occurs from 4–5 WBF, with the gibberellin-mediated activation of genes controlling sugar and polysaccharide metabolism. Subsequently, at 3–4 WBF, ethylene and cytokinin likely accelerate the expression of *MYB*, *ERF*, and *NAC* transcription factor genes, resulting in the initiation of cell wall organization and cell division (Nakano et al., 2015). Some *MYB* and *NAC* genes are also involved in gametophyte or ovule integument development (Kunieda et al., 2008; Renak et al., 2012). According to the observations in sweet cherry (*Prunus avium*), the 3–4 WBF timepoint corresponds to the formation of tetrads or microspores after meiosis in the pollen (Fadon et al., 2019). At this stage, amino acid metabolism is also upregulated. Metabolome analysis of the flower buds of Japanese apricot (*Prunus mume*) revealed significant changes in the abundance of several amino acids before the bud break stage (Zhuang et al., 2015). Amino acids play important roles in bud break. At 1–2 WBF, the reproduction process, together with metabolism, is likely activated. Auxin-related genes were also activated at this stage in this current study. On the flowering day, we speculate that the auxin signal alters pectinesterase gene expression to initiate cell wall remodeling, leading to rapid petal enlargement and movement, which is the most dynamic movement in flowering plants (Habu and Tao, 2014; Nakano et al., 2015). The high-level expression of aquaporin genes, which encode water channel proteins localized to the plasma membrane or tonoplast membranes, on the day of flowering might facilitate the water flux into the petals to allow smooth petal movement (Azad et al., 2013). The relationship between auxin signaling and pectinesterase gene expression has been demonstrated in the ripening process of strawberry fruit (Castillejo et al., 2004). Similar hormonal regulations might be conserved in the Rosaceae species. The role of the auxin signal in the dynamic movement of petals during flower opening needs further investigation. After blooming, genes involved in the stimulus response and cell-to-cell communication are expressed the flower, especially in the pistil and ovary, which might reflect the pollination and fertilization process, even though 'Somei-Yoshino' is self-incompatible. Similar transcriptome profiles were observed in other *Prunus* species (Habu and Tao, 2014; Gómez et al., 2019; Iqbal et al., 2021).

Because the flowering time of 'Somei-Yoshino' is important for the tourism industry in Japan, methods for predicting the flowering date of this cultivar have been developed in accordance with the cumulative temperature before flower opening (Aono and Murakami, 2017). Gene expression analysis could upgrade the current forecasting model, for which genes expressed on specific days before flowering could be employed as diagnostic markers. However, practical RNA quantification methods, such as real-time quantitative PCR and the next-generation sequencing (NGS)-based RNA-seq, are time-consuming and labor-intensive. Since diagnostic markers can be detected easily within

247 a short time, easy-to-use target-RNA qualifying methods are required to predict the flowering time in
248 'Somei-Yoshino', based on gene expression data.

249 Comparative time-course transcriptome analysis would enable the selection of high-confidence
250 diagnostic markers for various purposes. For example, time-course transcriptome analysis can be
251 applied to floral buds of vegetable and fruit crops and cut flowers. Moreover, this comparative
252 analysis is not only limited to the selection of markers for forecasting flowering time but is also
253 applicable to the prediction of disease, fertilization, and harvest time in the field. Indeed,
254 transcriptome profiling has been employed to monitor nutritional responses and adaptation in rice
255 (Takehisa and Sato, 2021).

256 Overall, this study provides insights into the genetic mechanisms controlling petal movement and
257 blooming in cherry. Further studies are required to connect the genetic insights with the physiological
258 mechanisms (van Doorn and Van Meeteren, 2003). Once the connection or relationship is validated,
259 the transcriptome-based prediction would serve as a powerful tool for monitoring the plant phenotype
260 under controlled cultivation conditions as well as in the field. For example, genome-based prediction
261 has been used to predict progeny phenotypes in breeding programs. The transcriptome- and genome-
262 based predictions would promise high-confidence forecasting plant traits in the near and distant
263 future, respectively.

264 **5   Conflict of Interest**

265 The authors declare that the research was conducted in the absence of any commercial or financial
266 relationships that could be construed as a potential conflict of interest.

267 **6   Author Contributions**

268 KS, TE, and AI conceived this study and collected the samples; KS and SI performed experiments
269 and collected data; KS, TE, and AI analyzed and interpreted the data; KS wrote the manuscript with
270 contributions from TE; and all authors read and approved the final manuscript.

271 **7   Funding**

273 **8   Acknowledgments**

276 **9   References**

277 Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J.
278 (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.
279 *Nucleic Acids Res* 25**,** 3389-3402.

280 Aono, Y., and Murakami, N. (2017). A simplified method to estimate cherry blossom phenology
281 considering temperature during endodormancy process. *Climate in Biosphere* 17**,** 25-33.

282 Azad, A.K., Hanawa, R., Ishikawa, T., Sawa, Y., and Shibata, H. (2013). Expression profiles of
283 aquaporin homologues and petal movement during petal development in Tulipa gesneriana. *Physiol*
284 *Plant* 148**,** 397-407.

285 Castillejo, C., De La Fuente, J.I., Iannetta, P., Botella, M.A., and Valpuesta, V. (2004). Pectin
286 esterase gene family in strawberry fruit: study of FaPE1, a ripening-specific isoform. *J Exp Bot* 55**,**
287 909-918.

288 Fadon, E., Herrero, M., and Rodrigo, J. (2019). Anther and pollen development in sweet cherry
289 (Prunus avium L.) in relation to winter dormancy. *Protoplasma* 256**,** 733-744.

290 Frazee, A.C., Pertea, G., Jaffe, A.E., Langmead, B., Salzberg, S.L., and Leek, J.T. (2015). Ballgown
291 bridges the gap between transcriptome assembly and expression analysis. *Nat Biotechnol* 33**,** 243-
292 246.

293 Gómez, E.M., Buti, M., Sargent, D.J., Dicenta, F., and Ortega, E. (2019). Transcriptomic analysis of
294 pollen-pistil interactions in almond (Prunus dulcis) identifies candidate genes for components of
295 gametophytic self-incompatibility. *Tree Genetics & Genomes* 15.

296 Habu, T., and Tao, R. (2014). Transcriptome Analysis of Self- and Cross-pollinated Pistils of
297 Japanese Apricot (Prunus mume Sieb. et Zucc.). *Journal of the Japanese Society for Horticultural*
298 *Science* 83**,** 95-107.

299 Higuchi, Y. (2018). Florigen and anti-florigen: flowering regulation in horticultural crops. *Breed Sci*
300 68**,** 109-118.

301 Huntley, R.P., Binns, D., Dimmer, E., Barrell, D., O'donovan, C., and Apweiler, R. (2009).
302 QuickGO: a user tutorial for the web-based Gene Ontology browser. *Database (Oxford)* 2009**,**
303 bap010.

304 Iketani, H., Ohta, S., Kawahara, T., Katsuki, T., Mase, N., Sato, Y., and Yamamoto, T. (2007).
305 Analyses of Clonal Status in 'Somei-yoshino' and Confirmation of Genealogical Record in Other
306 Cultivars of Prunus × yedoensis by Microsatellite Markers. *Breeding Science* 57**,** 1-6.

307 Innan, H., Terauchi, R., Miyashita, N.T., and Tsunewaki, K. (1995). DNA fingerprinting study on the
308 intraspecific variation and the origin of Prunus yedoensis (Someiyoshino). *Jpn J Genet* 70**,** 185-196.

309 Iqbal, S., Pan, Z., Hayat, F., Bai, Y., Coulibaly, D., Ali, S., Ni, X., Shi, T., and Gao, Z. (2021).
310 Comprehensive transcriptome profiling to identify genes involved in pistil abortion of Japanese
311 apricot. *Physiol Mol Biol Plants* 27**,** 1191-1204.

312 Izawa, T., Takahashi, Y., and Yano, M. (2003). Comparative biology comes into bloom: genomic
313 and genetic comparison of flowering pathways in rice and Arabidopsis. *Current Opinion in Plant*
314 *Biology* 6**,** 113-120.

315 Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory
316 requirements. *Nat Methods* 12**,** 357-360.

317 Kunieda, T., Mitsuda, N., Ohme-Takagi, M., Takeda, S., Aida, M., Tasaka, M., Kondo, M.,
318 Nishimura, M., and Hara-Nishimura, I. (2008). NAC family proteins NARS1/NAC2 and

319 NARS2/NAM in the outer integument regulate embryogenesis in Arabidopsis. *Plant Cell* 20**,** 2631-
320 2642.

321 Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network
322 analysis. *BMC Bioinformatics* 9**,** 559.

323 Nakamura, I., Takahashi, H., Ohta, S., Moriizumi, T., Hanashiro, Y., Sato, Y., and Mill, M. (2015).
324 Origin of *Prunus* x *yedoenins* 'Somei-yoshino' based on sequence analysis of *PolA1* gene. *Adv Hort*
325 *Sci* 29**,** 17-23.

326 Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N.A., and Ohtani, M. (2015). NAC-MYB-based
327 transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front Plant Sci* 6**,** 288.

328 Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression
329 analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* 11**,** 1650-1667.

330 Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015).
331 StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*
332 33**,** 290-295.

333 Renak, D., Dupl'akova, N., and Honys, D. (2012). Wide-scale screening of T-DNA lines for
334 transcription factor genes affecting male gametophyte development in Arabidopsis. *Sex Plant Reprod*
335 25**,** 39-60.

336 Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets.
337 *Bioinformatics* 27**,** 863-864.

338 Shirasawa, K., Esumi, T., Hirakawa, H., Tanaka, H., Itai, A., Ghelfi, A., Nagasaki, H., and Isobe, S.
339 (2019). Phased genome sequence of an interspecific hybrid flowering cherry, 'Somei-Yoshino'
340 (Cerasus x yedoensis). *DNA Res* 26**,** 379-389.

341 Takehisa, H., and Sato, Y. (2021). Transcriptome-based approaches for clarification of nutritional
342 responses and improvement of crop production. *Breed Sci* 71**,** 76-88.

343 Takenaka, Y. (1963). The Origin of the Yoshino Cherry Tree. *Journal of Heredity* 54**,** 207-211.

344 Van Doorn, W.G., and Van Meeteren, U. (2003). Flower opening and closure: a review. *J Exp Bot*
345 54**,** 1801-1812.

346 Yamane, H. (2014). Regulation of Bud Dormancy and Bud Break in Japanese Apricot (Prunus mume
347 Siebold ^|^amp; Zucc.) and Peach [Prunus persica (L.) Batsch]: A Summary of Recent Studies.
348 *Journal of the Japanese Society for Horticultural Science* 83**,** 187-202.

349 Zhuang, W., Gao, Z., Wen, L., Huo, X., Cai, B., and Zhang, Z. (2015). Metabolic changes upon
350 flower bud break in Japanese apricot are enhanced by exogenous GA4. *Hortic Res* 2**,** 15046.

351

10   Figure Legends

Figure 1 Floral buds and open-flowers of 'Somei-Yoshino'. The numbers in each picture indicate days before (-) or after (+) the flowering day (0). Photos were taken at Kazusa DNA Research Institute in 2021.

Figure 2 Expression patterns of genes categorized into seven modules. X-axis indicates the number of days before (-) or after (+) the flowering day (0). Lines show moving averages of gene expression (sliding window size = 8, walking speed = 4). Symbols indicate biological replicates: KDRI in 2017 (circle); KDRI in 2019 (triangle); SU1 in 2019 (plus); SU2 in 2019 (cross); SU3 in 2019 (diamond); KPUA in 2019 (upside-down triangle); KPUB in 2019 (square cross); and KPUC in 2019 (star).

Figure 3 Genes enriched in the biological process (BP) category and involved in the flower opening mechanism in 'Somei-Yoshino'. Upper panels show the time frame (from 5 weeks before flowering [WBF] to 2 weeks after flowering [WAF]) and the corresponding gene modules. Arrows indicate the Gene Ontology (GO) terms enriched in the BP category (see Table 1). Boxes indicate the properties of the genes highly expressed in the seven characterized modules (see Supplementary Table S3).

366 **Table 1** Gene Ontology (GO) terms enriched at the different flowering stages.

Module	No. of genes	Stage[1]	Biological process (BP)	Molecular function (MF)	Cellular component (CC)
Dark-red	110	4–5 WBF	'developmental process' related GO terms including 'aging', 'embryo development', 'developmental maturation', 'developmental process', and 'anatomical structure development'	'oxidoreductase activity'	'plasma membrane' and 'cell periphery'
Tan	267	4 WBF	'anatomical structure formation involved in morphogenesis', 'anatomical structure development', 'sulfur compound metabolic process', 'biosynthetic process', 'developmental process', and 'cell wall organization or biogenesis'	'DNA binding', 'DNA-binding transcription factor activity', 'catalytic activity', 'oxidoreductase activity', 'transferase activity, transferring acyl groups', 'hydrolase activity, acting on glycosyl bonds', 'lyase activity', and 'transcription regulator activity'	'extracellular region'
Pink	332	3 WBF	'lipid metabolic process', 'autophagy', 'small molecule metabolic process', and 'process utilizing autophagic mechanism'	'catalytic activity', 'transporter activity', 'lipid binding', 'oxidoreductase activity', 'transferase activity, transferring acyl groups', and 'transmembrane transporter activity' activity,	'extracellular region', 'extracellular space', 'vacuole', 'plasma membrane', 'membrane', and 'cell periphery'

				transmembrane transporter activity	
Royal blue	127	2–3 WBF	'mitotic cell cycle', 'carbohydrate metabolic process', 'catabolic process', 'secondary metabolic process', 'cell division', and 'cell wall organization or biogenesis'	not detected	'extracellular region', 'cell wall', 'Golgi apparatus', 'plasma membrane', 'endomembrane system', 'membrane', 'external encapsulating structure', 'cell periphery', and 'cellular anatomical entity'
Midnight blue	176	1–2 WBF	'pollen development' and 'gametophyte development'	'transporter activity' and 'transmembrane transporter activity'	not detected
Black	457	Flowering day	'carbohydrate metabolic process', 'cytoskeleton organization', 'catabolic process', and 'cell wall organization or biogenesis'	'lipid binding', 'lyase activity', 'enzyme regulator activity', and 'molecular function regulator'	'extracellular region' and 'extracellular space'
Sky-blue	72	1–2 WAF	'carbohydrate metabolic process', 'cellular protein modification process', 'response to stress', 'cell communication', 'signal transduction', 'metabolic process', 'catabolic process', 'signaling', 'protein modification	'lipid binding', 'hydrolase activity, acting on glycosyl bonds', 'enzyme regulator activity', and 'molecular function regulator'	'nucleus'

12

process', 'macromolecule modification', 'primary metabolic process', 'regulation of biological process', 'regulation of cellular process', 'response to stimulus', 'cellular response to stimulus', 'biological regulation', 'cell wall organization or biogenesis', and 'organic substance metabolic process'

367   [1]WBF, weeks before flowering; WAF, weeks after flowering.

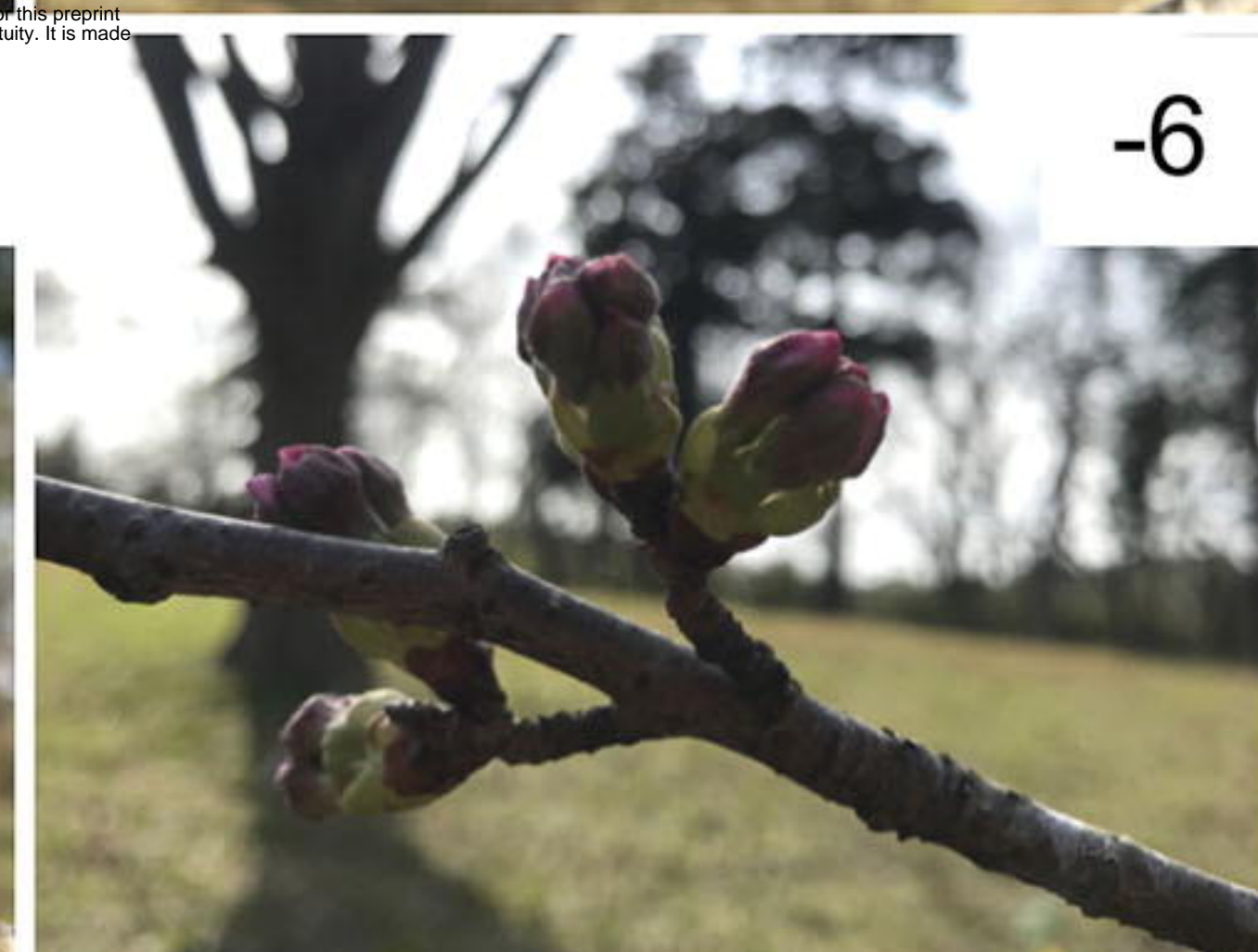| Stage	5WBF	4WBF	3WBF	2WBF	1WBF	0 (flowering day)	1WAF	2WAF

Gene module
- darkred
- pink
- midnightblue
- black
- skyblue
- tan
- royalblue

Biological process
- Development
- Reproduction
- Response to stimulus
- Cell wall organization
- Cell communication
- Metabolism

Transcription factor genes
- ERF
- NAC
- MYB

Phytohormone-related genes
- Gibberellin
- Auxin
- Ethylene
- Cytokinin

Transporter and aquaporin genes
- Sugar
- Amino acid
- K, P, Zn
- Aquaporin

Cell wall-related genes
- Fasciclin-like arabinogalactan
- Arabinogalactan peptide
- Pectinesterase