Manuscript
Click here to access/download;Manuscript;MT_Hls-manuscript_rev1.docx
Click here to view linked References

Original Article

Up-regulation of defense-related gene expressions associated with lethal growth failure in the hybrid seedlings of Japanese flowering cherry

Momi Tsuruta[1, 2] *, Chunlan Lian[2], Yuzuru Mukai[3]

[1] Forestry and Forest Products Institute, Forest Research and Management Organization, 1 Matsunosato, Tsukuba, Ibaraki 305-8687, Japan

[2] Asian Research Center for Bioresource and Environmental Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Midoricho, Nishitokyo, Tokyo 188-0002, Japan

[3] Faculty of Applied Biological Sciences, Gifu University, 1-1 Yanagido, Gifu, Gifu 501-1193, Japan

* Corresponding author (Momi Tsuruta): m.tsuruta.lfme@gmail.com

Abstract

Knowledge of post-zygotic hybrid incompatibility is essential to understand speciation. Although the genes and molecular mechanisms involved in hybrid incompatibility are being elucidated in model plants and crops, the information on woody non-model plants is lacking. In the seedlings of a cross between the most famous ornamental cherry cultivar *Cerasus* × *yedoensis* 'Somei-yoshino' and its closely related wild species *Cerasus itosakura*, we discovered a hybrid incompatibility characterized by a phenotype in which growth stops after the expansion of the first true leaves and the seedling eventually dies. To elucidate the molecular mechanisms related to this seedling necrosis, we performed a comprehensive expressed genes analysis on normal-growth and necrotic weak-growth (SW) hybrid seedlings. The RNA-seq results showed over 1,500 differentially expressed genes (DEGs) specified for the SW. Numerous genes associated with plant defense response, such as pathogenesis-related genes, and several receptor-like protein kinases were included in SW-specific up-regulated DEGs. The Gene Ontology (GO) enrichment analysis also showed the significant association of "defense response" in SW seedlings. These up-regulated defense-

related gene expressions were particularly observed in the hypocotyls. On the contrary, the reduction of ~~genes expression related to~~ photosynthesis-related gene expression as well as reduction ~~and a partial decrease~~ in the gene expressions of cell division and cell cycle at specific parts of seedlings were also observed in the SW. Our results suggest that an up-regulated defense-related gene expression suppress the meristem growth and deviation, resulting in growth failure as an autoimmune response in hybrid cherry seedlings. ~~This study is the first report inferring the molecular mechanisms of hybrid necrosis in woody plants.~~

Key words: autoimmunity, defense response, hybrid necrosis, Somei-yoshino, transcriptome

Declarations

Funding: This work was supported by the JSPS KAKENHI Grant Number JP18K14489.

Conflicts of interest: The authors declare no competing interests.

Availability of data and materials: The raw sequence data of RNA-seqs were deposited in the DDBJ Sequence Read Archive (DRA) with DRA accession No. DRA011866 and DRA011867.

Code availability: Not applicable

Authors' contributions: M.T. designed and carried out the experiments and wrote draft manuscripts. C.L. and Y.M. provided critical advice for experiments and analysis. All authors discussed the results and approved the final manuscript.

Ethics approval: Not applicable

Consent for participate: Not applicable

Consent for publication: Not applicable

Introduction

In a plant taxon that exhibits no distinct reproductive barrier before fertilization, post-zygotic reproductive isolation, such as hybrid sterility, weakness, necrosis, and hybrid breakdown (hybrid incompatibility that occurs in later generation after F_1), significantly affects the maintenance of species characteristics, i.e., speciation. Numerous studies have searched for genes involved in reproductive isolation, and several genes behind the post-zygotic reproductive isolation have been identified (Rieseberg and Willis 2007; Rieseberg and Blackman 2010). In the model plants *Arabidopsis* and rice, the expressions of genes derived from other lineages cause hybrid weakness and incompatibility (Bomblies et al. 2007; Alcázar et al. 2009; Yamamoto et al. 2010; Chae et al. 2014). Here, the heterologous allele interacts with another locus located in a cluster of receptor-like kinase (RLK) genes with leucine-rich repeat domain (LRR). The RLKs precipitate an autoimmune response and inhibit the growth of hybrid seedlings. The mechanism of hybrid incompatibility by the interaction of these two loci is understood as the Bateson–Dobzhansky–Muller model assuming that two genes originating from an independently mutated species or population cause harmful interactions (Orr 1996; Bomblies and Weigel 2007).

RLK is associated with plant defense. Especially, nucleotide-binding and LRR receptors (NB-LRR) are involved in the recognition of external stimuli, such as various pathogens, and trigger the expression of downstream defense-related genes. NB-LRR is also involved in the recognition of pathogen effectors (avirulence protein) and triggers more intense defense responses, such as the production of reactive oxygen species, accumulation of salicylic acid (SA), and even hypersensitive response (HR). They are known as microbe-associated molecular pattern ~~(MAMP)~~-triggered immunity (PTI) and effector-triggered immunity (DeYong and Innes 2006; Jones and Dangl 2006; Cesari 2018). In the hybrid seedlings of model plants, the up-regulated expressions of genes involved in the defense response, such as pathogenesis-related (PR) genes, and glutathione-S transferase (GST), were observed (Alcázar et al. 2009; Yamamoto et al. 2010). Moreover, high expressions of similar defense-related genes and transcription factors were observed in various types of growth failure, including chlorosis, HR, and severe growth abortion (SGA), in hybrid wheat seedlings (Mizuno et al. 2010, 2011; Hatano et al. 2012; Nakano et al. 2015; Takamatsu et al. 2015). Currently, the involvement of defense-related genes and autoimmunity in hybrid breakdown has been observed in several plant taxa (Wan et al. 2021). However, these studies are limited to herbal model plants and crops, and knowledge about hybrid incompatibility in woody species remains lacking.

Flowering cherry (Rosaceae, genus *Cerasus* or *Prunus* subgenus *Cerasus*) is one of the most famous ornamental tree species in Japan. Since the Edo period (18th century), numerous cultivars have been produced, many of which are considered to originate by hybridization (Ohba et al. 2007; Kato et al. 2014). Among cherry species, considerable hybrid seeds have been obtained by artificial pollinations (Watanabe and Yoshikawa 1967), and no physiological incompatibility has been observed before fertilization. In nature, frequent pollen transfer between species has been observed, and hybrid zones have been maintained in contact areas (Tochigi et al. 2021). Thus, flowering cherry is considered to be a taxon with low barriers to inter-specific hybridization. Nevertheless, distinct morphological and genetic differentiation of each species has been observed (Ohba et al. 2007; Kato et al. 2014), and more than 10 wild species are distributed in the small-island Japan. Understanding how speciation is maintained despite such frequent inter-specific gene flow requires elucidating the mechanism of post-zygotic isolation and hybrid breakdown in the generation after F_1 ~~later generations~~.

In our previous study, we observed that the lethal growth failure (necrotic) phenotype appeared in about half of the seedlings from crosses between cherry cultivar *Cerasus* × *yedoensis* (Matsum.) Masam. et S.Suzuki 'Somei-yoshino' and wild cherry species *C. itosakura* (Siebold) Masam. et S.Suzuki (syn. *C. spachiana* Lavallée ex Ed.Otto) (Tsuruta and Mukai 2015, 2019). 'Somei-yoshino' is thought to have originated from a cross between two wild species, namely, *C. speciosa* (Koidz.) H.Ohba and *C. itosakura* (Ohba et al. 2007). The genetic studies indicated the significant contribution of the two species (Kato et al. 2014; Tsuruta et al. 2017). The phased genome sequence of 'Somei-yoshino' also showed that the genome can be assigned to two kind of contigs derived from *C. speciosa* and *C. itosakura* (Shirasawa et al. 2019). Given that the seedlings from the crossing of 'Somei-yoshino' and *C. itosakura* can be considered as a backcross progeny at the species-level, the seedling growth failure may therefore be due to hybrid breakdown. From the fine mapping, one of the candidate loci for the presumable hybrid inviability of seedlings 1, *HIs-1*, has been mapped to linkage group 4 of the 'Somei-yoshino' map (Tsuruta and Mukai 2019). The candidate region corresponds to an approximate 240 Kb region of the peach genome, where RLKs and genes with LRR were concentrated. The mapping of inter-specific reproductive barriers were also reported in other Rosaceae species (inter-specific pear: Montanari et al. 2016, apple × pear: Morimoto et al. 2020). Interestingly, lethal quantitative trait loci were located in the region with clusters of pest resistance genes in the inter-specific pear hybrid (Montanari et al. 2016). From these results, the NB-LRR-mediated mechanism of hybrid incompatibility also exists in woody species as in herbal plants.

In this study, we aimed to accumulate knowledge on the physiological mechanisms of hybrid necrosis in tree species. Recently, information on genome sequencing and annotation for various Rosaceae species, including *Cerasus* and *Prunus*, has been published, and a genomic analysis environment has been established (Jung et al. 2019). We used this genome information as a reference to conduct a comprehensive gene expression analysis for the normal-growth (SN) and necrotic weak-growth (SW) seedlings of the cross between 'Somei-yoshino' and *C. itosakura*. Furthermore, with the addition of transcriptome analysis of different parts of the SW seedlings, we clarified the physiological mechanism responsible for the lethal hybrid seedlings of flowering cherries. Here, we also attempted to show that the autoimmune response due to high expressions of defense-related genes is also a common causal phenomenon in hybrid incompatibility in woody plants.

Materials and Methods

Normal and necrotic plant materials

We artificially pollinated *C. itosakura* (pollen donor) to 'Somei-yoshino' (seed parent), which were both planted in the University of Tokyo Tanashi Forest. The pollination procedures and seed handling ~~protocols followed the~~ methods were described in the work of Tsuruta and Mukai (2015). After three months of wet and chilling incubation, seeds were planted on a vermiculite bed. Of the 31 germinated seedlings, 12 (38.7%) grew normally, and 13 (41.9%) showed the same lethal phenotype as in previous studies (Tsuruta and Mukai 2015, 2019). The phenotype of the remaining six seedlings was uncertain because of insufficient growth after rooting. The lethal phenotype was characterized as the plants that grew until the emergence of first true leaves but failed to ~~further~~ develop further (Fig. 1B). Finally, the SW seedlings died after several weeks or a month. Additionally, we collected naturally pollinated seeds from a wild cherry tree, *C. jamasakura* (Siebold ex Koidz.) H.Ohba in Tanashi Forest, and seedlings were obtained after treating the seeds as described above. The necrotic phenotype was never observed in these wild seedlings.

Approximately two weeks after germination, SN, SW (necrotic fate), and control wild cherry (CJ) seedlings were collected with three replicates each, immediately frozen by liquid N_2, and stored at –80 °C until RNA extraction. Additionally, true leaves (Leaf), cotyledons (Coty), hypocotyls (Hypo), and root (Root) were collected separately (Fig. 1C) from three other weakly growing hybrid seedlings and stored in

the same manner.

RNA extraction and sequencing

From the whole plant or each part of the sample seedlings, the total RNA was extracted using the modified cetyltrimethylammonium bromide (CTAB) method. The sample was powdered in liquid N_2 using mortar and pestle, suspended in CTAB buffer (2% CTAB, 0.1 M Tris-HCl, 20 mM EDTA, 2 M NaCl, and 0.1% dithiothreitol), and incubated at 65 ℃ for 10 min. An equal volume of chloroform/isoamyl (CIA) was added and mixed gently, and the supernatant was collected by centrifugation (15,000 rpm, room temperature, 5 min). This CIA isolation procedure was repeated twice. After adding a quarter volume of 10 M LiCl and mixing gently, the solution was left at –30 ℃ for at least 2 hours. The precipitate collected by centrifugation (15,000 rpm, 4 ℃, 15 min) was dissolved in TE buffer. CIA isolation and LiCl precipitation were repeated. After rinsing with 70% ethanol, the RNA precipitate was dissolved in nuclease-free water. The quality of isolated RNA was checked by electrophoresis. All samples had sufficient amount and quality for RNA-seq except for the root sample, in which only one sample (W03-Root) passed the quality check. The qualified samples were used for further RNA-seq analysis.

RNA-seq was performed following Macrogen's regular workflow (Macrogen Inc, Seoul, South Korea). Briefly, high-throughput sequencing libraries were constructed using the TruSeq stranded mRNA Library Kit (Illumina, San Diego, CA, USA) following the quality control of RNA. Each library was sequenced by the Illumina platform (NovaSeq6000) at 100 or 150 bp pair-end reads (over four or six G bases par sample). All sequencing data have been submitted to the DDBJ Sequence Read Archive (accession No: DRA011866 and DRA011867).

RNA-seq data processing and identification of DEGs

The sequenced reads were first analyzed with fastp (Chen et al. 2018) for quality checking and trimming of low-quality bases and reads (using -3, -q 20, -l 50, -t 1, and -T 1 options). Three genome sequences, namely, the deeply re-sequenced peach genome (*Prunus persica* Genome v2.0.a1, Verde et al. 2017), sweet cherry genome (*Prunus avium* Tieton Genome v2.0.a1, Wang et al. 2020), and 'Somei-yoshino' genome (*Cerasus* × *yedoensis* Somei-yoshino Genome v1.0, Shirasawa et al. 2019) including two phasing contigs (SPA: CYEspachiana_r3.0, and SPE: CYEspeciosa_r3.0) registered in the Genome Database for Rosaceae (GDR: www.rosaceae.org, Jung et al. 2019), were used as reference genomes. The

clean reads were mapped to the reference genomes using bowtie2 (Langmead and Salzberg 2012) with default settings of the "--sensitive-local". Samtools v1.11 (Li et al. 2009) was used to convert the file format.

Uniquely mapped reads were used for gene expression quantification with featureCounts (Liao et al. 2014) with GFF files for each reference species. The differentially expressed genes (DEGs) between the two treatments of CJ, SN, and SW seedlings and between the SN seedlings and each seedling part were identified using a TCC-GUI pipeline (Su et al. 2019). TMM (Robinson and Oshlack 2010) and edgeR (Robinson et al. 2010) were used for normalization and DEG identification, respectively, with the number of iterations equal to 3, FDR cut-off at 0.05, and $P < 0.05$. Significant DEGs that were common to CJ vs. SW and SN vs. SW but not present in CJ vs. SN were determined as SW-specific DEGs (Fig. S1).

The expression levels of RNA-seq were verified by quantitative RT-PCR (qPCR) in three genes (Table S1). Actin (CYE_r3.1SPA5_g019640.1) was selected as a reference housekeeping gene. Primers for qPCR were developed using Primer-BLAST (Ye et al. 2012) using the 'Somei-yoshino' genome as a reference and are listed in Supplementary Table S1. A total of 1 µg total RNA (same as what was used for RNA-Seq) was converted to cDNA using iScript gDNA Clear cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA). Using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad), PCR reactions were conducted using CFX96 Real-Time System (Bio-Rad). The relative expression levels of each gene were calculated from the average of the three biological replicates using the efficiency correlated model (Pfaffl et al. 2002). The correlation between the relative expression level and the M-value obtained from RNA-seq was examined.

TCC-normalized expression data were also used to compare the expression levels of 12 ~~specific~~ genes (eight SW-specific up-regulated DEGs related to defense response, one up-regulated DEG unrelated to defense response, and three down-regulated DEGs in SW seedlings). Differences in the expressions among different seedling parts were tested by pairwise t-test with paired = TRUE option. Volcano plots of DEG identification were also drawn by the TCC-GUI. The Venn diagrams were drawn by R package VennDiagram v1.6.20 (Chen and Boutros 2011).

Annotation of DEGs and enrichment analysis

Gene information, such as annotations, BLAST results, InterPro, Gene Ontology (GO) terms, and GO accessions, were searched in the GDR database. The R package goseq v1.38.0 (Young et al. 2010) was used to identify significantly enriched GO terms in the SW-specific up- and down-regulated DEGs.

For data referencing of the peach genome, additional singular enrichment analysis (SEA) and

parametric analysis of gene set enrichment (PAGE) were performed by agriGO v2.0 web tool (http://systemsbiology.cau.edu.cn/agriGOv2/, Tian et al. 2017). Furthermore, the enrichment of each gene family was analyzed using the web application GenFam (www.mandadilab.com/genfam/, Bedre and Mandadi 2019).

Phasing of expressed genes and assignments of origin

We selected 8, 4, and 4 peach genes expressed in all seedlings from the *HIs-1* candidate region of chromosome 4 (Tsuruta and Mukai 2019), outside the candidate region, and another chromosome, respectively (Table S2~~1~~). Using bam data mapped to the peach genome, each haplotype allele sequence of the SN and SW seedlings was detected manually by phasing of the SNP variant with Integrative Genome Viewer v2.8.9 (Robinson et al. 2017). For each phased allele sequence, similarities with the paralog genes in the SPA and SPE genomes (*Cerasus* × *yedoensis* Somei-yoshino genome v1.0 cds) were searched using the BLAST tool on the GDR website.

Results

Gene expression profiling and particular DEGs in SW seedlings

From the RNA-seq, over 50 M sequencing reads with 8.1–13.2 G bases per seedling sample were obtained (Table S3~~1~~). Over 96.6%, 95.3%, and 97.8% of the quality filtered clean reads were mapped to the peach, sweet cherry, and 'Somei-yoshino' reference genomes, respectively (Table S3~~2~~).

Using the peach genome containing 47,089 transcribable genes as a reference, 22,218 genes were expressed in any seedling. Among genes, 19,148 (86.2%) were commonly expressed in all seedlings (Table 1). TCC analysis determined that 1,894, 3,841, and 3,638 genes were significant DEGs in the comparison of CJ vs. SN, CJ vs. SW, and SN vs. SW, respectively. A total of 1,856 genes were identified as specific DEGs in the SW seedlings (Fig. S1), and of these genes, 661 showed high expression (up-regulated DEGs), and 1,195 presented decreased expression compared with the SN and CJ seedlings (down-regulated DEGs). When the sweet cherry and 'Somei-yoshino' genomes were used as references, 2,644 (up-regulated: 935, down-regulated: 1,709) and 1,535 (up-regulated: 871, down-regulated: 664) genes were identified as SW-specific DEGs (Fig. S1). The expression pattern well corresponded with the qPCR results (r^2 = 0.90, Fig. S2).

Annotation of DEGs

From the results of the BLAST search, several plant defense-related genes were found in the up-regulated DEGs (Table 2, S4~~3~~). One of them~~gene~~ is a member of the PR gene group. The most up-regulated DEG, Prupe.8G152900 which shows similarity to the *Arabidopsis* AT2G14580.1 gene, was a member of PR1 in the peach genome referencing analysis. The paralogous PR genes FUN_030145-T1, CYE_r3.1SPA8_g020410, and CYE_r3.1SPE8_g024490 were identified as the most, second, and third most up-regulated DEGs in the sweet cherry, 'Somei-yoshino' phased SPA, and SPE genome referencing analysis, respectively. Nine PR1 homologous genes were detected in the peach genome. Of these genes, eight were significant DEGs in SN vs. SW. Furthermore, six genes were detected as SW-specific up-regulated DEGs (data not shown). Several other PR genes, such as thaumatin-like protein (TLP), a member of PR5, and Bet v1 superfamily gene belonging to PR10, were also found in the top 50 up-regulated DEGs (Table 2, S4~~3~~).

Another remarkable defense-related gene was a member of RLK. Of the 'Somei-yoshino' SPE genome genes, CYE_r3.1SPE4_g004570.1, which has RLK and LRR domains, was the most up-regulated DEG (Table ~~S3A~~ S4~~3~~D). Other LRK genes, including cysteine-rich LRK (CRK) and cysteine-rich repeat secretary protein (CRRSP), were also found in the DEG lists (Table 2, S4~~3~~). Additionally, several genes with oxidoreductase activity, such as GST, cytochrome P450 (Cyt_P450), and 2-oxoglutarate and Fe(II)-dependent oxygenase (2OG-FeII), and WRKY transcription factors were listed in the up-regulated DEGs (Table 2, S4~~3~~).

The most down-regulated DEG was the Prupe.3G208300.1 gene annotated as GDSL-like in the peach genome referencing analysis. Genes annotated as laccase 17 (LAC17), fasciclin-like arabinogalactan, and trichome birefringence-like frequently ranked the top in the down-regulated DEG list (Table S4~~3~~). The down-regulation of several photosynthesis-related genes was also observed (Table S4~~3~~).

Enrichment analyses

In the peach genome referencing analysis, 397 up- and 690 down-regulated genes that were annotated with GO terms were used for the enrichment analysis. From the results of goseq, we determined that 5, 15, and 2 GO terms belonging to biological process (BP), molecular function (MF), and cellular component (CC), respectively, were significantly enriched in the up-regulated DEGs (FDR < 0.05, Fig. 2A). Of the BP

terms, GO:0009607 "response to biotic stimulus" was identified as the most significantly enriched term in the up-regulated DEG list. The next most significant GO, i.e., GO:0006952 "defense response," was consistently detected as a ~~a significantly enriched GO term with 4.20e^{-09} FDR level, in which 16 of 99 genes were significantly highly expressed in SW seedlings. A~~ remarkably ~~significant enrichment of "defense response" was consistently observed~~ enriched GO in the sweet cherry and 'Somei-yoshino' genome referencing analyses (Fig. 2A). In addition, ~~GO:0009607 "response to biotic stimulus" was identified as the most significantly enriched term in the up-regulated DEG list.~~ two GO terms "protein phosphorylation" and "oxidation-reduction process" were also consistently enriched in three genome referencing analyses. The "abscisic acid-activated signaling pathway" term was the most enriched BP GO in the 'Somei-yoshino' genome referencing analysis, but it ~~showed~~ exhibited no significance in the peach reference (Fig. 2A). In MF and CC, GOs similar to the significant GOs in BP or associated with the results indicated~~As shown~~ in the annotation of up-regulated DEGs were significantly enriched, e.g., ~~, the significant enrichment of several GO terms, such as "oxidation-reduction process" in BP and~~ "oxidoreductase activity," "protein kinase activity," and "DNA-binding transcription factor activity" ~~in MF was also commonly detected~~ (Fig. 2A). The~~This~~ significan~~t enrichment~~ce of two plant defense-related GO terms ("response to biotic stimulus" and "defense response") was ~~consistently observed~~verified both in SEA and PAGE analyses (Figs. S3~~2~~ and S4~~3~~, respectively).

On the other hand, 15, 19, and 12 GO terms for BP, CC, and MF were significantly enriched in down-regulated DEGs of peach genome referencing analysis, respectively (Fig. 2B). Here, the most significantly enriched GO term was "photosynthesis" (GO:0015979). Additional ~~The~~ photosynthesis-related GO terms, such as ~~GO:0015979 "photosynthesis" and~~ GO:0009765 "photosynthesis, light harvesting" in BP and GO:0009522 "photosystem I" and GO:0009523 "photosystem II" in CC, were detected as significantly enriched GO terms (Fig. 2B). The "oxidation-reduction process" was also observed in the significant GO enrichment of down-regulated DEGs. Besides this, considerable enrichment of the five BP GOs (GO:0005975, GO:0007017, GO:0009664, GO:0030244, and GO:0046274) was consistently observed in the three reference analyses (Fig. 2B).

The GenFam analysis determined that 17 and 29 gene families were significantly enriched in the up- and down-regulated DEGs, respectively (FDR < 0.05, Table S5~~4~~). The results were similar with the results of DEG annotation and GO enrichment analysis. For example, a significant enrichment was found in gene families, such as GST, Bet v1, Cyt_P450, CRK, WRKY, 2OG-FeII, "cysteine-rich secretory proteins,

antigen 5, and pathogenesis-related 1 protein (CAP)," and TLP, for the SW up-regulated DEGs and "Chlorophyll a/b-binding" for the down-regulated DEGs (Table S5~~4~~). The phytocyanin, multicopper oxidase, and trichome birefringence-like gene families were also significantly enriched (Table S5).

Expression analysis for each seedling part

More than 93.2% of the sequenced reads for each seedling part were mapped to the peach genome (Table S3~~1~~). The count data of uniquely mapped read were normalized by TCC-GUI and then used for the expression pattern comparison in 12 genes. Of the nine identified SW-specific up-regulated DEGs, all of eight defense-related genes, such as Prupe.8G152900 (PR1), showed a significantly high expression in Hypo (P <0.05, Table 2, Fig. 3). Another SW-specific DEG, Prupe.7G051600 which was annotated as BG3, was highly expressed not only in the Hypo but also in Leaf. On the other hand, the expressions of down-regulated DEGs in SW seedlings, namely, Prupe.3G208300 (GDSL), Prupe.3G076300 (2OG-FeII), and Prupe.6G257700 (LAC17), were low across all parts (Fig. 3).

A comparison of SN seedlings with each seedling part identified 4,996, 2,102, and 5,405 DEGs in Coty, Leaf, and Hypo, respectively (Fig. S5~~4~~). A total of 7,659 DEGs were significantly enriched in any seedling part, and 1742 of them were common to the SW-specific DEGs. From the PAGE results, 245 GO terms (BP: 142, CC: 33, and MF: 70) were significantly enriched in any seedling part (Coty: 92, Leaf: 193, and Hypo: 172; Fig. 4, S6~~5~~). Of these GO terms, 62 were commonly enriched in all parts. The enrichment of "defense response" and "response to biotic stimulus" GO terms, which showed significant enrichment in the above comparison of SN vs. SW, was the significant in the Hypo (FDR: 0.00037 and $5.60e^{-05}$, respectively) but not in other parts of SW seedlings (Fig. 4), similar to the expression level data. On the other hand, GO terms related to photosynthesis were the most down-regulated in Hypo (Fig. 4, S6~~5~~). In addition, the significantly down-regulated enrichment of several GO terms associated with cell division, such as "cell cycle" and "mitotic cell cycle," was observed in Leaf (Fig. 4).

Origin of phased alleles

The phasing of the 16 expressed genes determined two to five haplotypes for each gene (Table S6~~5~~). From sequence homology, the haplotypes were divided into two groups, namely, allele_A and allele_B in Table S6~~5~~, respectively. The Allele_A was more homologous to the paralogous gene in the SPA genome compared to the SPE genome. On the other hand, Allele_B tended to have a higher identity to SPE genome

genes. The exceptions, for example, were Prupe.4G062600.1 and Prupe.4G063700.1 genes which had no paralogous sequences in SPE genome (Table S2~~1~~). In the 12 genes located in chromosome 4, SN seedlings always had two allele_As, meanwhile, the SW seedlings had one allele_A (derived from paternal *C. itosakura*) and another allele (allele_B) derived from maternal 'Somei-yoshino' (Table S6~~5~~). For genes located on another chromosome, the heterozygosity~~te~~ of allele_A and Allele_B was observed (Table S6~~5~~).

Discussion

Characteristics of gene expression in the hybrid seedlings of 'Somei-yoshino' and *C. itosakura*

Here, we performed a comprehensive gene expression analysis of hybrid seedlings of 'Somei-yoshino' and *C. itosakura* based on three reference genomes (Verde et al. 2017; Wang et al. 2020; Shirasawa et al. 2019). The phased 'Somei-yoshino' genomes, which are rich in sequence and structural polymorphisms derived from two ancestral species, aided in the mapping of hybrid seedling RNA-seq data. However, the assembly and annotation of this genome are still first versions, and the gap-free chromosome length is shorter compared with that of other plant genomes (Shirasawa et al. 2019, 2021). Therefore, the genomes of peach and closely related sweet cherry with updated assemblies and annotations were added to the analysis. The peach genome is well re-sequenced and annotated and has a wealth of database information applicable to GO analysis tools, such as AgriGO and GenFam. The results for mapping and the number of expressed genes were reflected as the characteristics of the three references. Although the mapping rate to the 'Somei-yoshino' genomes was high (Table S3~~1~~), the number of DEGs annotated with GO terms was higher in peach and sweet cherry genome analyses (data not shown). Through three referencing functional analyses (DEG annotations, enrichment analyses for the GO term and gene family), we consistently identified the gene expression features of the cherry hybrid seedlings, that is, an evidence of significant up-regulation of the expression of genes associated with defense response in the SW seedlings. On the other hand, a decreased gene expression in the growth failure seedlings was characterized by photosynthesis-related genes.

High expression of defense-related genes in the SW hybrid seedlings

Various PR genes associated with plant defense against pathogens (van Loon et al. 2006) were found in the up-regulated DEG list specified in the growth failure seedlings. In particular, PR1 homologous to

Arabidopsis basic PR1-like genes was detected as a major up-regulated DEG of SW seedlings in every analyses. PR1 is expressed by pathogen infection and associated with plant defense (van Loon et al. 2006). Another group of PR protein PR5, also referred as TLP, has also been indicated to participate in pathogen resistance (van Loon et al. 2006). The involvement of these PR1 and PR5 genes in the defense response has also been shown in peach (Sherif et al. 2012). Bet v1 family gene, a member of PR10, is also associated with plant defense (Liu and Ekramoddoullah 2006). The pathogen responsible for the expression of PR10 genes was also identified in several Rosaceae species (apples: Pühringer et al. 2000; Ziadi et al. 2001, peach: Zubini et al. 2009, plum: El-kereamy et al. 2009). They may be paralogs to Bet v1 genes identified in our DEG list.

The NB-LRR gene CYE_r3.1SPE4_g004570.1 was identified as the most up-regulated DEG in the 'Somei-yoshino' SPE genome. NB-LRR, which triggered an autoimmune response in *Arabidopsis* and rice, was assumed to promote the participation of other defense-related genes (Bomblies et al. 2007; Yamamoto et al. 2010). Although several genes having LRR domain were significantly up-regulated in the SW seedlings, the NB-LRR gene family exhibited no substantial enrichment in the peach genome referencing analysis (7/558 genes, FDR = 0.927; not significant in the GenFam result). In addition, paralog gene of CYE_r3.1SPE4_g004570.1 in peach did not become DEG (data not shown). Additional consideration is needed to determine whether LRR-NB triggers growth failure in flowering cherries. Meanwhile, other RLKs, namely, CRK and CRRSP, were also found in SW up-regulated DEGs. These genes are also associated with pathogen resistance and cell death in *Arabidopsis* (Acharya et al. 2007; Yadeta et al. 2017) and in apples responding to fungal inoculation (Zuo et al. 2020). High expressions of genes associated with oxidation-reduction, such as GST, Cyt_P450, and 2OG-FeII, were also found in the SW seedlings. These genes are involved in the detoxification of peroxides and are highly expressed during the immune response in plants (Mauch and Dudler 1993; Wagner et al. 2002). Therefore, GST and Cyt_P450 are occasionally used as indicators of immune responses in *Arabidopsis* and rice hybrid inviability (Alcázar et al. 2009; Yamamoto et al. 2010). The high expression of these series of defense-related genes without pathogen inoculation strongly indicates that an autoimmune response also occurs in the hybrid seedlings of flowering cherry.

The DEG list included several transcription factors called WRKYs. The WRKYs regulate various biological processes, several of which are highly expressed upon pathogen infection, and up- or down-regulate the expression of themselves, other transcription factors, and downstream defense-related genes

(Rushton et al. 2010; Chen et al. 2019). GenFam analysis showed that WRKY domains were significantly enriched in SW seedlings, where 9 out of 93 genes were detected as up-regulated DEGs. Among these genes, WRKYs with high homology to *Arabidopsis* WRKY40, WRKY70, and WRKY75 were remarkably up-regulated in the SW seedlings. AtWRKY70 increases the PR gene expression and regulates plant defense and senescence through ~~salicylic acid (~~SA~~)~~ mediated pathways (Li et al. 2004, 2006; Ülker et al. 2007). Strawberry FaWRKY1 is homologous to AtWRKY75, highly expressed upon pathogen infection, and associated with the accumulation of peroxides and increased GST expression (Encinas-Villarejo et al. 2009). Meanwhile, AtWRKY40 is also associated with PTI ~~MAMP-triggered immunity~~ but suppresses the resistance-like feedback response (Xu et al. 2006; Birkenbihl et al. 2017). The upward expression of the series of defense-related genes may be coordinated by these transcription factors.

The excess of these defense responses probably resulted in the reduction of other physiological activities (Tian et al. 2003; van Hulten et al. 2006). Various gene families, including LAC17, were significantly suppressed in the SW seedlings. Photosynthesis-related GO terms also showed significant enrichment in the down-regulated DEGs of SW seedlings. The up-regulation of the defense-related genes and the decrease in bioactivities strongly suggested that autoimmune response triggers the growth arrest in hybrid cherry seedlings, as shown in previous studies.

Gene expression in different seedling parts

~~Quantitative PCR is often used to quantify and validate the expression of genes of interest. However, in this method, the number of genes to be tested is limited in labor. Thus, we~~ We performed additional RNA-seq on different parts of the SW seedlings to identify the expression site of the overall defense-related genes described above. The results showed that the increased expression of genes related to defense response consistently occurred in the hypocotyl. The comprehensive analysis further captured a site-specific gene expression that was not found in the comparisons of whole plant of SN and SW seedlings. Decreased cell division or differentiation-related gene expression was confirmed in the leaf samples. The growth failure of hybrid seedlings in flowering cherry is characterized by the absence of further development and differentiation of shoots after first true leaf expansion. Therefore, cell division-related processes are expected to be inhibited in shoot apical meristems. For example, wheat type II necrosis was characterized by a remarkable reduction of cell division in crown tissues (Mizuno et al. 2011). A decrease in cell cycle-related gene expressions was also observed in the crown tissues of wheat SGA line, in which

shoot growth stopped after the appearance of the second or third leaf (Hatano et al. 2012). The seedling phenotype observed in necrotic cherry seedlings was similar to the morphological characteristics in the SGA wheat.

The SW seedling-specific enrichment of defense-related genes was observed especially in the hypocotyl~~Hypo~~, just below the shoot apex. This result may be related to growth arrest, but the detailed mechanism is still unknown. A sampling that is more specific to the seedling parts, such as the shoot apex alone, and differential time sampling will be useful in elucidating the growth failure mechanism. Furthermore, the concentration of certain plant hormones, such as abscisic acid, SA, and jasmonic acid, acting as signaling pathways through defense response, peroxide accumulation, and morphological observation of cell division in the shoot apex will be required.

Candidate genes for trigger of defense response

A locus involved in seedling necrosis was mapped to the candidate region of LG4 in our previous fine mapping (Tsuruta and Mukai 2019). In the peach genome, the region expanded by approximately 240 Kb with 45 transcribable genes, but whether it was the causative gene was unclear. The candidate region included several RLKs and genes with the LRR domain. Although several of these genes were highly homologous (86%–88%) to CYE_r3.1SPE4_g004570.1, which was the most significant up-regulated DEG in the 'Somei-yoshino' genome referencing analysis, they were never determined to be DEGs (e.g., Prupe.4G062600 and Ppe4G063700; data not shown). Two genes in the candidate region were identified as minor DEGs (Prupe.4G060400 and Prupe.4G064100), but their functions are unknown. In addition, CYE_r3.1SPE4_g004570.1 was located outside the candidate region. Several gaps remained in the candidate region of the 'Somei-yoshino' genome sequence and aligning the region with the SPE and SPA genomes or genomes in closely related species was difficult (Table S2~~1~~; Tsuruta and Mukai 2019). These features suggest the presence of structural polymorphisms in the genome sequences of the candidate region. The genome region of RLKs frequently highly varied, including copy number variants (Bergelson et al. 2001; Chae et al. 2014). In the case of rice hybrid breakdown, the comparison of the sequences of causative NB-LRR clusters between homologous chromosomes was difficult (Yamamoto et al. 2010). Thus, the genome sequence of 'Somei-yoshino' in the candidate region will also need to be re-sequenced carefully.

The estimation of the allele origin of several genes located on the *HIs-1* candidate region is also suggestive. In the SN seedlings, both alleles were derived from *C. itosakura*, whereas in SW seedlings,

heterozygotes of *C. itosakura*- and *C. speciosa*-derived alleles were observed. This genotype is compatible with the seedling inviability, which was thought to be an incompatibility between *C. speciosa*, an ancestor of 'Somei-yoshino', and *C. itosakura*, the cross parent (Tsuruta and Mukai 2015); here, a causable locus was mapped to the SPE genome. These results reminded us that a candidate CYE_r3.1SPE4_g004570.1 was located on the SPE genome and suggested that the NB-LRR remains the most likely candidate for triggering hybrid incompatibility in flowering cherries. Otherwise, from the complex origin of 'Somei-yoshino' (not a simple F_1 hybrid of the two ancestor species but also a line derived from complicated crosses discussed in Tsuruta et al. 2017), possibly, the partial genome introgression of different lineages resulted in an autoimmune response of defense-related genes due to epistasis disruption. The causative gene of seedling necrosis may still need to be discussed. On the other hand, the up-regulation of defense-related gene expression probably commonly leads to hybrid weakness and hybrid breakdown, which is a phenomenon observed not only in herbal plants but also in woody species. Diversified defense-related genes may also contribute to the maintenance of speciation.

Acknowledgments

We would like to thank the staff of The University of Tokyo Tanashi Forest for providing the cherry trees for artificial pollination, Dr. S-I. Miyazawa, Senior Researcher of Forestry and Forest Products Research Institute, for the advising in qPCR analysis, and as Dr. D. Tsugama, Assistant Professor of ARC-BRES, the University of Tokyo, for the kind support in the first step of next generation sequencing analysis. This work was supported by the JSPS KAKENHI, Grant-in-Aid for Early-Career Scientists (JP18K14489).

Data Archiving Statement

The raw sequence data of RNA-seqs were deposited in the DDBJ Sequence Read Archive (DRA) with accession numbers DRA 011866 and DRA011867.

References

Acharya BR, Raina S, Maqbool SB, Jagadeeswaran G, Mosher SL, Appel HM, Schultz JC, Klessig DF, Raina R (2007) Overexpression of CRK13, an Arabidopsis cysteine-rich receptor-like kinase, results in enhanced resistance to *Pseudomonas syringae*. Plant J 50:488–499. https://doi.org/10.1111/j.1365-313X.2007.03064.x

Alcázar R, García AV, Parker JE, Reymond M (2009) Incremental steps toward incompatibility revealed by Arabidopsis epistatic interactions modulating salicylic acid pathway activation. Proc Natl Acad Sci USA 106:334–339. https://doi.org/10.1073/pnas.0811734106

Bedre R, Mandadi K (2019) GenFam: A web application and database for gene family-based classification and functional enrichment analysis. Plant Direct 3:e00191. https://doi.org/10.1101/pld3.191

Bergelson J, Kreitman M, Stahl EA, Tian D (2001) Evolutionary dynamics of plant *R*-genes. Science 292:2281–2285. https://doi.org/10.1126/science.1061337

Birkenbihl RP, Kracher B, Somssich IE (2017) Induced genome-wide binding of three Arabidopsis WRKY transcription factors during early MAMP-triggered immunity. Plant Cell 29:20–38. https://doi.org/10.1105/tpc.16.00681

Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, Weigel D (2007) Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol 5:e236. https://doi.org/10.1371/journal.pbio.0050236

Bomblies K, Weigel D (2007) Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. Nat Rev Genet 8:382–393. https://doi.org/10.1038/nrg2082

Cesari S (2018) Multiple strategies for pathogen perception by plant immune receptors. New Phytol 219:17–24. https://doi.org/10.1111/nph.14877

Chae E, Bomblies K, Kim ST, Karelina D, Zaidem M, Ossowski S, Martín-Pizarro C, Laitinen RAE, Rowan BA, Tenenboim H, et al. (2014) Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. Cell 159:1341–1351. https://doi.org/10.1016/j.cell.2014.10.049

Chen H, Boutros PC (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bbioinformatics 12:35. https://doi.org/10.1186/1471-2105-12-35

Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chen X, Li C, Wang H, Guo Z (2019) WRKY transcription factors: evolution, binding, and action. Phytopathol Res 1:13. https://doi.org/10.1186/s42483-019-0022-x

DeYoung BJ, Innes RW (2006) Plant NBS-LRR proteins in pathogen sensing and host defense. Nat Immunol 7:1243–1249. https://doi.org/10.1038/ni1410

El-kereamy A, Jayasankar S, Taheri A, Errampalli D, Paliyath G (2009) Expression analysis of a plum pathogenesis related 10 (PR10) protein during brown rot infection. Plant Cell Rep 28:95–102.

https://doi.org/10.1007/s00299-008-0612-z

Encinas-Villarejo S, Maldonado AM, Amil-Ruiz F, de los Santos B, Romero F, Pliego-Alfaro F, Muñoz-Blanco J, Caballero JL (2009) Evidence for a positive regulatory role of strawberry (*Fragaria* × *ananassa*) Fa WRKY1 and *Arabidopsis* At WRKY75 proteins in resistance. J Exp Bot 60:3043–3065. https://doi.org/10.1093/jxb/erp152

Hatano H, Mizuno N, Matsuda R, Shitsukawa N, Park P, Takumi S (2012) Dysfunction of mitotic cell division at shoot apices triggered severe growth abortion in interspecific hybrids between tetraploid wheat and *Aegilops tauschii*. New Phytol 194:1143–1154. https://doi.org/10.1111/j.1469-8137.2012.04125.x

Jones JD, Dangl JL (2006) The plant immune system. Nature 444:323–329. https://doi.org/10.1038/ nature05286

Jung S, Lee T, Cheng CH, Buble K, Zheng P, Yu J, Humann J, Ficklin SP, Gasic K, Scott K, et al. (2019) 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. Nucleic Acids Res 47:D1137–D1145. https://doi.org/10.1093/nar/gky1000

Kato S, Matsumoto A, Yoshimura K, Katsuki T, Iwamoto K, Kawahara T, Mukai Y, Tsuda Y, Ishio S, Nakamura K et al (2014) Origins of Japanese flowering cherry (*Prunus* subgenus *Cerasus*) cultivars revealed using nuclear SSR markers. Tree Genet Genome 10:477–487. https://doi.org/10.1007/s11295-014-0697-1

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357–359. https://doi.org/10.1038/nmeth.1923

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li J, Brader G, Palva ET (2004) The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. Plant Cell 16:319–331. https://doi.org/10.1105/tpc.016980

Li J, Brader G, Kariola T, Palva ET (2006) WRKY70 modulates the selection of signaling pathways in plant defense. Plant J 46:477–491. https://doi.org/10.1111/j.1365-313X.2006.02712.x

Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30: 923–930. https://doi.org/10.1093/bioinformatics/btt656

Liu JJ, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. Physiol Mol Plant Pathol 68:3–13. https://doi.org/10.1016/j.pmpp.2006.06.004

Mauch F, Dudler R (1993) Differential induction of distinct glutathione-S-transferases of wheat by xenobiotics and by pathogen attack. Plant Physiol 102:1193–1201. https://doi.org/10.1104/pp.102.4.1193

Mizuno N, Hosogi N, Park P, Takumi S (2010) Hypersensitive response-like reaction is associated with hybrid necrosis in interspecific crosses between tetraploid wheat and *Aegilops tauschii* Coss. PLoS ONE 5:e11326. https://doi.org/10.1371/journal.pone.0011326

Mizuno N, Shitsukawa N, Hosogi N, Park P, Takumi S (2011) Autoimmune response and repression of miotic cell division occur in inter-specific crosses between tetraploid wheat and *Aegilops tauschii* Coss. That show low temperature-induced hybrid necrosis. Plant J 68:114–128. https://doi.org/10.1111/j.1365-313X.2011.04667.x

Montanari S, Brewer L, Lamberts R, Velasco R, Malnoy M, Perchepied L, Guérif P, Durel CE, Bus VGM, Gardiner SE, Chagné D (2016) Genome mapping of postzygotic hybrid necrosis in an interspecific pear population. Hortic Res 3:15064. https://doi.org/10.1038/hortres.2015.64

Morimoto T, Inaoka M, Banno K, Itai A (2020) Genetic mapping of a locus controlling the intergeneric hybridization barrier between apple and pear. Tree Genet Genome 16:5. https://doi.org/10.1007/s11295-019-1397-7

Nakano H, Mizuno N, Tosa Y, Yoshida K, Park P, Takumi S (2015) Accelerated senescence and enhanced disease resistance in hybrid chlorosis lines derived from interspecific crosses between tetraploid wheat and *Aegilops tauschii*. PLoS ONE 10:e0121583. https://doi.org/10.1371/journal.pone.0121583

Ohba H, Kawasaki T, Tanaka H, Kihara H (2007) Flowering cherries of Japan, Newth edn. Yama-kei Publishers, Tokyo (in Japanese).

Orr HA (1996) Dobzhansky, Bateson, and the genetics of speciation. Genetics 144:1331–1335.

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30:e36. https://doi.org/10.1093/nar/30.9.e36

Pühringer H, Moll D, Hoffmann-Sommergruber K, Watillon B, Katinger H, da Câmara Machado ML (2000) The promoter of an apple *Ypr10* gene, encoding the major allergen Mal d 1, is stress- and pathogen-inducible. Plant Sci 152:35–50. https://doi.org/10.1016/S0168-9452(99)00222-8

Rieseberg LH, Blackman BK (2010) Speciation genes in plants. Ann Bot 106:439–455. https://doi.org/10.1093/aob/mcq126

Rieseberg LH, Willis JH (2007) Plant speciation. Science 317:910–914. https://doi.org/10.1126/science.1137729

Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP (2017) Variant review with the integrative genomics viewer. Cancer Res 77:e31–e34. https://doi.org/10.1158/0008-5472.CAN-17-0337

Robinson MD, McCarthy DJ, Smyth GK. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. https://doi.org/10.1093/bioinformatics/btp616

Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11:R25. https://doi.org/10.1186/gb-2010-11-3-r25

Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010) WRKY transcription factors. Trends Plant Sci 15:247–258. https://doi.org/10.1016/j.tplants.2010.02.006

Sherif S, Paliyath G, Jayasankar S (2012) Molecular characterization of peach PR genes and their induction kinetics in response to bacterial infection and signaling molecules. Plant Cell Rep 31:697–711. https://doi.org/10.1007/s00299-011-1188-6

Shirasawa K, Esumi T, Hirakawa H, Tanaka H, Itai A, Ghelfi A, Nagasaki H, Isobe S (2019) Phased genome sequence of an interspecific hybrid flowering cherry, 'Somei-Yoshino' (*Cerasus* × *yedoensis*). DNA Res 26:379–389. https://doi.org/10.1093/dnares/dsz016

Shirasawa K, Harada D, Hirakawa H, Isobe S, Kole C (2021) Chromosome-level *de novo* genome assemblies of over 100 plant species. Breed Sci 71:117–124. https://doi.org/10.1270/jsbbs.20146

Su W, Sun J, Shimizu K, Kadota K (2019) TCC-GUI: a Shiny-based application for differential expression analysis of RNA-seq count data. BMC Res Notes 12:133. https://doi.org/10.1186/s13104-019-4179-2

Takamatsu K, Iehisa JC, Nishijima R, Takumi S (2015) Comparison of gene expression profiles and responses to zinc chloride among inter-and intraspecific hybrids with growth abnormalities in wheat and its relatives. Plant Mol Biol 88:487–502. https://doi.org/10.1007/s11103-015-0338-6

Tian D, Traw MB, Chen JQ, Kreitman M, Bergelson J (2003) Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. Nature 423:74–77. https://doi.org/10.1038/nature01588

Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, Xu W, Su Z (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res 45:W122–W129. https://doi.org/10.1093/nar/gkx382

Tochigi K, Shuri K, Kikuchi S, Naoe S, Koike S, Nagamitsu T (2021) Phenological shift along an elevational gradient and dispersal of pollen and seeds maintain a hybrid zone between two cherry tree species. Plant Species Biol 36:230–245~~(in press)~~. https://doi.org/10.1111/1442-1984.12311

Tsuruta M, Mukai Y (2015) Hybrid seedling inviability locus (*HIs1*) mapped on linkage group 4 of the Japanese flowering cherry, *Cerasus* × *yedoensis* 'Somei-yoshino'. Tree Genet Genome 11:88. https://doi.org/10.1007/s11295-015-0910-x

Tsuruta M, Mukai Y (2019) Fine mapping of a locus presumably involved in hybrid inviability (*HIs-1*) between flowering cherry cultivar *Cerasus* × *yedoensis* 'Somei-yoshino' and its wild relative *C. spachiana*. Breed Sci 69:658–664. https://doi.org/10.1270/jsbbs.19078

Tsuruta M, Wang C, Kato S, Mukai Y (2017) Map based estimation of the origin of Japanese flowering cherry cultivar, *Cerasus* × *yedoensis* 'Somei-yoshino' Fujino with an assignment for each chromosome. J Japan For Soc 99:210–213 (In Japanese with English abstract). https://doi.org/10.4005/jjfs.99.210

Ülker B, Mukhtar MS, Somssich IE (2007) The WRKY70 transcription factor of *Arabidopsis* influences both the plant senescence and defense signaling pathways. Planta 226:125–137. https://doi.org/10.1007/s00425-006-0474-y

van Hulten M, Pelser M, Van Loon LC, Pieterse CM, Ton J (2006) Costs and benefits of priming for defense in *Arabidopsis*. Proc Natl Acad Sci USA 103:5602–5607. https://doi.org/10.1073/pnas.0510213103

van Loon LC, Rep M, Pieterse CM (2006) Significance of inducible defense-related proteins in infected plants. Annu Rev Phytopathol 44:135–162. https://doi.org/10.1146/annurev.phyto.44.070505. 143425

Verde I, Jenkins J, Dondini L, Micali S, Pagliarani G, Vendramin E, Paris R, Aramini V, Gazza L, Rossini L, et al. (2017) The Peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. BMC Genomics 18:225. https://doi.org/10.1186/s12864-017-3606-9

Wagner U, Edwards R, Dixon DP, Mauch F (2002) Probing the diversity of the Arabidopsis glutathione *S*-transferase gene family. Plant Mol Biol 49:515–532. https://doi.org/10.1023/A:1015557300450

Wan WL, Kim ST, Castel B, Charoennit N, Chae E (2021) Genetics of autoimmunity in plants: an evolutionary genetics perspective. New Phytol 229:1215–1233. https://doi.org/10.1111/nph.16947

Wang J, Liu W, Zhu D, Hong P, Zhang S, Xiao S, Tan Y, Chen X, Xu L, Zong X, et al. (2020) Chromosome-scale genome assembly of sweet cherry (*Prunus avium* L.) cv. Tieton obtained using long-read and Hi-C sequencing. Hortic Res 7:122. https://doi.org/10.1038/s41438-020-00343-8

Watanabe K, Yoshikawa K (1967) Notes on variation and self-incompatibility in Japanese flowering cherries. Bot Mag Tokyo 80:257–260. https://doi.org/10.15281/jplantres1887.80.257

Xu X, Chen C, Fan B, Chen Z (2006) Physical and functional interactions between pathogen-induced

Arabidopsis WRKY18, WRKY40, and WRKY60 transcription factors. Plant Cell 18:1310–1326. https://doi.org/10.1105/tpc.105.037523

Yadeta KA, Elmore JM, Creer AY, Feng B, Franco JY, Rufian JS, He P, Phinney B, Coaker G (2017) A cysteine-rich protein kinase associates with a membrane immune complex and the cysteine residues are required for cell death. Plant Physiol 173:771–787. https://doi.org/10.1104/pp.16.01404

Yamamoto E, Takashi T, Morinaka Y, Lin S, Wu J, Matsumoto T, Kitano H, Matsuoka M, Ashikari M (2010) Gain of deleterious function causes an autoimmune response and Bateson-Dobzhansky-Muller incompatibility in rice. Mol Genet Genomics 283:305–315. https://doi.org/10.1007/s00438-010-0514-y

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 13:134. https://doi.org/10.1186/1471-2105-13-134

Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 11:R14. https://doi.org/10.1186/gb-2010-11-2-r14

Ziadi S, Poupard P, Brisset MN, Paulin JP, Simoneau P (2001) Characterization in apple leaves of two subclasses of PR-10 transcripts inducible by acibenzolar-S-methyl, a functional analogue of salicylic acid. Physiol Mol Plant Pathol 59:33–43. https://doi.org/10.1006/pmpp.2001.0343

Zubini P, Zambelli B, Musiani F, Ciurli S, Bertolini P, Baraldi E (2009) The RNA hydrolysis and the cytokinin binding activities of PR-10 proteins are differentially performed by two isoforms of the Pru p 1 peach major allergen and are possibly functionally related. Plant Physiol 150:1235–1247. http://doi.org/10.1104/pp.109.139543

Zuo C, Liu H, Lv Q, Chen Z, Tian Y, Mao J, Chu M, Ma Z, An Z, Chen B (2020) Genome-wide analysis of the apple (*Malus domestica*) cysteine-rich receptor-like kinase (CRK) family: Annotation, genomic organization, and expression profiles in response to fungal infection. Plant Mol Biol Rep 38:14–24. https://doi.org/10.1007/s11105-019-01179-w

Table and Figure Legends

Table 1 Gene expression number in each referenced genome for the normal (SN), necrotic weak growth (SW), and wild cherry (CJ) seedlings.

Table 2 Annotations of defense-related DEGs listed in the top 50 up-regulated DEGs in the necrotic hybrid

seedlings identified by TCC analysis, referencing the peach genome and their expression patterns in each part of the seedlings.

DEGs were represented with log fold change (M value) and FDR values (Q value) on the comparison of normal seedling (SN) vs. SW, SN vs. true leaves (Leaf), SN vs. cotyledons (Coty), and SN vs. hypocotyl (Hypo). Non-significant values (FDR > 0.05) were grayed out.

Fig. 1 (A) The normal-growth (SN) and (B) necrotic weak-growth phenotypes (SW) of the hybrid seedlings approximately two weeks after germination. (C) Sampling scheme of each part of the SW seedlings. Leaf: true leaves, Coty: cotyledons, Hypo: hypocotyl, Root: root.

Fig. 2 Significantly enriched GO terms detected by goseq in the up- (A) and down-regulated DEGs (B) of the peach (Ppe), sweet cherry (Pav), and 'Somei-yoshino' (Pye) genome referencing analysis. Significant enrichment of biological process (BP), cell component (CC), and molecular function (MF) GO terms was represented with a q value (FDR) and the number of DEGs (dot size). The GO terms "defense response" and "response to biotic stimulus" are indicated by red arrows.

Fig. 3 An example of gene expression pattern on each seedling part. The normalized gene expression levels for eight up-regulated and four down-regulated SW-specific DEGs were calculated by TMM method using TCC-GUI pipeline. SN: normal-growth seedlings, Coty: cotyledons, Leaf: true leaves, Hypo: hypocotyl, Root: root.

Fig. 4 A heatmap for the enrichment pattern of biological process GO terms in the comparison of normal seedlings (SN) vs. cotyledons (Coty), SN vs. true leaves (Leaf), and SN vs. hypocotyl (Hypo). Significant enrichment terms were represented with the Z score for up-regulation (red) and down-regulation (blue) detected by PAGE analysis. Several GO terms related to plant defense response (red), photosynthesis (blue), and cell cycles (orange) are indicated by arrows.

Supplementary Table S1 Primer sequences for the quantitative RT-PCR.

Supplementary Table S2~~1~~ Genes used for the haplotype allele sequence detection and their paralogs in the phased SPA and SPE genomes of 'Somei-yoshino'.

Supplementary Table S3~~2~~ Summary of the sequencing data and mapping rates to the reference genomes of each sample.

Supplementary Table S4~~3~~ List of top 50 up-regulated DEGs and their annotations in the necrotic hybrid

seedlings identified by TCC analysis referencing the *Prunus persica* genome v2.0.a1.

Supplementary Table S54 Significant enrichment of gene family in the up- and down- regulated DEGs detected by GenFam.

Supplementary Table S65 Identities of the phased allele and its paralogous sequences on the SPA and SPE genomes of 'Somei-yoshino'.

Supplementary Fig. S1 Significantly up- and down-regulated differentially expressed genes (DEGs) and detection of weak growth seedling (SW)-specific expression genes in the peach (A–D), sweet cherry (E–H), and 'Somei-yoshino' genome referencing analyses (I–L). TCC-GUI pipeline was used for the identification of DEGs in the comparison of wild cherry seedlings (CJ) vs. normal growth seedlings (SN) (A, E, and I), CJ vs. SW (B, F, and J), and SN vs. SW (C, G, and K). Of these DEGs, those that were common to CJ vs. SW and SN vs. SW but not present in CJ vs. SN (squared) were determined as SW-specific DEGs (D, H, and L). Several characteristic DEGs, such as PR1, PR5, LRR-RLK, CRK, WRKY, and LAC17, listed among the top significant DEGs (Table 2) are indicated by arrows.

Supplementary Fig. S2 Relative expression levels of SW to SN seedlings in three genes evaluated by quantitative RT-PCR (A) and its correlation to the M-values obtained from RNA-seq analysis (B). The relative expression levels are represented in mean value with confidence intervals (error bars).

Supplementary Fig. S32 Significantly enriched GO terms in the up- (A) and down-regulated (B)DEG list detected by SEA analysis using AgriGO v2. The GO terms were represented with q value (FDR) and the number of DEG (dot size). The terms "defense response" and "response to biotic stimulus" are indicated by red arrows.

Supplementary Fig. S43 Significantly enriched GO terms detected by AgriGO v2 PAGE analysis. The significant enrichment for up- (Z score > 0) and down-regulation (Z score < 0) was represented with FDR and the gene number (dot size). The terms "defense response" and "response to biotic stimulus" are indicated by red arrows.

Supplementary Fig. S54 Venn diagrams of significant DEGs detected in the comparison of the normal seedling (SN) and each seedling part, cotyledons (Coty), true leaves (Leaf), and hypocotyl (Hypo).

Supplementary Fig. S65 Heatmaps of the enrichment GO terms for cell component (A) and molecular function (B) in the comparison of the normal seedlings (SN) vs. cotyledons (Coty), SN vs. true leaves (Leaf), and SN vs. hypocotyl (Hypo). The significant enrichment terms were represented with Z scores

for the up- (red) and down-regulation (blue) detected by PAGE analysis.

Tables Click here to access/download;Table;MT_Tables.docx

Table 1 Gene expression number in each referenced genome for the normal (SN), necrotic weak growth (SW), and wild cherry (CJ) seedlings.

Reference genome	Total number of transcribable gene	Number of gene expressed in				
		any sample	SN	SW	CJ	all samples
Peach	47,089	22,218	20,790	20,771	20,594	19,148 (86.2%)
Sweet cherry 'Tieton'	40,338	31,181	28,828	28,690	28,483	26,195 (84.0%)
'Somei-yoshino'	95,076	70,746	63,483	63,896	62,580	55,749 (78.8%)
CYE_SPA	48,280	36,392	33,097	33,073	31,470	28,484 (78.3%)
CYE_SPE	46,796	34,354	30,386	30,823	31,110	27,265 (79.4%)

Table 2 Annotations of defense response-related DEGs listed in the top 50 up-regulated DEGs in the necrotic hybrid seedlings (SW) identified by TCC analysis, referencing the peach genome and their expression patterns in each part of the seedlings.

Rank	Gene name	BLAST search result	Symbol	SN vs. SW		SN vs. Coty		SN vs. Leaf		SN vs. Hypo	
				M value	Q value	M value	Q value	M value	Q value	M value	Q value
1	Prupe.8G152900	AT2G14580 | basic pathogenesis-related protein 1	PR1	8.87	$6.40e^{-42}$	3.53	$7.94e^{-04}$	7.35	$1.50e^{-14}$	10.1	$1.09e^{-40}$
4	Prupe.8G011000	AT2G36690 | 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily protein (2OG-FeII)	2OG-FeII	8.04	$3.24e^{-21}$	3.93	$1.73e^{-05}$	5.06	$4.09e^{-06}$	9.00	$6.33e^{-10}$
10	Prupe.5G094000	AT4G11650 | osmotin 34; thaumatin-like protein	TLP	6.05	$1.02e^{-13}$	2.36	0.0815	5.17	0.0202	8.43	$1.81e^{-27}$
15	Prupe.1G054800	AT2G29420 | glutathione S-transferase tau 7	GST	3.69	$2.10e^{-12}$	1.50	0.0690	2.47	0.0040	4.72	$4.53e^{-14}$
18	Prupe.4G263100	AT1G52800 | 2OG-FeII	2OG-FeII	3.88	$8.50e^{-12}$	2.80	0.0278	1.44	0.207	5.15	$1.75e^{-06}$
20	Prupe.8G210600	AT2G29420 | glutathione S-transferase tau 7	GST	6.11	$1.02e^{-11}$	3.81	0.0010	3.87	0.0043	7.74	$6.86e^{-09}$
23	Prupe.4G027700	AT4G05200 | cysteine-rich RLK 25	CRK	3.75	$3.29e^{-11}$	0.28	1	2.68	0.0013	5.31	$2.39e^{-13}$
26	Prupe.1G362600	AT1G77280 | Protein kinase protein with adenine nucleotide alpha hydrolases-like domain	LRR-RLK	5.40	$3.72e^{-11}$	6.46	$6.80e^{-05}$	3.90	0.0259	7.31	$5.05e^{-07}$
29	Prupe.1G054900	AT3G09270 | glutathione S-transferase TAU 8	GST	3.71	$9.56e^{-11}$	1.57	0.0795	2.51	0.0080	4.74	$5.47e^{-13}$
31	Prupe.4G262800	AT1G52800 | 2OG-FeII	2OG-FeII	5.92	$1.21e^{-10}$	4.19	$3.31e^{-04}$	6.16	$5.12e^{-06}$	7.15	$7.67e^{-18}$
32	Prupe.2G095000	AT2G36690 | 2OG-FeII	2OG-FeII	5.97	$1.44e^{-10}$	4.42	0.0693	2.74	0.281	6.67	$2.07e^{-05}$
33	Prupe.3G110300	AT3G22060 | RLK-related family protein	CRK	3.59	$1.56e^{-10}$	1.50	0.308	3.06	$4.45e^{-04}$	5.47	$1.19e^{-16}$
35	Prupe.1G129000	PRU1_PRUAR Major allergen Pru ar 1	Bet v1	5.75	$2.06e^{-10}$	3.73	0.0011	2.86	0.0773	7.47	$9.55e^{-06}$
37	Prupe.1G387500	AT4G31940 | cytochrome P450, 82C4	Cyt_P450	2.89	$2.25e^{-10}$	2.13	0.0354	2.14	0.179	2.92	$8.11e^{-07}$
40	Prupe.6G094400	PRU1_PRUAR Major allergen Pru ar 1	Bet v1	3.21	$3.84e^{-10}$	2.51	$7.15e^{-06}$	1.42	0.0559	4.88	$7.85e^{-07}$
41	Prupe.7G102900	AT2G30830 | 2OG-FeII	2OG-FeII	3.97	$4.28e^{-10}$	0.63	1	1.43	0.438	5.44	$1.92e^{-08}$
42	Prupe.1G393100	AT1G80840 | WRKY DNA-binding protein 40	WRKY40	6.36	$5.07e^{-10}$	3.01	0.124	3.10	0.0531	7.83	$4.53e^{-09}$
43	Prupe.8G210500	AT2G29420 | glutathione S-transferase tau 7	GST	5.92	$5.35e^{-10}$	1.83	0.265	2.82	0.0452	7.15	$9.27e^{-08}$
45	Prupe.1G127300	AT1G24020 | MLP-like protein 423; Bet v1-like	Bet v1	3.42	$7.30e^{-10}$	1.18	0.154	0.92	0.659	4.78	$3.32e^{-05}$
47	Prupe.7G103000	AT2G30830 | 2OG-FeII	2OG-FeII	4.12	$1.27e^{-09}$	0.85	1	1.57	0.452	5.67	$4.15e^{-08}$
50	Prupe.1G515800	AT1G35710 | Protein kinase family protein with leucine-rich repeat domain	LRR-RLK	4.37	$2.91e^{-09}$	0.31	1	4.47	$2.27e^{-04}$	5.55	$1.58e^{-12}$

DEGs were represented with log fold change (M value) and FDR values (Q value) on the comparison of normal seedling (SN) vs. SW, SN vs. true leaves (Leaf), SN vs. cotyledons (Coty), and SN vs. hypocotyl (Hypo). Non-significant values (FDR > 0.05) were grayed out.

Click here to access/download;Figure;MT_Figures.docx

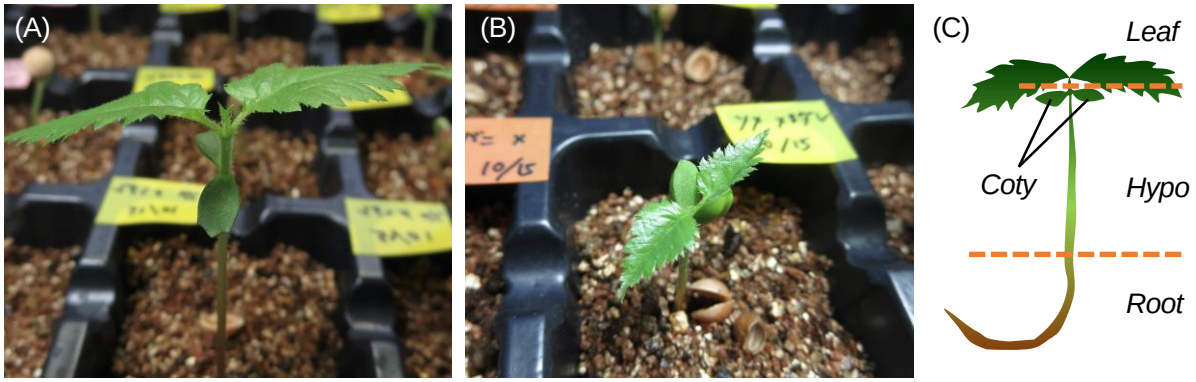

Fig. 1 (A) The normal-growth (SN) and (B) necrotic weak-growth phenotypes (SW) of the hybrid seedlings approximately two weeks after germination. (C) Sampling scheme of each part of the SW seedlings. Leaf: true leaves, Coty: cotyledons, Hypo: hypocotyl, Root: root.

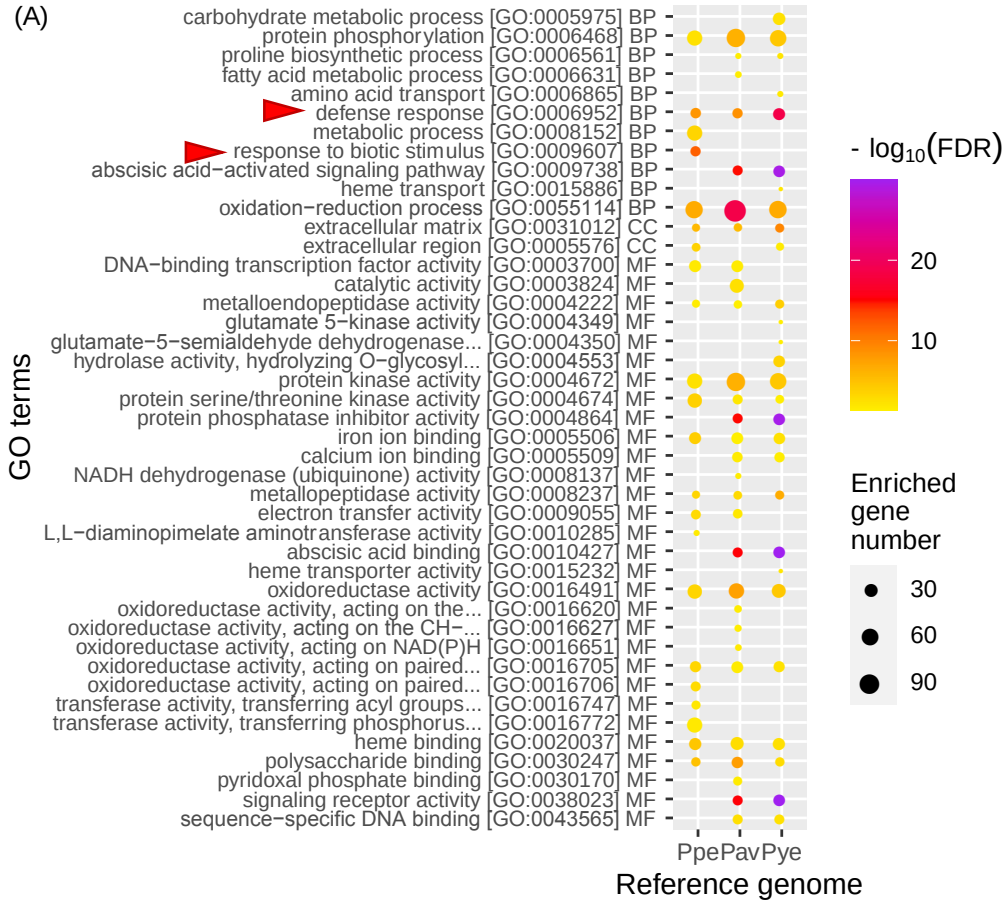

Fig. 2 Significantly enriched GO terms detected by goseq in the up- (A) and down-regulated DEGs (B) of the peach (Ppe), sweet cherry (Pav), and 'Somei-yoshino' (Pye) genome referencing analysis. Significant enrichment of biological process (BP), cell component (CC), and molecular function (MF) GO terms was represented with a q value (FDR) and the number of DEGs (dot size). The GO terms "defense response" and "response to biotic stimulus" are indicated by red arrows.

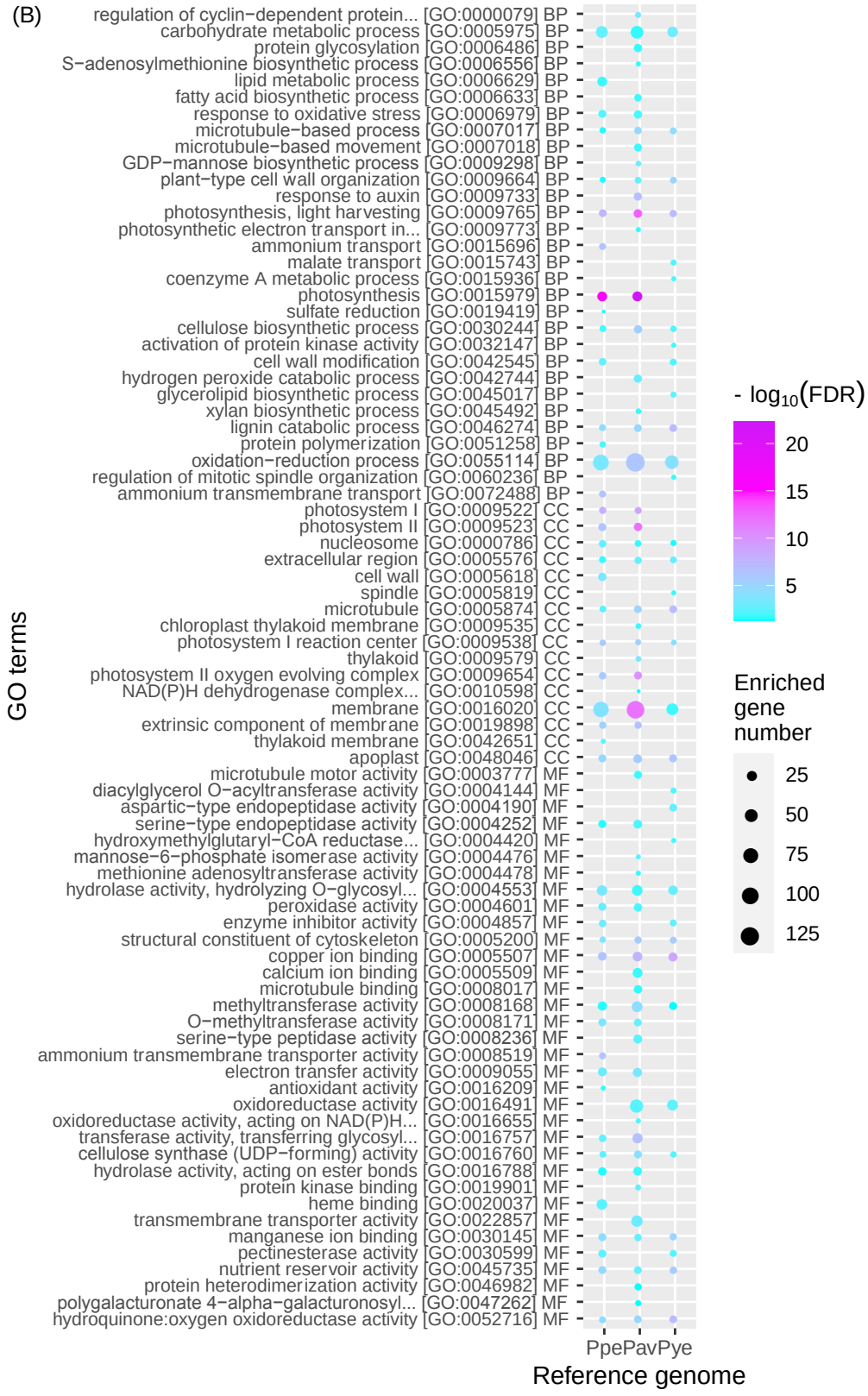

Fig. 2 (Continued)

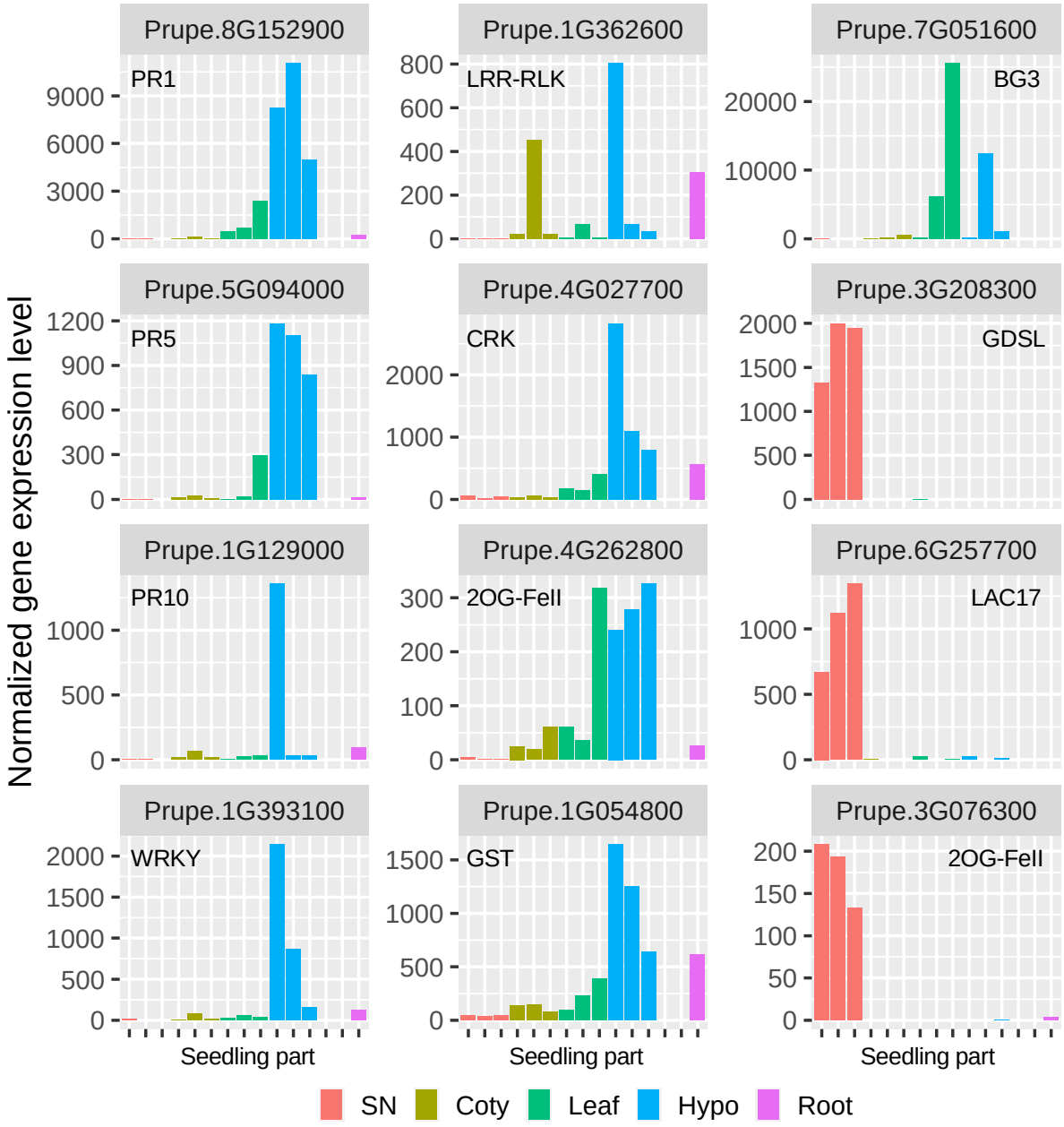

Fig. 3 An example of gene expression pattern on each seedling part. The normalized gene expression levels for eight up-regulated and four down-regulated SW-specific DEGs were calculated by TMM method using TCC-GUI pipeline. SN: normal-growth seedlings, Coty: cotyledons, Leaf: true leaves, Hypo: hypocotyl, Root: root.

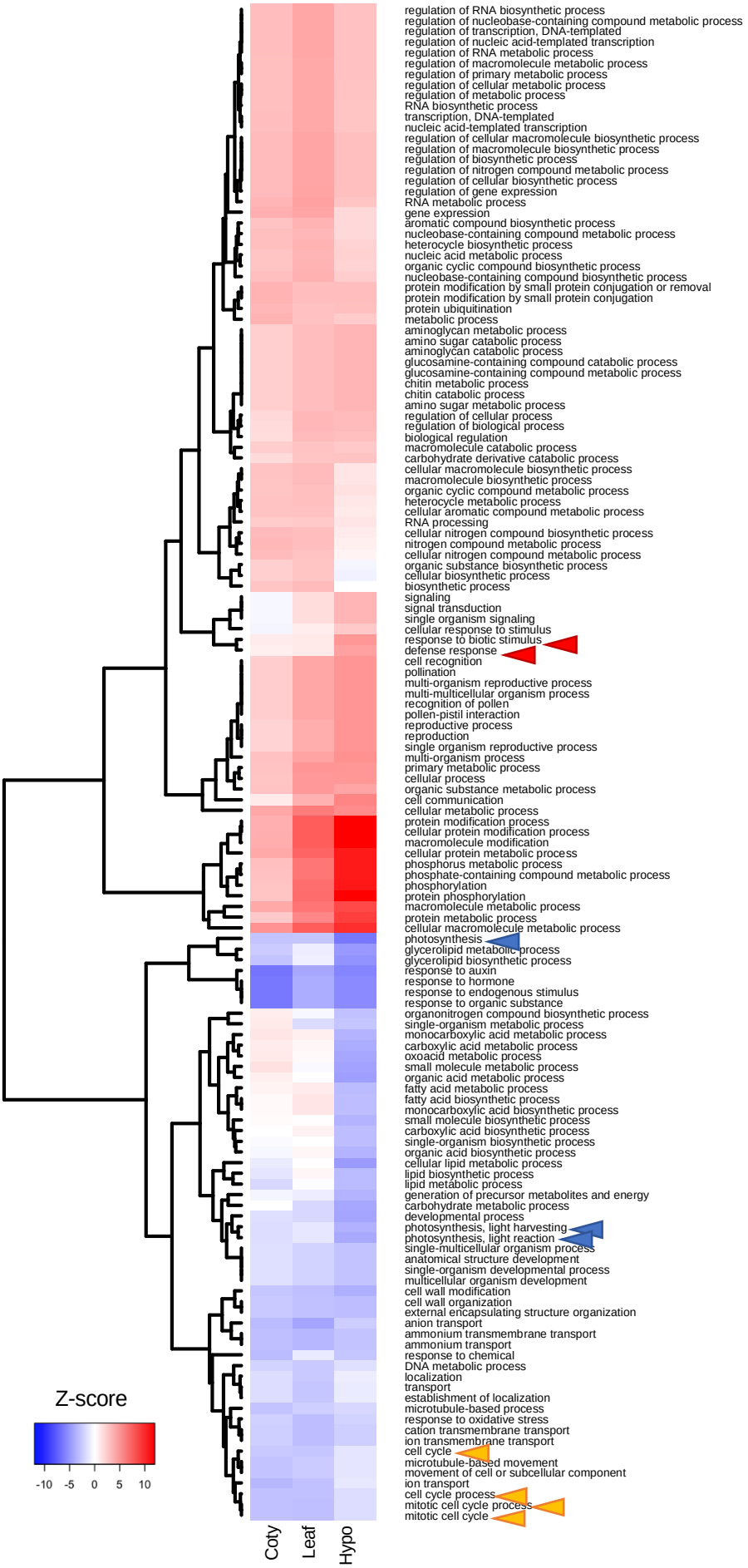

Fig. 4 A heatmap for the enrichment pattern of biological process GO terms in the comparison of normal seedlings (SN) vs. cotyledons (Coty), SN vs. true leaves (Leaf), and SN vs. hypocotyl (Hypo). Significant enrichment terms are represented with the Z score for up-regulation (red) and down-regulation (blue) detected by PAGE analysis. Several GO terms related to plant defense response (red), photosynthesis (blue), and cell cycles (orange) are indicated by arrows.

Supplementary Material 1_Figures

Click here to access/download
Supplementary Material
ESM1_Supplementary Figures_rev1.pdf

Click here to access/download
Supplementary Material
ESM2_Supplementary Table_S1-S3.rev1.pdf

Click here to access/download
Supplementary Material
ESM3_Supplementary Table S4_rev1.xlsx

Click here to access/download
Supplementary Material
ESM4_Supplementary Table_S5-S6.rev1.pdf