Molecular mechanisms for plastic sex expression in hexaploid persimmon

2022

Kanae Masuda

Contents

General Introduction

In contrast to animal, most flowering plants are hermaphrodite, which is thought to be the ancestral state of the sexual system in angiosperm (Ainsworth 2000). A minority of species have evolved separate sexuality, with separate male and female individuals (dioecy), or with the formation of male and female flowers in one individual (monoecy), in a lineage-specific manner (Renner 2014). Some lineages have independently and repeatedly evolved dioecy through intermediate sexualities. Although dioecy might be often viewed as an irreversible state, called “a dead-end”, in reproductive evolution (Heilbuth 2000), plant species have undergone frequent transitions into and out of dioecy (Goldberg et al. 2017) (Fig. 1), which have potentially acted as a key strategy to insure the balance between genetic diversity and stable reproduction (Käfer et al. 2017). Evolutionary processes to establish separate sexualities, especially dioecy, have been well modeled theoretically in the past 40 years (Charlesworth and Charlesworth 1978; Renner and Ricklefs 1995; Barret 2002). Then, they have recently been experimentally validated for the evolution of sex chromosomes in papaya

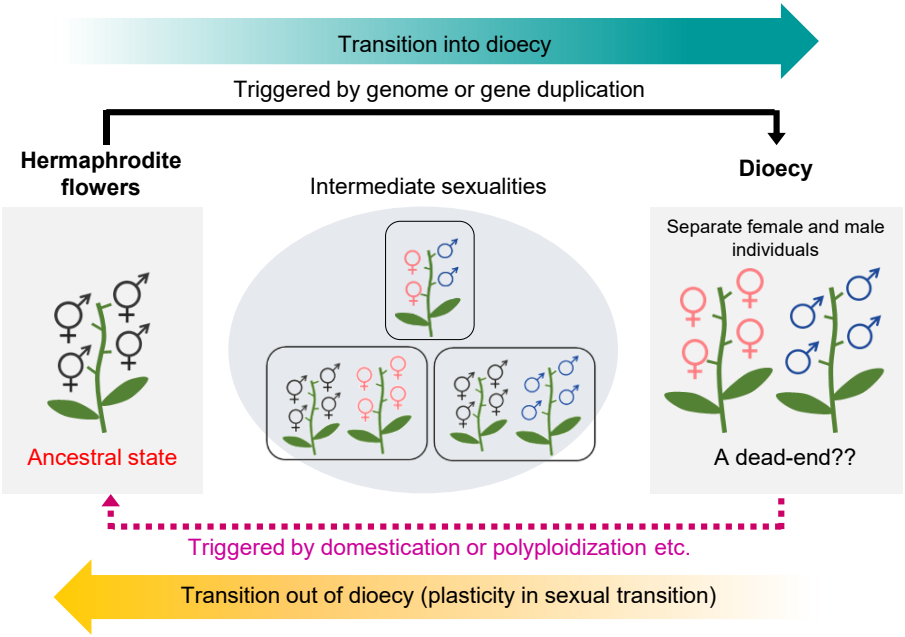

Figure 1 Evolutionary model for transitions into and out of dioecy.

(Liu et al. 2004, Wang et al. 2012), Silene (Kazama et al. 2016, Krasovec et al. 2018), or date palm (Torres et al. 2021), and with the identification of genetic determinants in persimmon (*Diospyros* spp.; Akagi et al. 2014, Akagi et al. 2020), garden asparagus (*Asparagus officinalis* L.; Harkess et al. 2017, 2020, Murase et al. 2017, Tsugama et al. 2017), kiwifruit (*Actinidia* spp.; Akagi et al. 2018; Akagi et al. 2019), and so on. Notwithstanding, the evolutionary or molecular mechanisms to invent the plant-specific flexible transitions among sexual systems have been little understood.

From the studies identifying the sex determinants described above, it has been unveiled that gene or genome-wide paleo-duplication (or ancient polyploidization) events frequently played important roles in the establishment of the new functions to facilitate each sex determinant (Akagi et al. 2020 for persimmon, Akagi et al. 2018, 2019 for kiwifruit, Harkess et al. 2017 for asparagus). On the other hand, recent genome duplications, or polyploidization events have also enabled plasticity in plant sexual reproduction systems, including sexuality (Comai 2005, Ashman et al. 2013, Akagi et al. 2016a). Thus, genome-duplication or polyploidization can potentially act as keys for both transitions into and out of dioecy, as theoretically discussed (Ashman et al., 2013 and Goldberg et al. 2017). Plant specifically have undergone frequent polyploidization cycles, in a lineage-specific manner (Van de Peer et al. 2017), with various phenotypic consequences associated with gene expression variation, either from altered copy number or regulatory interactions, as well as genetic and epigenetic modifications (Wendel 2000, Osborn et al. 2003). While, the actual mechanisms that polyploidizations can alter the sexual systems have been limitedly uncovered so far. Particularly, in contrast to the transitions into dioecy, the molecular mechanisms to trigger the transitions out of dioecy by polyploidization are still poorly understood.

Polyploidization would be fundamentally associated with changes in the mating system,

especially if the origin is a dioecious species. A newly polyploidized individual, which would undergo the reproductive isolation from the original dioecious population, should adjust their mating system to selfing immediately, to properly make the next generations. Some reports have provided actual evidence for the transition out of (or breakdown of) dioecy in polyploid plants. For example, artificially induced autopolyploidy in dioecious *Silene* and *Rumex* gave rise to potentially selfing hermaphrodite progeny (reviewed in Westergaard, 1958). In natural populations, *Empetrum* (Miller and Venable, 2000), *Bryonia* (Volz and Renner 2008), *Mercurialis annua* (Pannell et al., 2004), and *Diospyros* including oriental persimmon (Akagi et al. 2016a), exhibited breakdown of dioecy into various selfing systems. On the other hand, polyploidization has not necessarily led to breakdown of dioecy in natural population, such as in the genus *Actinidia*, where independently polyploidized species, including *A. deliciosa*, *A. macrosperma*, and *A. arguta*, all exhibited a common dioecious sex determination system, as well as the diploid species (Chat et al. 2004, Akagi et al. 2018). Evolutionary contexts to break or maintain dioecious systems (or simply separated sexual systems) in polyploidization events have been long discussed (Miller and Venable, 2000, Glick et al. 2016) while not have been fully understood yet, presumably because of the lack of insights into their concrete mechanisms to alter the sexual systems via polyploidization.

The genus *Diospyros* would exhibit an interesting transition of sexual systems, providing good materials for association in polyploidy and sexuality. Most *Diospyros* species are dioecious, while limited numbers of species often exhibit more plastic sexuality, with polygamous (i.e. male, female, and hermaphrodite flowers in a plant) or monoecious (i.e. male and female flowers in a plant) systems (Duangjai et al. 2006). These breakdowns of dioecy tend to be found in some polyploidy species, and are supposed to be originated from independent polyploidization events. For example, diploids *D. lotus*, *D. glaucifolia*, *D. dygina*,

etc. are thought to be perfect dioecious, while independently derived polyploid *D. kaki* (basically 6X), *D. virginiana* (4X or 6X), *D. rhombifolia* (6X), and *D. ebenum* (6X), often show the described plastic sexualities (Duangjai et al. 2006, Yonemori et al. 2008), with minor exceptions, such as a diploid *D. oleifera* which is reported to occasionally produce male, female, and hermaphrodite flowers in a tree (Suo et al. 2020). Recently, the whole genome of *D. lotus*, which is a diploid wild relative of hexaploid oriental persimmon (*D. kaki*), was sequenced. Evolutionary interpretation on this genome suggested that the genus *Diospyros* underwent at least one paleo-polyploidization event in lineage-specific manner, which would trigger the transition into dioecy (Akagi et al. 2020). The basic mechanism underlying dioecious sex determination in the genus *Diospyros* depends on the Y-encoded small-RNA gene *OGI* and its targeted autosomal feminizing gene *MeGI* (Akagi et al. 2014). Evolutionary analyses on the *D. lotus* genome suggested that a lineage-specific genome-wide duplication (or paleo-polyploidization) derived the first twin, *MeGI* and *Sister of MeGI* (*SiMeGI*), and then *MeGI* specifically underwent an adaptive evolution to positively acquire the feminization function. Later, a segmental duplication event spawned the *MeGI*'s regulator, *OGI*, on the Y-chromosome, to be dioecy (Akagi et al. 2020). These results suggested that an original non-sex determining gene could be rapidly evolved to acquire the new functions involving sex determination via ancient polyploidization and segmental duplications, which might explain the importance of paleo-duplications in the transition into dioecy. The association of paleo-duplications and transition into dioecy have been reported also in other lineages, such as in the genus *Asparagus* (Harkess et al. 2017), *Actinidia* (kiwifruit) (Akagi et al. 2018, 2019), or *Populus* (Müller et al. 2020). On the other hand, the molecular mechanisms involving the transition out of the dioecy in polyploidized *Diospyros* species have been limitedly uncovered.

Here, we focus on mainly three aspects of transitions out (or breakdown) of dioecy in

hexaploid *D. kaki* (Fig. 2). Firstly, the genetically male individuals carrying the Y-encoded *OGI* in *D. kaki* exhibit monoecious sex determination potentially producing both male and female flowers in one tree. Akagi et al. (2016a) suggested that a sex determination gene *MeGI* has established a novel DNA methylation status in monoecious hexaploid *D. kaki*, adding a layer of epigenetic regulation to the genetic sex determination system. The basic regulatory network and the function of *OGI/MeGI* for sex determination are conserved also in hexaploid *D. kaki*, while their expression patterns are quite distinct from the diploid relatives, such as *D. lotus*. In contrast to the stable annual expression pattern of *OGI* in *D. lotus* (Akagi et al. 2014), the *OGI* expression in *D. kaki* is thought to be semi-silenced putatively due to the highly methylated retrotransposon-like insertion, named *Kali*, within the *OGI* promoter region (Akagi et al. 2016a). Although the timing of *OGI* activation in *D. kaki* remains to be solved, once small-RNAs of *OGI* can access the *MeGI* transcript, it is thought to start transitive RNAi to repress the *MeGI* expression and establish DNA methylations in the *MeGI* promoter. Then, the maintenance and release of the DNA methylations in the *MeGI* promoter act as a switch for flower sex conversion in monoecious cultivars. However, the mechanisms to determine the male and female balances in this monoecious system, or namely, to regulate the *OGI* reactivation and the epigenetic switches of *MeGI* is little understood. Secondly, a few genetically female cultivars in *D. kaki*, which do not carry the *OGI* gene in the genome, occasionally exhibit flower sex conversion into male (Yakushiji et al. 2007, Esumi et al. 2015), while this sex conversion is thought to be uninheritable (personal communications, Dr. Akihiko Sato). This situation might suggest an agriculturally beneficial possibility of cross breeding, only with genetically female individuals. Thirdly, in addition to the conversion from dioecy to monoecy, hexaploid *D. kaki* further exhibits conversion from male to hermaphrodite flowers, in response to natural environmental signals (Bawa 1980), or to artificial cytokinin

treatment (Yonemori et al. 1990). Importantly, these transitions out of dioecy into monoecy or hermaphroditism have not generally been observed in diploid *Diospyros* species. In this study, we aimed to identify the molecular pathways associated with these three aspects of polyploid-specific breakdown of dioecy in hexaploid *D. kaki* (Fig. 2).

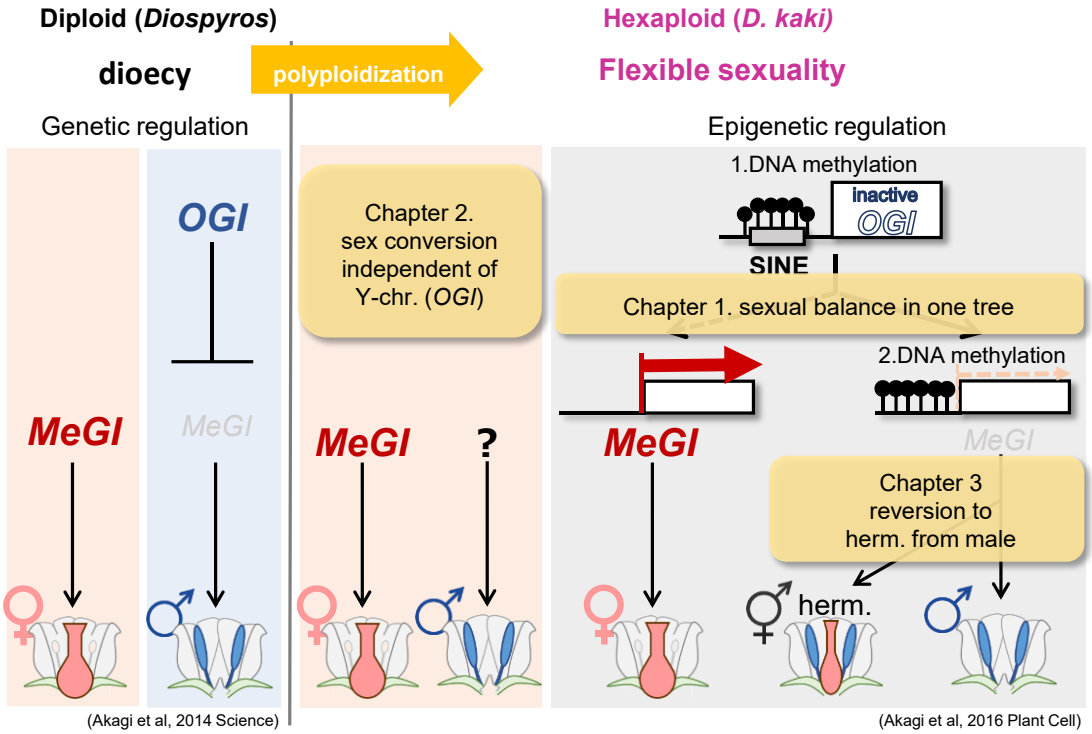

Figure 2 Three aspects of transitions out of dioecy in hexaploid *D. kaki*.

Chapter 1 The conversion from dioecy to monoecy in genetically male tree

Section 1.1 Molecular mechanism underlying unsilenced male production

1.1.1 Introduction

In hexaploid *D. kaki*, as described, *OGI* expression is substantially inactivated because of the presence of a highly methylated SINE-like transposon, named *Kali*, in the promoter region (Akagi et al., 2016a). Hypothetically, once *OGI* occasionally (or accidentally) works to trigger small-RNA production in *MeGI*, the resultant DNA methylation in the promoter region represses *MeGI* expression to form initial male flowers. Then, the flower sexualities from male parental branches are dependent on the maintenance or release of DNA methylation in *MeGI*. In previous studies, all of the assessed genetically male cultivars/accessions in hexaploid *D. kaki* exhibit monoecy or female, which carry at least one copy of *OGI* gene, included the *Kali* insertion in the *OGI* promoter region (Akagi et al., 2016 a,b). This suggested that the (semi-)silenced *OGI* had the evolutionary significance to establish the hexaploid *D. kaki* population. However, among a wide variety of persimmon cultivars (*D. kaki*) deposited to the germplasm repository in the Genebank Project, NARO, Japan, cv. Kumemaru (https://www.gene.affrc.go.jp/databases-plant_search_detail_en.php?jp=117219) is thought to specifically produce only male flowers. This cultivar was derived from a selection of chance seedlingin Toyama prefecture, Japan. Although the genetic factors or inheritance mode involving this unsilenced male production system have not been identified, the previous study has already indicated that cv. Kumemaru carries the *OGI* gene with *Kali*-SINE insertion in its promoter region, as well as in other cultivars (Akagi et al., 2016a). To characterize the mechanism underlying the unsilenced male production in cv. Kumemaru, we assessed *OGI* and *Kali*-SINE regulation, and then detected co-expression network and comprehensive

polymorphisms by comparing cv. Kumemaru and predominantly female-producing cultivars carrying the *OGI* genes.

1.1.2 Materials and Methods

Plant Materials

Developmental stages of diploid (*Diospyros lotus*) and Oriental persimmon (*D. kaki* Thunb.) buds/flowers were defined in previous studies (Akagi et al., 2016a). Flower buds in 14 predominantly female-producing cultivars, which carry at least one copy of the *OGI* gene but have produced no or rare male flowers, and cv. Kumemaru were collected in early June 2017, which correspond to the male/female primordia differentiation (Akagi et al., 2014, 2016a). The 13 dominantly female-producing cultivars (Atago, Beniwase, Chagone, Kunitomi, Nagara, Nanshi, Ogosho, Okugosho, Oniwa, Saisho, Suruga, Yashima, Yoshino) planted in the experimental orchard of Kyoto University (Kyoto, Japan), and cv. Kumemaru and cv. Suruga planted in the grape and persimmon research station, NIFTS, NARO (Hiroshima, Japan), were used. To construct mRNA-Seq libraries, the buds were collected separately from medial and top branch positions with three to four biological replicates abbreviated as #1 and #2, respectively. The collected samples were frozen at −80°C until used for RNA extractions.

Characterization of MeGI and OGI

The cDNAs of developing flower buds of cv. Taishu, cv. Kumemaru, and *D. lotus* cv. Kunsenshi-male were synthesized from total RNA with ReverTra Ace qRT-PCR Master Mix (TOYOBO, Osaka, Japan). The relative expression level of *OGI* and *MeGI* (Akagi et al., 2014, 2016b for the primer sequences) was detected by qRT-PCR using a LightCycler 480 (Roche

Diagnostics, Mannheim, Germany), with four biological replicates. A constitutively expressed *DkActin* gene in hexaploid persimmon (Akagi et al., 2009) was used for standardizing expression levels. qPCR analyses were conducted under the following conditions: 95°C for 5 min followed by 45 cycles of 95°C for 15 s, 60°C for 15 s, and 72°C for 30 s, using THUNDERBIRD SYBR qPCR Mix (TOYOBO).

Genomic DNA was extracted from young flower buds with the CTAB method (Akagi et al., 2014). The full-length and 5′ promoter regions of the *OGI* in cv. Kumemaru and a wide variety of *Diospyros* species, *D. kaki*, *D. lotus*, *D. glaucifolia*, *D. digyna*, *D. virginiana*, *D. mespiliformis*, and *D. rhombifolia*, were amplified from gDNAs, with the OGI-prom-2F-TOPO and OGI-spR primer set (Akagi et al., 2016a). Quantitative genotyping of *OGI* in cv. Kumemaru was conducted according to the previous report (Akagi et al., 2016b). Alleles of *Kali*-SINE insertion on the *OGI* promoter and *OGI* transcript region, in cv. Kumemaru and *D. kaki* cultivars, were amplified from gDNAs, with primer sets OGI-441-kali-F (5′ -GCTCCAGTTTAGTTATGGAAGAG-3′ , Tm: 60) and OGI-441-kali-R (5′ -ATTGCCAACGTTCACATCAC -3′ , Tm: 63) for the *Kali*-SINE, and OGI-candF and OGI-spR for the *OGI* transcript region (Akagi et al., 2016b). For the detection of DNA methylation status in *OGI*, gDNA extracted from developing flower buds was treated with an EZ DNA Methylation-Gold kit (Zymo Research, United States) to deaminate and convert non-methylated cytosine residues into uracil residues. Bisulfite-treated *OGI* 5′ promoter sequences were amplified with Epi Taq HS (TaKaRa) and each four sense- and antisense-direction primer sets (Akagi et al., 2016a). The bisulfite amplicons were subjected to Illumina library construction described below.

Illumina Sequencing

To construct mRNA-Seq libraries, total RNA was purified with the Dynabeads mRNA Purification Kit (Invitrogen, United States). The mRNA-Seq libraries were prepared with the KAPA RNA HyperPrep Kit (Roche, Switzerland) as previously described (Yang et al., 2019), followed by a DNA cleanup step with AMPure XP beads (Beckman Coulter, United States; AMPure: reaction = 0.8:1). To construct bisulfite-PCR-Seq libraries, the promoter region of *OGI* was amplified from bisulfite-treated gDNA of cv. Kumemaru, according to the previous report (Akagi et al., 2016a). The bisulfite amplicons were purified with AMPure and then applied to KAPA HyperPrep Kit (Roche) as previously described (Akagi et al., 2016a). The purified libraries were quantified with a Qubit 2.0 fluorometer (Invitrogen) and then analyzed with the Illumina HiSeq 4000 system at the QB3 Genomic Sequencing Laboratory of UC Berkeley. The resulting SR50 sequencing reads were analyzed at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley. Raw sequencing reads were processed using custom Python scripts developed in the Comai Laboratory, which are available online, as previously described (Akagi et al., 2014). All Illumina sequencing data have been deposited in the DDBJ database: Short Read Archives (SRA) database (BioProject ID PRJDB9564, Run ID DRR233540-DRR233569).

Transcriptomic Data Profiling

The mRNA-Seq Illumina reads were aligned to the reference coding sequences (CDSs) of the diploid Caucasian persimmon, *D. lotus* (Akagi et al., 2020), using the default parameters of the Burrows-Wheeler Aligner (BWA) (version 0.7.12) (Li and Durbin, 2009). These settings allow the mapping of variable allele sequences in hexaploid *D. kaki* (Yang et al., 2019). The read counts per CDS were determined from the aligned SAM files using a custom R script to calculate the RPKM for each gene. To examine the gene expression dynamics in 14

predominantly female-producing cultivars and cv. Kumemaru, a PCA was conducted using the genes with RPKM > 1, with prcomp in R.

Differentially expressed genes (DEGs) between predominantly female-producing cultivars and cv. Kumemaru were detected by DESeq analysis with “fit-only” for sharing mode (FDR < 0.01). The DEGs were further filtered based on RPKM and bias values (RPKM > 1.0, bias >2). Furthermore, Kumemaru-specific DEGs were detected by comparing between cv. Kumemaru and each predominantly female-producing cultivar with edgeR (Robinson et al., 2010; McCarthy et al., 2012) using the paired-test option depending on the positions in branches (N = 2 × 14 for biological replicates). The DEGs were filtered according to RPKM and p-values (RPKM > 1.0, P < 0.1 for cv. Nagara and cv. Ogosho, and P < 0.05 for the rest of the 12 cultivars). Putative functions of each gene were annotated with a BLASTX search of the TAIR10 database.

Construction of the Co-expression Network

The genes detected as DEGs in the 14 comparisons of predominantly female-producing cultivars and cv. Kumemaru (N = 20,216) were selected to construct a gene co-expression network with the WGCNA package (Langfelder and Horvath, 2008). The soft-thresholding power for a signed network was set at 5, with a scale-free model fitting index R^2 > 0.6. A relatively large minimum module size (30) and a medium sensitivity (deepSplit = 2) to cluster splitting were applied. In co-expression networking, genes were represented by nodes, and the correlation values (weight) between two genes were calculated by raising the Pearson’s correlation coefficient. The heatmap for expression was designed with ggplot2 packages (Csardi and Nepusz, 2006; Wickham, 2009). The genes in the same module were first visualized with the Cytospace program (Shannon et al., 2003) and only lines with a weight

greater than 0.43.

Cataloging of cv. Kumemaru-Specific Subsequences From mRNA-Seq Reads

For the comprehensive caption of Kumemaru-specific polymorphisms in the expressed transcripts, we conducted “kmer cataloging” in the cDNA reads, according to the previous reports characterizing Y-linked polymorphisms in persimmon (Akagi et al., 2014) and kiwifruit (Akagi et al., 2018). We cataloged 35-bp subsequences triggered with an “A” nucleotide in the mRNA-Seq reads of cv. Kumemaru and of the predominantly female-producing cultivars pools using custom Python scripts. The subsequence catalogs from cv. Kumemaru and from female cultivars were compared to extract the Kumemaru-specific subsequences (“0” counts in the female cultivars pool). The Illumina reads including cv. Kumemaru-specific subsequences (k-mers) (KSK) were assembled with a CLC assembler, as described (Akagi et al., 2014). The derived contigs were filtered with the remapping of the KSK to exclude minor polymorphisms or sequencing noises (nos. of KSK < 10).

1.1.3 Results and Discussion

Morphological characterization of male flowers in cv. Kumemaru

The architectures of flowering branches and inflorescences in cv. Kumemaru were similar to those in other monoecious cultivars, where male flowers set in medial positions of the parental branches, with cyme-like trifurcated inflorescence (Fig. 1-1a,b). In the flower developing/maturing stages (stages 2-4) previously defined (Yang et al., 2019), developmental processes of male flowers/organs in cv. Kumemaru were consistent with other monoecious cultivars (Fig. 1-1c-n, compared with Yang et al., 2019), producing fertile anther/pollens (Fig. 1-1l-n). Flower structure/components in cv. Kumemaru were identical to

male flowers in monoecious cv. Taishu (Fig. 1-1o-p). These observations suggested that the mechanism of male flower production in cv. Kumemaru is consistent with monoecious cultivars, where gene networks orchestrated by *MeGI* play an important role (Yang et al., 2019).

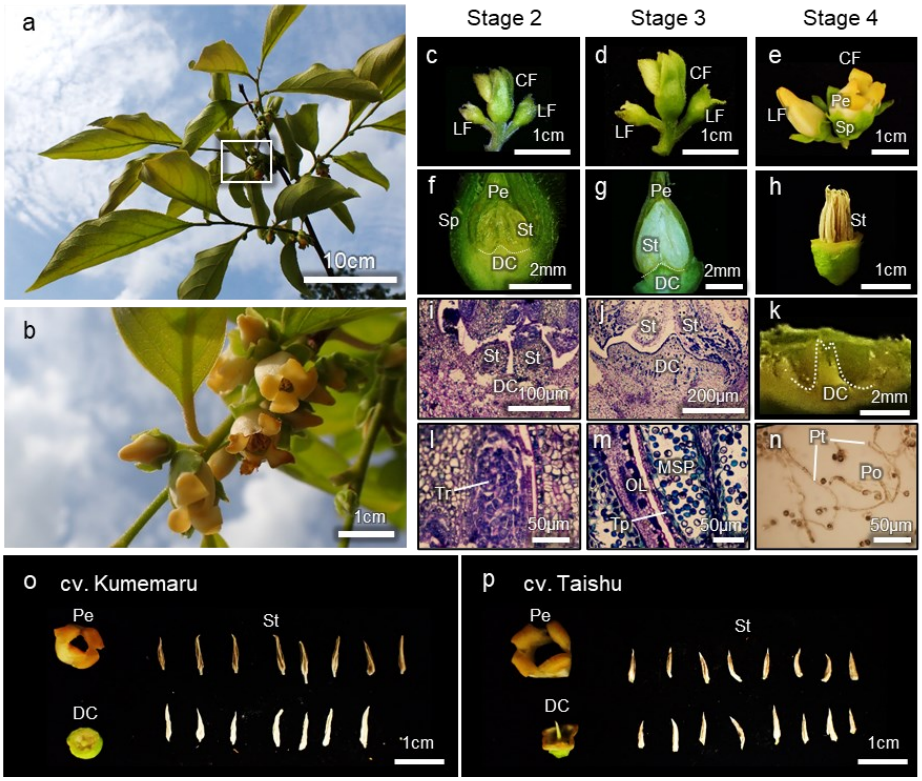

Figure 1-1. Morphological characterization of male flower development in cv. Kumemaru.
a-b, The appearance of flowering branches in cv. Kumemaru (**a**), and the closing up of the white box in panel (**a**), showing cyme-like trifurcated inflorescence (**b**). **c–n**, The developmental processes of male flowers in cv. Kumemaru, in the stages defined according to Yang et al. (2019). In stage 2, tetrads (Tr) were formed inside stamens (St) (**l**). In stage 3, microspores (MSP) were developed, while tapetal layers (Tp) have not been degenerated (**m**). In stage 4, flowers opened with mature stamens, producing fertile pollens (Po) (**h, n**). Sp, sepal; Pe, petal; DC, deficient carpel; Pt, pollen tube. **o-p**, Alignment of male flower components in cv. Kumemaru (**o**) and monoecious cv. Taishu (**p**), which included four gamosepalous petals, 1 deficient carpel, and ca 16 stamens.

Stable male production system caused by derepression of the OGI

To find out this basic molecular mechanism of stable male production, we first analyzed the regulations of *OGI* and *MeGI* in cv. Kumemaru. The expression level of *MeGI* in cv. Kumemaru was substantially lower than in the predominantly female-producing cultivars, at

the onset of primordia development (Fig. 1-2a), which was consistent with male flower buds in other monoecious cultivars (Akagi et al., 2016a). For the structure of *OGI*, cv. Kumemaru has the *Kali*-SINE insertion on the *OGI* promoter like in other cultivars (Fig. 1-2b) and simplex *OGI*. The sequences of *Kali*-SINE insertion in the promoter and transcript region showed no specific polymorphisms in comparison to other cultivars (Akagi et al., 2016b). Furthermore, bisulfite PCR-Seq analysis targeting the *OGI* promoter in flower bud primordia showed that the cytosine residues in the *Kali*-SINE insertion were highly methylated in the CG and CHG sequences (62.8 and 59.4%, respectively) (Fig. 1-2c). The methylation pattern at each cytosine residue was also consistent with male flower buds in the monoecious cultivar (*P* >

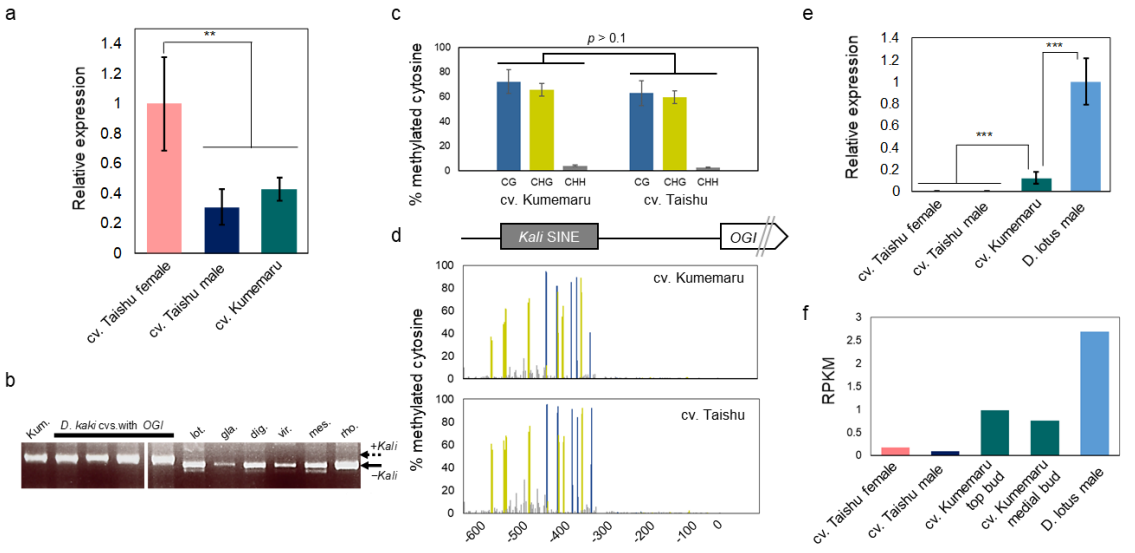

Figure 1-2. Regulation of *OGI* and *MeGI* in cv. Kumemaru.
a, *MeGI* expression level in flower buds in cv. Kumemaru, and monoecious cv. Taishu, from qRT-PCR analysis. cv. Kumemaru showed substantial reduction in the *MeGI* expression, similar to male flowers of cv. Taishu (***P* < 0.01, N = 4 for biological replicates). **b**, PCR analysis detecting the *Kali*-SINE in *D. kaki*, and the relatives. cv. Kumemaru (Kum.) carries only the *Kali* inserted *OGI*, as well as other *D. kaki* cultivars (Akagi et al., 2016a). Other representative *Diospyros* species (lot. *D. lotus*, gla: *D. glaucifolia*, dig: *D. digyna*, vir: *D. virginiana*, mes: *D. mespiliformis*, and rho: *D. rhombifolia*) showed no *Kali*-SINE insertion in the *OGI* promoter region. **c-d**, Cytosine methylation level (**c**) and methylation pattern at each cytosine residue (**d**) in the cv. Kumemaru and cv. Taishu *OGI* promoter region, in flower primordia. cv. Kumemaru showed similar methylation pattern with cv. Taishu (*P* > 0.1). **e-f**, *OGI* expression level in flower buds in cvs. Kumemaru, Taishu, and diploid male *D. lotus* (cv. Kunsenshi-male), from qRT-PCR (**e**) and mRNA-Seq (**f**) analyses. *OGI* is not fundamentally expressed in monoecious cultivars including cv. Taishu (Akagi et al., 2016a), while cv. Kumemaru showed significant increase in the *OGI* expression (E, ****P* < 0.001, N = 4 for biological replicates). The *OGI* expression in cv. Kumemaru was still lower than in diploid *D. lotus*, which carries no *Kali*-SINE in the *OGI* promoter (**e**) (****P* < 0.001, N = 4 for biological replicates). Bars indicate standard errors.

0.1; Akagi et al., 2016a) (Fig. 1-2d). These results signified semi-silenced of the *OGI* expression, as observed in other cultivars (Akagi et al., 2016a). However, in cv. Kumemaru, *OGI* exhibited slight but fundamental expression in the flower bud primordia (RPKM = 0.87 and 1.18 for independent buds from different positions in mRNA-Seq analysis), which was significantly higher than in other monoecious cultivars (RPKM < 0.2 in mRNA-Seq analysis) (Fig. 1-2e-f). The expression level of *OGI* in cv. Kumemaru was still lower than that in diploid male Caucasian persimmon (*D. lotus*) (average RPKM = 2.69), which carries no *Kali*-SINE insertion on the promoter (Fig. 1-2f). These results suggested that the epigenetic sex regulation mechanism, in which the expression of *OGI* is suppressed under the control of the *Kali*-SINE, is shared in common among monoecious hexaploid persimmon cultivars, but *OGI* is induced with an unknown mechanism in cv. Kumemaru.

Differentially Expressed Genes in cv. Kumemaru-Specific Manner

To screen the genes associated with the mechanism for unsilenced of *OGI*, we compared mRNA-Seq data from the flower bud primordia in cv. Kumemaru and the 14 predominantly female-producing cultivars. Profiling with PCA using all the genes with RPKM > 1 (*N* = 32,143) indicated no specific clusters between cv. Kumemaru and the female-producing cultivars, although buds from some female-producing cultivars formed a specific cluster differentiated in the PC1 axis (Fig. 1-3). These results indicated that cv. Kumemaru showed no dynamic expression differentiation from the female-producing cultivars.

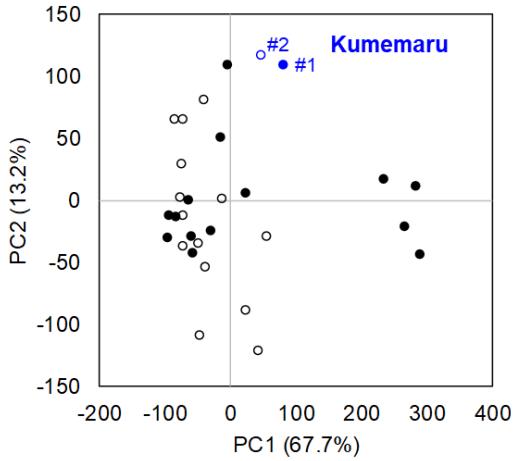

Figure 1-3 Transcriptomic profiles in comparison to cv. Kumemaru (Km) and 14 predominantly female-producing trees.
Characterization of gene expression dynamics by PCA on transcriptomic data (RPKM > 1) of each cultivar. The buds were sampled from two positions, substantially affecting flower sexuality in monoecious cultivars (Akagi et al., 2016a). The two positions of buds in branches such as medial (#1, filled circle) and top (#2, open circle) based on the criteria in the previous report (Akagi et al., 2016a) were used as replicates in each cultivar. No significant difference was observed between the cv. Kumemaru and the other dominantly female-producing cultivars ($p > 0.05$).

To analyze sex-biased gene expression, we conducted paired edgeR analysis between cv. Kumemaru and each female-producing cultivar, separately in 14 combinations ($N = 2 \times 2$ for biological replicates in all 14 combinations), to detect cv. Kumemaru-specific expressions in more stringent criteria ($P < 0.1$ for cvs. Nagara and Ogosho, $P < 0.05$ for the other 12 female cultivars, RPKM > 1). The 21 upregulated genes and 21 downregulated genes were shared in all 14 combinations, defined as Kumemaru-specific DEGs. We detected the candidate genes related to the unsilenced *OGI*, which probably acts as an upstream of *MeGI* during early primordia differentiation.

Core Gene Networks Correlated with the Derepressed OGI Expression System in cv. Kumemaru

To understand the regulatory paths of Kumemaru-specific unsilenced of *OGI*, the co-expression patterns were visualized by applying a weighted correlation network analysis (WGCNA), which were first clustered into 27 modules with total 20,216 DEGs in 14 combinations (Fig. 1-4a). Of them, 8 modules included the Kumemaru-specific DEGs. Here we focused on module 1 ($N = 993$), which has the most significant enrichment of the DEGs (N = 17) ($P = 2.5e\text{-}19$, respectively, for one-sided Fisher's exact test). Of the 612 biological

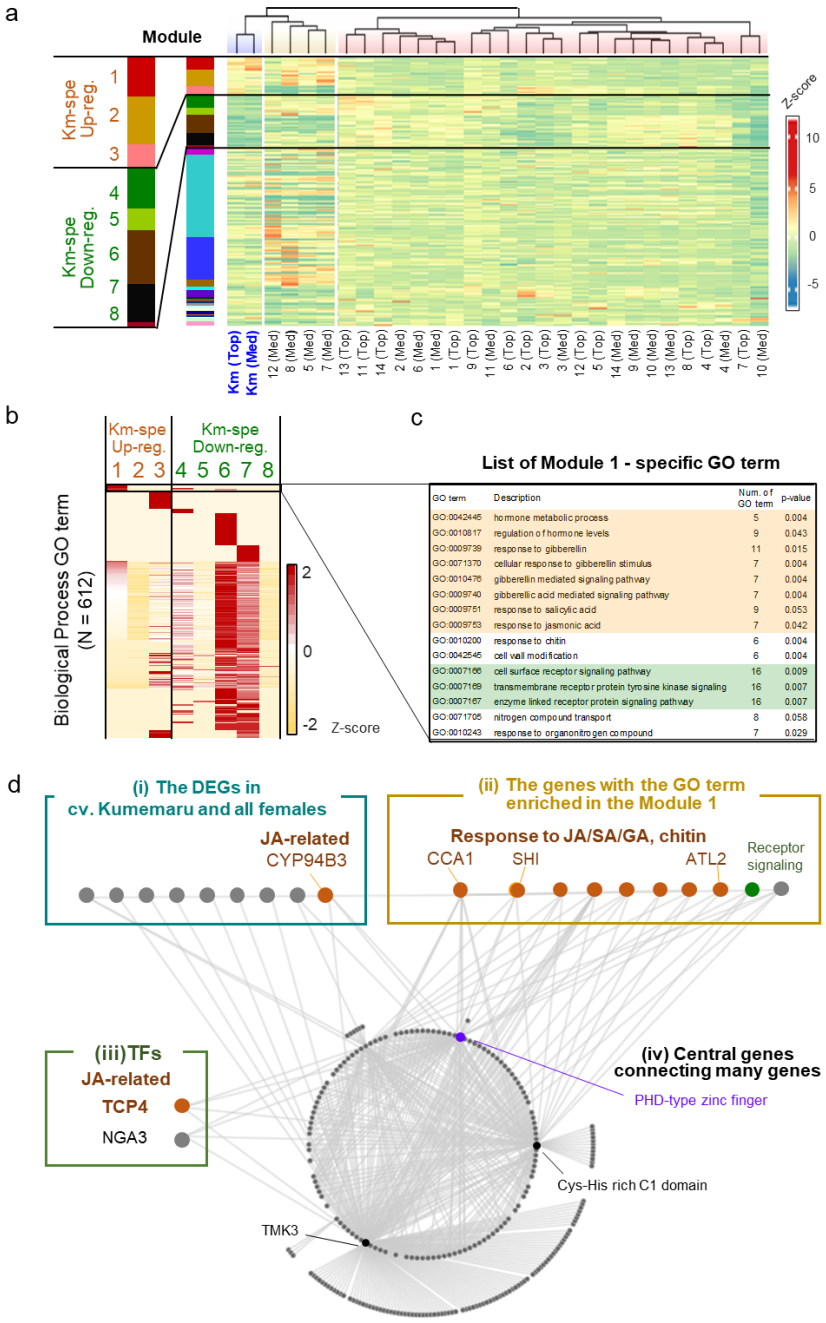

Figure 1-4 Co-expression network analysis in cv. Kumemaru and the predominantly female-producing cultivars.

a, Co-expression moduling of the genes based on expression patterns in cv. Kumemaru and 14 predominantly female-producing cultivars, with the WGCNA package. Modules 1–8 included the comparison II DEGs. **b**, Heatmap for the enrichment of GO terms involving biological processes in these eight modules. The numbers of annotated GO terms were standardized by Z-scoring. **c**, List of the GO terms specifically enriched in module 1 ($P < 0.06$), which included most of the cv. Kumemaru-specific DEGs. The GO terms relating to the hormonal process and receptor signaling pathway were highlighted in orange and green, respectively. **d**, Visualization of the coexpressed gene network in module 1. As described in the main text, they were categorized into the following four groups: (i) the DEGs specific to cv. Kumemaru, (ii) the genes with the GO terms specifically enriched in module 1, as given in the panel (**c**), (iii) transcription factors, and (iv) central genes in this network.

process GO terms annotated to the genes in the 8 modules, 15 terms were specifically enriched in module 1 ($P < 0.06$) (Fig. 1-4b,c). These terms included tyrosine kinase signaling pathway and response to gibberellin (GA), salicylic acid (SA), jasmonic acid (JA), and chitin, which overall implied association with stress signals (Xu et al., 1998; Devoto and Turner, 2004; Hu et al., 2017). These JA/GA/SA-associated genes were not included in the DEGs between male and female buds in diploid dioecious *D. lotus*, where *OGI* is genetically active in male (Akagi et al., 2014). Thus, these JA/SA/GA associated genes might contribute to the unsilenced *OGI* in a cv. Kumemaru-specific manner.

To further dissect the key structure potentially unsilencing *OGI* in cv. Kumemaru, we visualized the co-expression gene network in module 1 and focused on the genes highly correlated to other genes (correlation value $r > 0.43$) (Fig. 1-4d). Here we featured the following four categories: (i) the Kumemaru-specific DEGs, which were thought to be potential candidates to regulate Kumemaru-specific stable maleness (from 42 DEGs as described); (ii) the genes with the GO terms specifically enriched in module 1, which are mostly associated with stress signal and would reflect cv. Kumemaru-specific physiological reactions; (iii) transcription factors; and (iv) central genes connecting many genes in this module. In category (i), 9 out of the 17 Kumemaru-specific DEGs were visualized with high correlation values with the other genes (Table 1-1). Among them, *SIP1* mediates the abiotic stress-induced accumulation of raffinose (Egert et al., 2013), and *CYP94B3* is a key enzyme in the catalytic pathway of JA biosynthesis (Heitz et al., 2011). In category (ii), focusing on JA/SA signaling, *CIRCADIAN CLOCK-ASSOCIATED 1* (*CCA1*) and *ARABIDOPSIS TOXICOS EN LEVADURA 2* (*ATL2*) can act for the promotion of plant defense presumably via JA/SA signals (Salinas-Mondragón et al., 1999; Serrano and Guzmán, 2004; Lei et al., 2019). *ATL2* is often applied as a biomarker for JA/SA signals (Serrano and Guzmán, 2004).

Table 1-1. List of Kumemaru-specific DEGs in module 1.

gene	symbol	putative function	evalue	TAIR
Dlo_pri0856F.1_g00210.1	**CYP94B3**	cytochrome P450, family 94, subfamily B, polypeptide 3	0	AT3G48520.1
Dlo_pri0248F.1_g01000.1	ARO1	armadillo repeat only 1	0	AT4G34940.1
Dlo_pri0056F.1_g03770.1		alpha/beta-Hydrolases superfamily protein	2.0E-95	AT4G10955.1
Dlo_pri0000F.1_g05660.1		Protein kinase superfamily protein	8.5E-97	AT3G17850.1
Dlo_pri0061F.1_g03110.1		Eukaryotic aspartyl protease family protein	0	AT5G45120.1
Dlo_pri0205F.1_g00730.1		glyoxal oxidase-related protein	0	AT1G14430.1
Dlo_pri0236F.1_g00430.1	SIP1	Raffinose synthase family protein	9.3E-165	AT5G40390.1
Dlo_pri0281F.1_g00130.1		UDP-Glycosyltransferase superfamily protein	1.2E-129	AT2G22590.1
Dlo_pri0042F.1_g02140.1		unknown protein	4.9E-12	AT5G66440.1

Focusing on GA signaling, *SHORT INTERNODES* (*SHI*) is thought to involve gibberellin response to regulate internode length and structures (Fridborg et al., 2001). The other seven genes were annotated with the GO term associated with the regulation of hormone levels, receptor signaling, and nitrogen compound transport. For category (iii), we found two transcription factors. *TCP4* can affect JA regulation by directly controlling the expression of *LIPOXYGENASE2* (*LOX2*), which is one of the first steps of the JA biosynthesis pathway in *Arabidopsis* (Schommer et al., 2008). Moreover, the class II TCP family, including *TCP4*, was reported to regulate the expression of the other TF in category (iii), *NGATHA3* (*NGA3*) (Ballester et al., 2015). The results so far are reminiscent of the involvement of stress signals, especially with the effect of JA signals. As the core genes [or category (iv)] predominantly connecting the module 1 network (with > 50 connections), the genes annotated as the PHD-type zinc finger (Lee et al., 2009), Cys-His rich C1 domain (Bhaskar et al., 2015), and *TMK3* (Dai et al., 2013) were identified. Among the three genes, the PHD-zinc finger has been reported to be associated with both gene expression and repression through histone modification (Mellor, 2006), which is possibly related to unsilenced *OGI* in cv. Kumemaru.

Our result of co-expression network so far suggested the involvement of stress, particularly the JA-related pathway, for the activation of *OGI* expression in cv. Kumemaru. In addition to the genes involved in regulations of stress-induced plant hormones, other keywords including chitin response and/or tyrosine kinase signaling (see Fig. 1-4b,c) would imply the

involvement of plant defenses and immunity. Tyrosine kinase has been reported to be responsible for biotic or abiotic stresses (Miyamoto et al., 2019) and is also known as a regulator of gibberellin responses (Fu et al., 2002), abscisic acid (ABA) signaling (Ghelis et al., 2008), cold stress (Sangwan et al., 2001), and sugar responses (Ritsema et al., 2009). Furthermore, one of the core genes in module 1, TMK3, as an Auxin-Binding Protein (ABP) family gene (Xu et al., 2014), is a member of the transmembrane kinase TMK subfamily, which is reported to be inducible under the stress or JA signals in *Nicotiana* plants (Cho and Pai, 2000). These implications might be consistent with empirical observations of flower sex bias in a tree in monoecious persimmon cultivars. Monoecious trees are supposed to produce only female flowers in vigor branches and/or at young age, while older or stressed branches tend to produce a high frequency of male flowers (Hasegawa, 1995; Hasegawa et al., 2004), although they might depend on cultivars/accessions. Furthermore, the ratio of male/female flowers in trees depends on their parental branch characters, such as sexuality, length, and bud positions (Akagi et al., 2016a). These reports suggest that not only genetic factors but also internal and external environment conditions of cv. Kumemaru may be involved in stresses, and may be the causal factor(s) to express unsilenced male production.

Identification of cv. Kumemaru-Specific Transcript Polymorphisms

To identify candidate genetic factors related to the unsilenced *OGI*, kmer cataloging was conducted for the comprehensive caption of Kumemaru-specific polymorphisms. Assembly reads of Kumemaru-specific k-mers (KSK) derived 1,446 contigs (Table 1-2 for the representative 25 genes with > 80 KSK). These polymorphic (or cv. Kumemaru-specific) contigs showed no overlap with the Kumemaru-specific DEGs described above. The transcript contig with the most KSK showed no homologous sequences in the persimmon

genome (Akagi et al., 2020), while it showed undefined numerous virus-like sequences similar to marafiviruses. For other transcript contigs with fundamental numbers of the KSK, we could identify a structural variation of 38-bp insertion in RING-type E3 ubiquitin ligase. The other candidate contigs with the KSK showed no clear disruptions directly related to their protein functions. The detected viral-like sequence was not matched to any other known viruses but most likely to be classified into the genus *Marafivirus*. Although the pathogenicity of this virus is little known, viral proteins encoded by another plant virus are thought to often inhibit JA signaling in *Arabidopsis thaliana* plants (Lozano-Duran et al., 2011) or to interact with plant histone deacetylase (Wang et al., 2018). Thus, it may be also possible that virus infection can induce specific stress signaling to activate *OGI*, although that is not a case of genetic alternation specific to cv. Kumemaru.

Table 1-2. List of the genes including cv. Kumemaru-specific k-mers (KSK) with >80 KSK.

Contigs	KSK counts	Annotation		Highest hit in blastx with TAIR (or with nr database for no hits in TAIR)	
		By persimmon DB	By TAIR		e-value
Contig_19	1428	No hit	No hit	Blackberry virus S (*Marafivirus*)	0
Contig_492	140	No hit	No hit	Persimmon virus A (*Cytorhabdovirus*)	2E-158
Contig_905	133	Dlo_pri 0061 F.1_g 04290.1	AT4G19670.7	RING/U-box superfamily protein	0
Contig_171	128	Dlo_pri0084F.1_g01740.1	AT5G49910.1	Chloroplast heat shock protein 70-2	0
Contig_103	122	Dlo_pri0026F.1_g01940.1	AT5G45650.2	Subtilase family protein	0
Contig_841	119	Dlo_pri0002F. 1_g02880.1	AT1G79570.2	Kinase with octicosapeptide/Phox/Bem1p domain-containing	0
Contig_634	117	Dlo_pri0213F.1_g 00220.1	AT1G10430.2	Protein phosphatase 2A-2	0
Contig_221	111	Dlo_pri0044F1_g 02210.1	AT3G09630.1	Ribosomal protein L4/L1 family	0
Contig_739	109	Dlo_pri0176F1_g01650.1	AT5G62770.1	Membrane-associated kinase regulator%2C putative	3E-27
Contig_495	106	No hit	No hit	Nectarine marafivirus M (*Marafivirus*)	8E-42
Contig_275	104	Dlo_pri0004F. 1_g05490.1	AT3G16910.1	Acyl-activating enzyme 7	0
Contig_131	98	Dlo_pri0002F1_g 00210.1	No hit	No hit	
Contig_314	97	Dlo_pri0453F. 1_g00560.1	AT1G28520.5	Vascular plant one zinc finger protein	0
Contig_72	96	Dlo_pri0050F. 1_g02250.1	AT5G50360.1	Von Willebrand factor A domain protein	2E-84
Contig_114	93	Dlo_pri0072F1_g 00140.1	AT2G44730.1	Alcohol dehydrogenase transcription factor Myb/SANT-like family	2E-18
Contig_1165	93	Dlo_pri0890F-1.1_g00030.1	AT1G33970.3	P-loop containing nucleoside triphosphate hydrolases superfamily	e-119
Contig_576	93	Dlo_pri0015F1_g 00060.1	AT4G31880.2	Transcriptional regulator	e-147
Contig_292	90	Dlo_pri0225F1_g01580.1	AT2G44260.1	DUF946 family protein (DUF946)	0
Contig_319	90	Dlo_pri0138F1_g01600.1	AT4G24220.1	NAD(P)-binding Rossmann-fold superfamily protein	0
Contig_1778	89	Dlo_pri0606F. 1_g00080.1	AT4G16143.2	Importin alpha isoform 2	0
Contig_230	89	Dlo_pri0133F1_g01490.1	AT1G67 430.1	Ribosomal protein L22p/L17e family protein	e-116
Contig_324	83	Dlo_pri0198F1_g 00560.1	AT4G02590.2	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	2E-79
Contig_518	82	Dlo_pri0037F. 1_g04600.1	AT1G 12570.1	Glucose-methanol-choline (GMC) oxidoreductase family protein	0
Contig_594	82	Dlo_pri0439F. 1_g00380.1	AT5G43810.4	Stabilizer of iron transporter SufD/Polynucleotidyl transferase	0
Contig_105	81	Dlo_pri0337F. 1_g00950.1	AT3G 14240.1	Subtilase family protein	0

1.1.4 Summary

From the described results in cv. Kumemaru, direct effectors of the unsilenced *OGI* still remain to be solved. We can propose mainly two potential pathways to activate *OGI*: (i) release from the silencing by the *Kali*-SINE insertion on the promoter region and (ii) an undefined "non-canonical" pathway specific to cv. Kumemaru (Fig. 1-5). For the possibility (i), the *Kali*-SINE was also highly methylated in cv. Kumemaru, as well as other cultivars (Fig. 1-2d), suggesting that epigenetic conditions other than DNA methylation, such as histone methylation/acetylation, might involve release of the silenced *OGI*. These histone modifications are frequently associated with the condition of DNA methylation on transposons (Qian et al., 2012; Du et al., 2015; Zhang et al., 2018). Although we have not assessed the histone methylation/acetylation in the *OGI* promoter region, future comparison of them among a wide variety of the persimmon cultivars may unveil the importance of histone

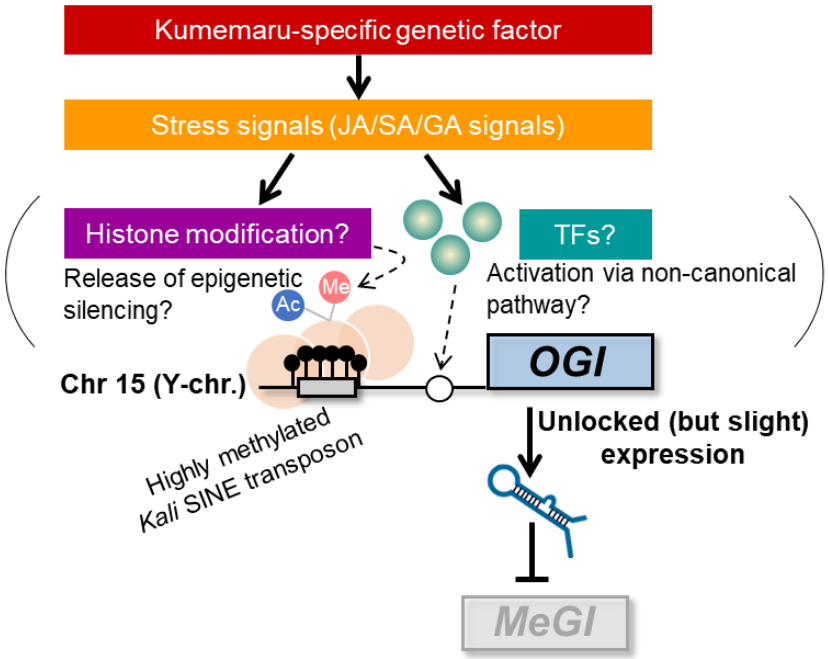

Figure 1-5 Model for derepressed *OGI* expression in cv. Kumemaru.
Inactivation of OGI due to the *Kali*-SINE insertion and the resultant DNA methylation can be activated by two potential pathways: (i) release from the epigenetic silencing, which might result from histone modification, and (ii) an undefined "non-canonical" pathway such as activation via cv. Kumemaru-specific transcription factors. Those are likely induced by stress signals, especially JA/SA/GA signals, under the control of hypothetical Kumemaru-specific genetic (or possibly environmental) factors.

modification. For the possibility (ii), *trans*-acting factors (or transcription factors) which were highly expressed specific to cv. Kumemaru and also related to JA signals, such as *TCP4* and *NGA3* (Fig. 1-4d), might be good candidates to activate *OGI* expression, independent of the effect of *Kali*-SINE. On the other hand, cv. Kumemaru-specific mutations in *cis*-factors (or mutations in the promoter sequences) would be unlikely to be the causal factor to establish the unsilenced *OGI* expression system. Mutations in *cis*-factors of *OGI* would affect only *OGI* (and its downstream pathways), while the JA/SA/GA-related genes detected as the DEGs specific to cv. Kumemaru, were distinct from the downstream pathways of the *OGI* (Akagi et al., 2014; Yang et al., 2019). As given in the model (Fig. 1-5), future assessment about the factors potentially connecting stress signals and the unsilenced *OGI* would have importance to reveal not only the mechanisms determining flower sex balances in a tree, but also the evolution to facilitate the current monoecious system in hexaploid persimmon.

Section 1.2 Genome-wide study on the polysomic genetic factors conferring the male and female flower balances

1.2.1 Introduction

Investigating the genetic architecture of traits with multiple alleles and mixed inheritance patterns in polyploid species is challenging (Dufresne et al. 2014). Genotyping in polyploids is complicated by the possibility of more than two alleles at each locus, and the existence of different heterozygous states. For instance, in the case of autohexaploid persimmon with autohexasomic inheritance (Akagi et al. 2012), we can define five heterozygous states; AAAAAa (5:1), AAAAaa (4:2), and AAAaaa (3:3) AAaaaa (2:4) to Aaaaaa (1:5). Quantitative genotyping of such loci using qPCR is a possible solution (Akagi et al. 2012; Henry et al. 2006), but it is time-consuming. Recent advances in sequencing and genotyping technologies now allow calling high-density single nucleotide polymorphism (SNP), and accurate determination of allele ratio and allele dosage for polyploid genomes. Genome-wide association analysis, using quantitative genotypes, coupled with realistic genetic models (Garcia et al. 2013; Grandke et al. 2016) could shed light on the genetic basis of complex traits.

To examine the genetic factors underlying female-to-male conversion in hexaploid Oriental persimmon, we collected high-density genotype information in a segregating population exhibiting biases in male flower ratio. Considering the nature of polysomic inheritance, we developed genome-wide correlation/association analyses for polyploid persimmon using two different models (Fig. 2-1): (i) an additive model using quantitative genotypes in the form of allelic ratio, and (ii) a diploidized additive model using diploidized genotypes (Rosyara et al. 2016). The results led to the identification of genetic regions and candidate genes potentially

involved in the regulation of female-to-male conversion, providing novel insights into the genetic basis of flexible sex expressions after adaptation to polyploidization or domestication.

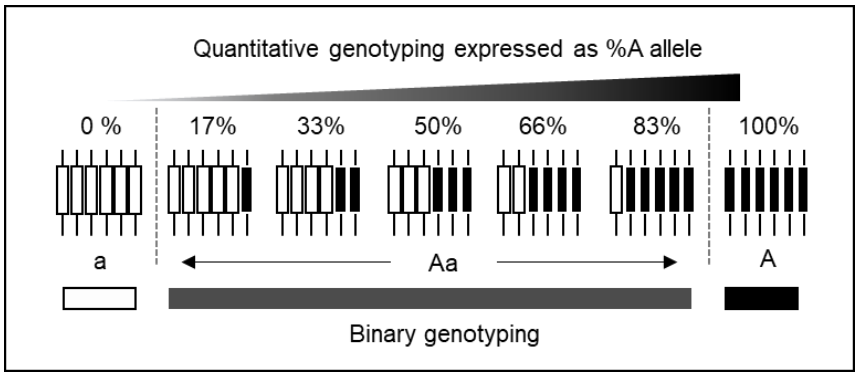

Figure 2-1. Approaches used to identify the genetic factors regulating *OGI* activity.
Two genotyping approaches are available for polyploid populations, that tailor to two different models for the effect of genetic inheritance: (i) in the additive model, heterozygous alleles are further characterized based on their copy number, (ii) in the diploidized additive model, only diploidized genotypes are considered (homozygous or heterozygous), irrespective of the ratio of alleles in the heterozygous states.

1.2.2 Materials and Methods

Plant materials and selection of individuals carrying OGI alleles

The segregating F1 population produced from a cross between 'Yamatogosho' (6 A + XXXXXX) × 'Taishu' (6 A + XXXXYY), named the YTF1 population, was developed in 2009, and exhibited variation in the frequency of male production since the first flowering in 2013-2014. Flower sexuality from a total of 5,016 non-male (female or no flower) parental branches was assessed for 5 years, from 2014 to 2018. A male conversion branch was defined as a non-male parental branch producing male flowers on at least one branch. Two phenotypes were recorded each year. 'Ability of male conversion', reflects the presence or absence of male flowers on any given branch. In other words, if any branch in a tree had switched to male that year, the tree was scored as 1 for this phenotype. If none of the branches had converted, it was scored as 0. The second phenotype, 'Male conversion rate', was calculated by dividing the number of branches that converted to male in a specific year, by the total number of non-male parental branches. Both phenotypes were scored each year. To identify

the genetic factors regulating male conversion, a total 83 individuals, each carrying at least one *OGI* allele were selected. To identify individuals carrying the *OGI* genes, genomic DNA was extracted from leaves of each individual using the CTAB method (Akagi et al. 2016b), and *OGI* was amplified by PCR using two primer pairs, with a program of 95°C for 3 min followed by 32 cycles of 95°C for 20 s, 56°C for 15 s, and 72°C for 1 min.

Genome-wide genotyping

ddRAD-sequencing libraries were prepared according to previous reports (Shirasawa et al. 2016) using 107 individuals of the YTF1 population and their two parents 'Yamatogosho' and 'Taishu'. The libraries were sequenced on an Illumina HiSeq4000 platform, and generating PE100 reads at the Vincent J. Coates Genomics Sequencing Laboratory, University of California Berkeley. The reads were aligned to the reference genome of *Diospyros lotus*, a wild relative close to Oriental persimmon (http://persimmon.kazusa.or.jp/index.html) (Akagi et al. 2020) using the Burrows–Wheeler Aligner with default parameters [version 0.7.12 (https://github.com/lh3/bwa) (Li et al. 2009)]. Based on these parameters, it is possible that some allelic and potential paralogous polymorphisms were occasionally derived from non-specific mapping. The resulting sam files were converted to bam files and subsequently vcf format using bcf/vcftools (Danecek et al. 2011) and Varscan (Koboldt et al. 2012). For genome-wide association/correlation analyses, only the 83 individuals carrying the *OGI* allele were used. Individual genotypes were only considered if coverage > 20. A total of 95,639 heterozygous markers were selected by using bcftools with the following options: minor allele frequency within the population > 0.05 and data were available for at least 50% of the individuals. For detection of transmission ratio distortion, only the 91 individuals of the YTF1 population with total coverage > 200 Mb were used, to decrease the possibility of detecting

the false transmission distortion when using low coverage data. A total of 39,344 loci were selected, with average coverage between 60 and 200 for each of the 91 individuals, minimum coverage in the parents > 60, and by using bcftools with the following options: minor allele frequency > 0.01 and max-missing = 1. The coverage threshold was used to filter uniquely mapped sequences. Transmission distortion ratios of alternative alleles were calculated as the log value of [allele ratio in the YTF1/allele ratio in the parents]. Here, for the 'allele ratio in the parents' values, we calculated the alternative allelic read coverages in the two parents independently, and averaged them. Significant transmission distortion was detected by standardizing the transmission distortion ratios using z-transformation with the threshold of *P*-value < 1E-10.

Genome-wide correlation/association analyses using additive and diploidized additive models

Allele composition (or quantitative genotype) at each SNP locus was estimated from the frequency of alternative alleles in mapped reads. Genome-wide correlation analysis was conducted using a total 40,111 of the 95,639 SNPs, in which duplex–pentaplex genotypes were predicted in at least 23/83 YTF1 individuals (< 28% individuals with at least one recessive allele). To test if a locus affected male conversion in a dosage-dependent manner, Pearson product moment correlation analysis was conducted between the ratio of alternative alleles at each locus and the phenotype values of each individual. The coefficient values were standardized using z-transformation, to estimate *P*-values. Significant association loci were detected with the threshold of −log(*P*-value) > 5.60, representing a Bonferroni-corrected significance threshold of 0.1.

Diploidized genotype at each locus was defined with the threshold of 5% (and 95%) of

alternative alleles in the mapped reads to call heterozygosity. Genome-wide association analysis was conducted using the R package rrBLUP, with a linear mixed model (LMM) (Endelman et al. 2011). Significant association loci were detected with the threshold of −log(*P*-value) > 5.98, representing a Bonferroni-corrected significant threshold of 0.1. Linkage disequilibrium (LD) surrounding the loci highly associated with the sex expression (*P* < 10E−6) was evaluated using squared Pearson's correlation coefficient (R^2) with the −r2 command in the software PLINK version 1.07.

Genome-wide analyses compensated with the effect of OGI allele dosage

Allele dosage of *OGI* was estimated with quantitative PCR using LightCycler 480 (Roche Diagnostics, Mannheim, Germany), according to a previous report (Akagi et al. 2016b). Briefly, using primer sets for *OGI* and the control locus, L5R, quantitative PCR analyses were conducted using THUNDERBIRD SYBR qPCR Mix (TOYOBO, Osaka, Japan). For six individuals of the YTF1 population, due to the lack of DNA, allele dosage on Chr. 15, the peak of 17,052,751-bp on the sex-chromosome was used as a proxy for *OGI* allele dosage (around 18.12-MB on Chr. 15).

To assess the effect of loci on the male conversion traits independently of the effect of *OGI* allele dosage, which substantially affect male production ability, 'partial correlation' analysis was conducted, based on quantitative genotypes. The partial correlation, considering the effect of *OGI* allele dosage, assumes the following equation:

$$y = \frac{(r_{by} - r_{ay} \times r_{ab})}{\sqrt{1 - r_{ay}{}^2} - \sqrt{1 - r_{ab}{}^2}} \qquad (1)$$

where r_{by} is the correlation coefficient between the quantitative genotype at the locus and the phenotype; r_{ay} is the correlation coefficient between *OGI* allele dosage and the phenotype; and r_{ab} is the coefficient between *OGI* allele dosage and the quantitative genotype at the

locus. The partial correlation coefficient values were standardized with z-transformation to estimate *P*-values.

For the diploidized additive model with diploidized genotypes, we assumed the following LMM:

$$y = \beta_o O + Z_g + S_T + \varepsilon \qquad (2)$$

where y is a vector of the phenotype; βo is a fixed effect for *OGI* allele dosage (O) and Z is an incident matrix relating y to g. The variable g models the genetic background of each line as a random effect with Var[g] = Kσ2g, where σ2g is the genetic variance. K is an additive kinship matrix calculated from the genotype data. S is a vector of genotype of each polymorphic site. The variable τ models the additive SNP effect. ε is a matrix of the residual effects with Var[ε] = Iσ2e, where I is an identity matrix and σ2e is the residual variance. To solve the mixed model, we used functions in the R package rrBLUP (Endelman et al. 2011). The additive kinship matrix K was calculated using the 'A.mat' function. The variance components σ2g and σ2e were estimated using the 'mixed.solve' function. *P*-values for the SNP markers were calculated using the 'GWAS' function, which we slightly modified to use quantitative values of *OGI* allele dosage as a fixed effect.

1.2.3 Results and Discussion

Characterization of flower sex expression in the YTF1 population

Within the YTF1 population, 53 of 83 individuals only produced female flowers despite carrying at least one copy of the *OGI* gene (Fig. 2-2b). The monoecious individuals (*N* = 30) showed a wide range of male conversion rate, i.e. conversion from non-male (female and no flowers) parental branches producing male flowers within the 5 years assessment (Fig. 2-2a,c). As suggested in previous reports (Akagi et al. 2016b, Yonemori et al. 1993), the

frequency of male flowers substantially increased as years went by, during the assessment (average male conversion rate are 0.07, 0.13, and 0.15 during the first 3 years). Most of the male parental branches also produced male flowers in subsequent years, which was hypothetically due to maintenance of epigenetic marks (or DNA methylation) on *MeGI* in the parental branches (Akagi et al. 2016a). Consistent with these results, the estimated narrow-sense heritability for male conversion rate, across 5 years, was 0.28. Therefore, we assessed the cumulative ability to produce male branches from female branches throughout the 5 years, to compensate for the limited number of replicates, and reflect the effect of environmental inputs.

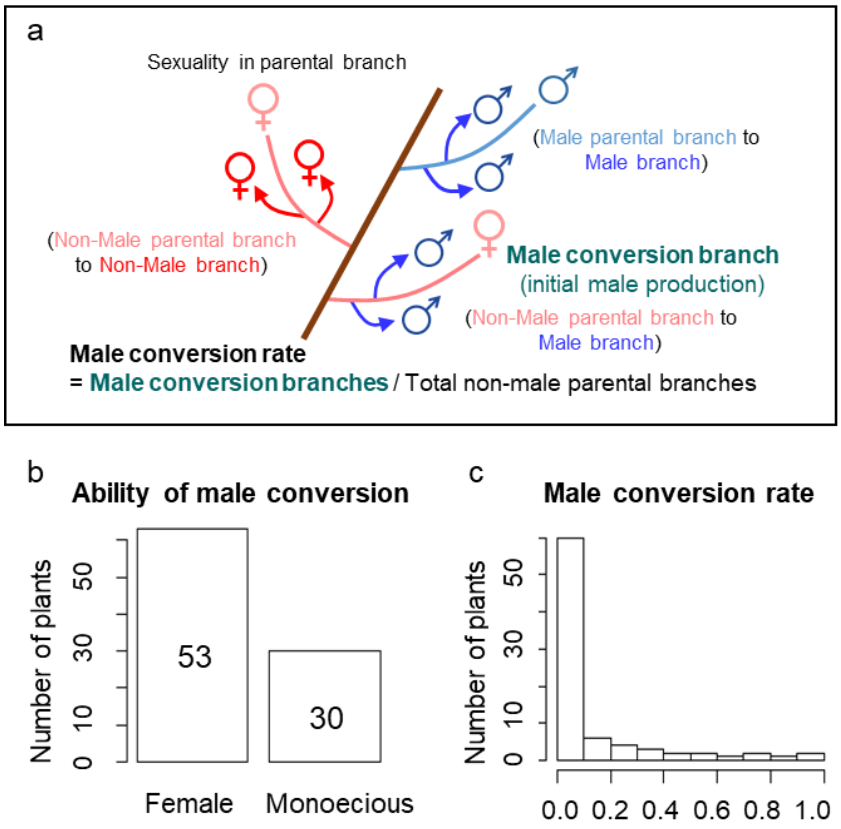

Figure 2-2. Phenotypic diversity in the YTF_1 population.
a, Schematic representation of flowering branches in hexaploid persimmon. Male conversion rate was calculated as the numbers non-male parental branches (female or previously non-flowering) producing male flowers, divided by the total number of non-male parental branches. A male conversion branch was defined as a non-male parental branch producing male flowers on at least one branch. **b-c**, Distribution of phenotypic values in the 83 individuals that carried at least one copy of the *OGI* locus: cumulative ability of male conversion (over 5 years) (**b**) and male conversion rate (**c**).

The Y chromosome contributes to male conversion in a dosage-dependent manner

Genome-wide correlation analyses based on quantitative genotypes using an additive model resulted in several peaks, the strongest of which was located on the sex-chromosome (Chr. 15) (Akagi et al. 2020). This was true both for the ability of male conversion and the male conversion rate traits (Fig. 2-3a,b). Although the male-specific region of the Y-chromosome (MSY) is not included in the reference pseudomolecule (DLO_r1.0.pseudomolecule), it has been genetically anchored within Chr. 15 (around 18.12 Mb) (Akagi et al. 2020). Two highly associated peaks flanked the MSY that includes *OGI* (Fig. 2-3e,f). Some loci within the sex chromosome commonly showed significant correlation with the ability of male conversion and the male conversion rate (Fig. 2-3e,f), suggesting that these two traits were regulated by the same allele combinations. When using the diploidized additive model based on diploidized genotypes, the region surrounding the MSY showed a weak association to male conversion rate (Fig. 2-3c,d). The location of the haploblock including MSY, based on diploidized genotypes (Fig. 2-3g), was consistent with the peaks based on quantitative genotypes (Fig. 2-3e,f).

OGI allele numbers, as determined by quantitative PCR, were highly correlated with the numbers of alleles in the nearest peak produced using the additive model ($r = 0.63$, 78.2% of individuals exhibited identical quantitative genotypes, Fig. 2-3h), and with male conversion abilities (Fig. 2-3i; $P = 0.003$ for male conversion rate with Student t-test). Multiple regression tests with the loci quantitatively associated with male conversion abilities indicated that, amongst these loci, *OGI* exhibited the strongest effect ($P < 0.0002$). These results supported the idea that the *OGI* locus significantly contributes to male conversion ability in a dosage-dependent manner, and suggested that *cis*-elements within or close to *OGI* can play

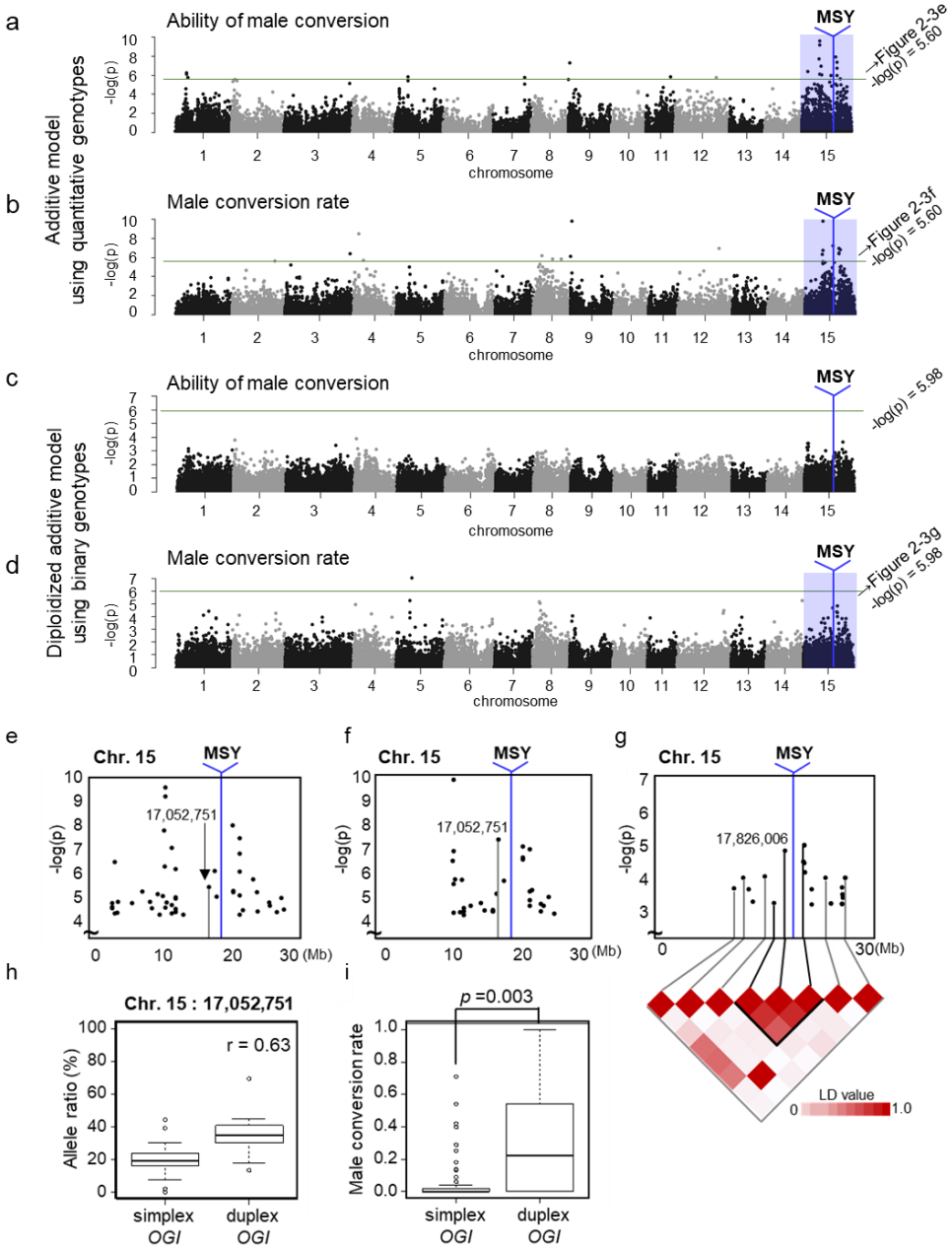

Figure 2-3. Genome-wide correlation/association analysis for male productivity.
a-b, Genome-wide correlation analysis using quantitative genotypes for the ability of male conversion (**a**) and male conversion rate (**b**) traits. **c-d**, Genome-wide association analysis using diploidized genotypes, for the ability of male conversion (**c**) and male conversion rate (**d**) traits. The sex-chromosome (Chr. 15) is highly association with both ability of male conversion and male conversion rates. The blue vertical lines correspond to the MSY, including the *OGI* locus, which is not included in the reference genome of the wild diploid species (*D. lotus*). The green lines indicate −log(*P*-value) > 5.60 for the additive model, and > 5.98 for the diploidized additive model, respectively, representing the Bonferroni-corrected threshold of 0.1. **e-g**, Close-up of the association loci on the sex chromosome, including the MSY region, for ability of male conversion (**e**) and male conversion rate (**f**) based on quantitative genotypes, and for male conversion rate based on diploidized genotypes (**g**). For (**g**), the MSY is contained within the peak locus. **h**, Correlation between allele ratio at the peak (17,052,751-bp) on the sex chromosome (Chr. 15, see e and f), and *OGI* allele dosage, as detected by qPCR. **i**, Association between male conversion rate and the *OGI* allele dosage. The individuals carrying duplex *OGI* alleles exhibited significantly higher male conversion rates (t-test *P*-value shown).

important roles for the activation of *OGI*. This concept is similar to the dosage sex-chr. systemin *Rumex* and *Humulus* (Parker and Clark, 1991), where dioecious sex determination is dependent on X/autosome balance. In these systems, individuals with intermediate sexuality, bearing hermaphrodite flowers, also appear in association with irregular X/autosome ratios between 0.5 and 1.0 (Parker and Clark, 1991).

Identification of candidate loci activating OGI

After compensating for the effect of *OGI* allele dosage (see Materials and methods for the details), the peaks on the sex chromosome, both with quantitative and diploidized genotypes, was significantly reduced (Fig. 2-4a–d). Although the trends were almost identical to those in the original analysis (Fig. 2-3a–d), some peaks, especially those associated with the male conversion rate (Fig. 2-4b,d), became sharper than in the original model. Some of the original putative peaks were also reduced after compensation for the *OGI* dosage effect, presumably because their genotypes were accidentally similar to those of *OGI*. Some of the major peaks, especially on Chr. 5, 8, and 9, were common to the two models (Fig. 2-4a–d). On the other hand, peaks on Chr. 1, 2, and 4 were specific to additive model (Fig. 2-4a,b). Here, we focused on the common candidate loci on Chr. 5, 8, and 9. To understand the combination effect of *OGI* and these three loci, multiplex regression tests were performed between the male conversion rate and the quantitative/diploidized genotypes, resulting in r^2 values of 0.611 and 0.548 for the quantitative and diploidized genotypes, respectively. Importantly, the accuracy of the regression was significantly increased in comparison to the test without compensation for the *OGI* allele dosage (r^2 = 0.561 and 0.492 for quantitative and diploidized genotypes, respectively). These results highlight the importance of quantitative genotyping and compensation for allele dosage in GWASs in polyploids.

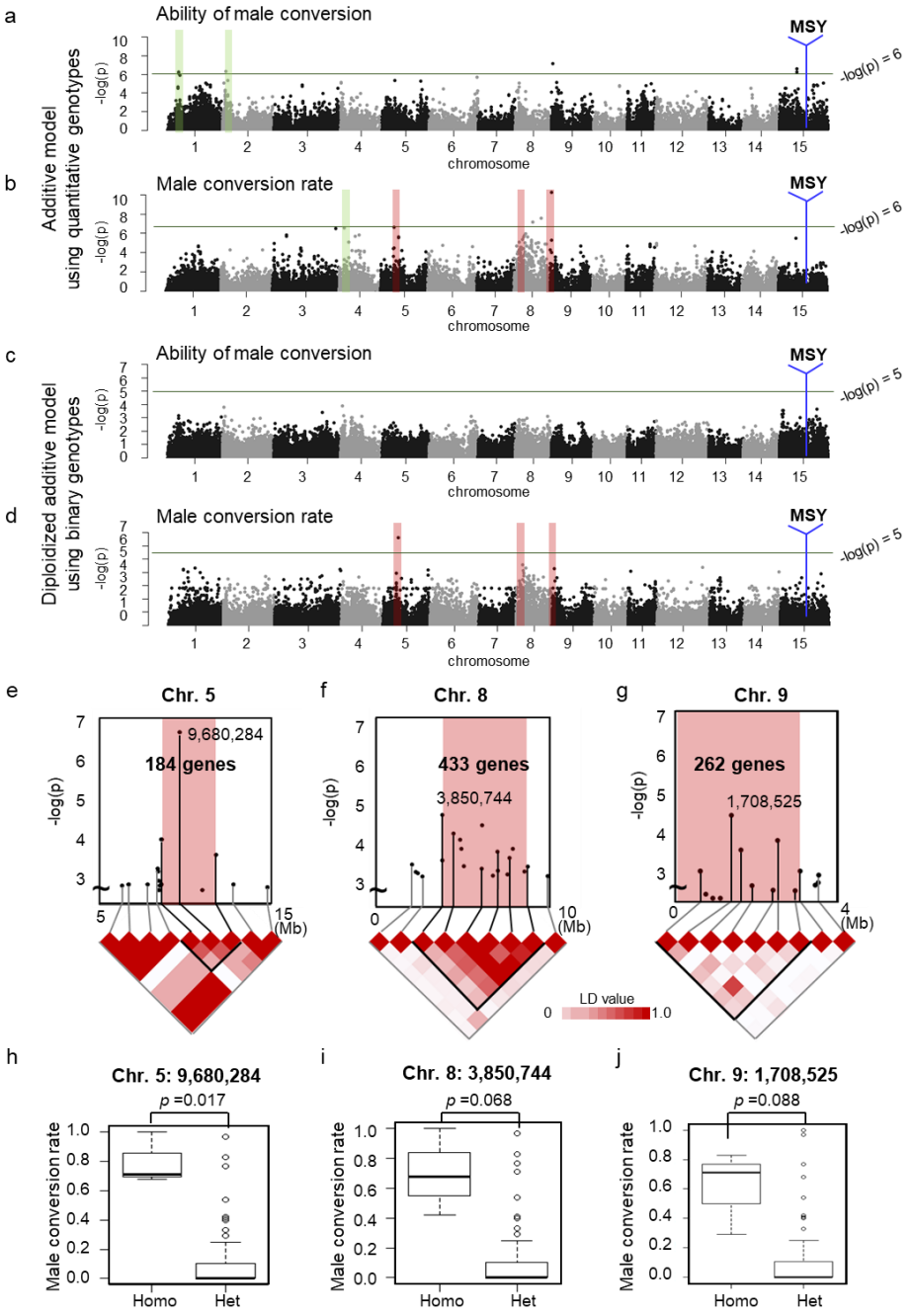

Figure 2-4. Genome-wide correlation/association analysis after normalization for the effect of *OGI* allele dosage.

a-b, Genome-wide partial correlation analysis of the ability of male conversion (**a**) and male conversion rate (**b**), using quantitative genotypes in the additive model after normalization for the *OGI* dosage effect. **c-d**, Genome-wide association analysis of the ability of male conversion (**c**) and male conversion rate (**d**) after normalization for the *OGI* dosage effect. The sex-chromosome (Chr. 15) was not association with either ability of male conversion or male conversion rate. The vertical red and green bars indicate the location of the association peaks. The red bars show the peaks commonly detected in the analyses using the additive and diploidized additive models, for male conversion rate. **e-g**, Close-up of the association peaks on Chr. 5 (**e**), 8 (**f**), and 9 (**g**), with LD, using diploidized genotypes. The main haploblocks on Chr. 5, 8, and 9 included 184, 433, and 262 genes, respectively. **h-j**, Genotypes (Homo: homozygous, Het: heterozygous) of the most significant loci on Chr. 5 (**h**), Chr. 8 (**i**), and Chr. 9 (**j**).

Both the additive and diploidized additive models showed consistent associations between Chr. 5 and male conversion rate (Fig. 2-4b,d). The haploblock including the peak summit was maintained over a 3 Mb region (ca 8.6–11.8 Mb), and contained 184 genes (Fig. 2-4e). For Chr. 8, the peak spanned the region between 3.8 and 8.8 Mb, included 433 genes (Fig. 2-4f), and was associated with male conversion rate in both the additive and diploidized additive models (Fig. 2-4b,d). The quantitative genotypes for the peak summit of the Chr. 9 observed in the additive model (Chr9:1,287,557) ranged from nulliplex to duplex (Fig. 2-4b). The peak covered 4 Mb of the sub-telomeric region and contained 262 genes (Fig. 2-4g). For all three of these peaks identified from the diploidized additive model analysis, nulliplex individuals showed significantly higher male conversion rate than heterozygous individuals (Fig. 2-4h-j), suggesting that loss of that particular allele is associated with positive regulation of male conversion.

Significant LD was maintained over ~2–5 Mb of the peaks on Chr. 5, 8, and 9 ($r^2 > 0.15$; Fig. 2-4e-g). These regions included 184, 433, and 262 genes, respectively. Next, we aimed to detect potential candidate with the genes underlying these peaks that hypothetically regulate the activation of *OGI*, which is normally silenced via SINE insertion on the promoter and the resultant DNA methylations (Akagi et al. 2016a). Considering that, in all cases, homozygous genotypes act recessively for *OGI* activation loss-of-function of genes regulating DNA/histone methylation to silence transcription would be good candidates. Histone remodeling genes such as Chromatin remodeling (Shaked et al. 2006), SET domain protein (Shafiq et al. 2014), and histone deacetylase 19, thought to be involved in jasmonic acid and ethylene signaling in Arabidopsis (Zhou et al. 2005), are all located under the Chr. 5 peak. Candidate regulators within the Chr. 8 peak included SWADEE homeodomain homolog 1 (Law et al. 2011), histone deacetylase2 (Ito et al. 2000), and Jumonji family (Chen

et al. 2011), while the peak of Chr. 9 included the PHD finger protein (Peña et al. 2006) and SET domain protein 38 (Hariganeya et al. 2009). Future in-depth sequence analysis of these candidates might help in the identification of candidate polymorphisms potentially involved in *OGI* activation/repression.

Examination of GWAS in polyploid genome

Genome-wide analyses in polyploid genomes are often challenging in terms of (i) genotype calling, and (ii) haplotype phasing (Grandke et al. 2014). For genotyping, we adopted the allele ratio per locus with high coverage for quantitative genotypes, as previously reported in potato and in blueberry (Rosyara et al. 2016, Ferrão et al. 2018). This approach, combined with the Pearson correlation was successful at detecting the effect of Y-chromosome dosage for *OGI* activation and other potential candidates. Furthermore, the compensation with *OGI* allele dosage with two approaches, (i) partial correlation (in correlation analysis using quantitative genotypes), and (ii) rrBLUP with the covariate of *OGI* dosage (in GWAS using diploidized genotypes), improved our ability to detect candidate loci. On the other hand, issues involving haplotype phasing still remain to be solved, since polyploids have more than two haplotypes per reference region, which are generally difficult to define, not only for conventional binary genotypes but also for quantitative genotypes. For instance, in our hexaploidy samples, the presence of many (>2) homologous haploblocks caused LD inconsistent with the order of the SNPs markers, despite of the F1 segregation line in which LD decay completely depends on recombinations (Fig. 2-4e), as observed in autotetraploid blueberry as well (Ferrão et al. 2018). In autohexaploid sweet potato (*Ipomoea batatas*) with 90 chromosomes (2n = 6x = 90), 96 linkage groups (LGs) were generated, using only double-simplex SNPs showing a Mendelian segregation ratio in an F2 progeny (Shirasawa et al.

2017). Such construction of well-separated LGs would allow QTL analysis with conventional tools, but the targets would be limited to simplex loci and, unless the genomes were significantly different, much of the genomic space would not be included. Here, we were able to include multiplex alleles in our genome-wide association analysis, and our results suggest that a simple GWAS approach can be used to identify polysomic candidates, if sufficient markers are available.

1.2.4 Summary

Our results indicated that, in the additive model based on quantitative genotypes, individuals with higher allele dosages of *OGI* tended to exhibit more frequent male conversion. This situation would propose the following two hypothesis for *cis*-regulation of the silenced *OGI* (Fig. 2-5a): (i) multiple *trans*-acting factors can access the *OGI* promoter, each with independent sequence binding affinity, or (ii) epigenetic *cis*-factors of *OGI*, such as DNA/histone methylation, are modified in each allele independently. On the other hand, in both the additive and the diploidized additive models, all the loci on Chr. 5, 8, and 9 were

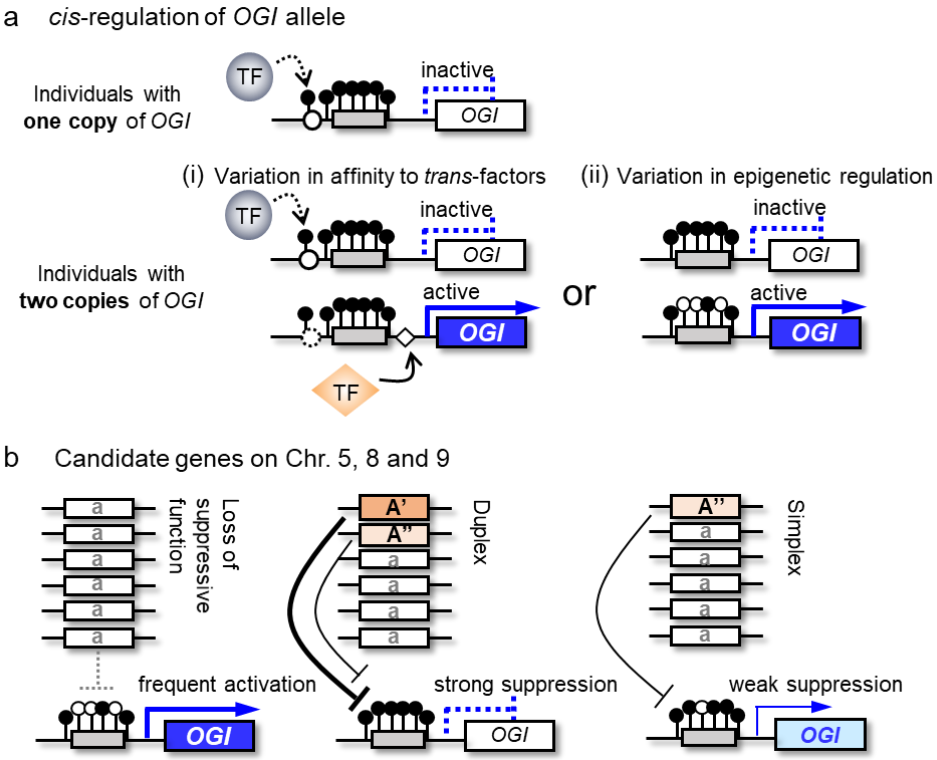

Figure 2-5. Hypothetical model for the mechanisms underlying *OGI* regulation in hexaploid persimmon.
a, Two hypothesis can be proposed for the dosage-dependent cis-regulation of *OGI* (see Fig. 2-3i): (i) variation in cis-factors within the *OGI* promoter sequences modulate the effect of the transcription factors accessing to the *OGI* promoter, or (ii) variation in the *OGI* promoter sequence results in a variation in epigenetic status. Both of these scenarios can be consistent with varying probability *OGI* activation, depending on the environmental conditions. **b**, Model for the action of candidate genes on Chr. 5, 8, and 9, acting in dosage-dependent manner. Since recessive homozygosity had the highest effect on male conversion (*OGI* activation) in all three cases (Fig. 2-4h-j), the model predicts that the dominant (functional) alleles act to suppress *OGI*. The dosage-dependent effect (Fig. 2-4b) could originate either from multiple dominant alleles (A' and A") with slightly different ability, or additive action of a single A allele to suppress *OGI* expression.

significantly associated male conversion rate (Fig. 2-4). This suggests that, although recessive alleles at these loci lack the function to maintain *OGI* suppression, the other alleles also vary in their ability to suppress *OGI* and/or have additive effects for silencing of *OGI* (Fig. 2-5b). These complex situations, which are not rare in handling polysomic genetic factors, would represent how they can contribute to the fine tuning of sexual balances in an individual, in a way that is not applicable to a diploid situation.

Chapter 2 Non-canonical somaclonal sex conversions in genetically female persimmon

2.1.1 Introduction

Long-lived (perennial) plants, especially vegetatively propagated tree crops, often produce visible somaclonal mutations (i.e. 'bud sports'). The mutant phenotypes may be maintained for at least a few years, but they occasionally revert back to the original phenotype, likely because of epigenetic regulation. In oil palm trees (*Elaeis guineensis*), the appearance of a somaclonally abnormal fruit type (i.e. mantled) is related to the decreased DNA methylation of transposons (Ong-Abdullah et al. 2015). In almond trees (*Prunus dulcis*), noninfectious bud failure is associated with clonal age and epigenetic status (Fresnedo-Ramírez et al. 2017). For both of these major tree crops, the appearance of altered phenotypes was thought to be mainly due to their long lives or long propagations. In addition, in the *D. kaki* cultivar Saijo, which lacks *OGI* and originally bears only female flowers, we detected somaclonal mutant male (SMM) branches and revertant female (RF) branches (Fig. 2-1a–f). This noncanonical male system, which does not involve *OGI*, has been observed in several independent Saijo trees in a specific area (Esumi et al. 2015), implying that epigenetic regulation contributes to this frequent sex conversion. Furthermore, the production of noncanonical male flowers has occasionally been observed in the cultivars Jiro and Fuyu, which also lack *OGI* (Yakushiji and Nakatsuka 2007, Akagi et al. 2014b). These findings suggest that certain epigenetic mutations to specific factors in genetically female individuals can induce noncanonical sex conversions in hexaploid persimmon species.

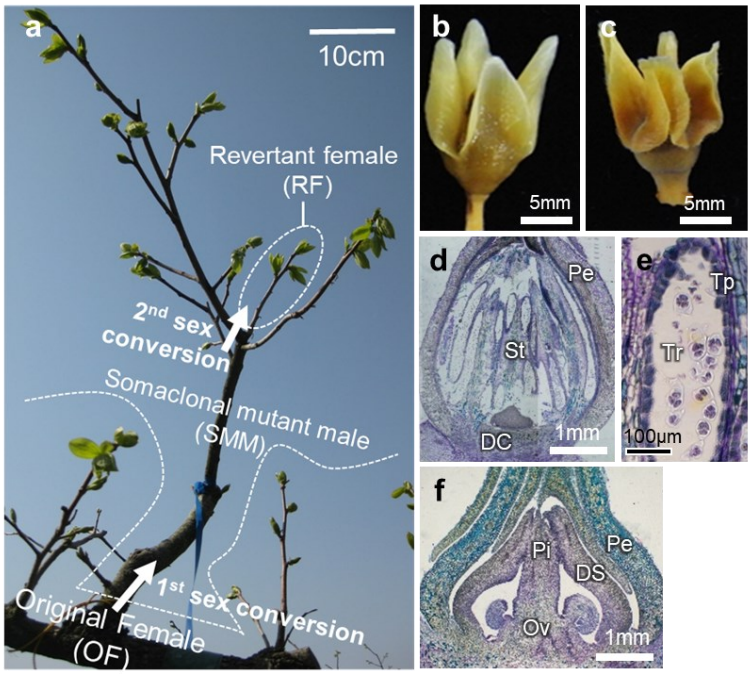

Figure 2-1. Characterization of somaclonal sex conversions in persimmon.
a, Appearance of an SMM branch from a branch of a genetically female tree (OF) (first sex conversion) and an RF branch arising from an SMM branch (second sex conversion) in a persimmon tree (Saijo) grown in Shimane prefecture, Japan. **b-f**, Appearance of SMM (**b**) and RF (**c**) flowers at maturity (stage 3; Yang et al. 2019). Dissected SMM (**d** and **e**) and RF (**f**) flowers were stained with toluidine blue. The SMM flower bears intact stamens (St) and a deficient carpel (DC) (**d**). Image of SMM stamen at stage 3 and enclosed tetrads (TR) with a semi-mature tapetum cell layer (Tp) (**e**), which is consistent with normal male flowers in stage 3 (Yang et al. 2019). The RF flowers formed an intact pistil (Pi) and ovary (Ov), but a deficient stamen (DS) and no developing gametes, which is consistent with normal female flowers in stage 3.

To characterize the mechanism underlying the production of noncanonical male flowers in a genetically female tree, we applied multi-omics approaches to examine early-stage flowers from the original female (OF), SMM and RF branches of Saijo trees. A comparison of the multitiered somaclonal mutations in an individual tree enabled us to elucidate the epigenetic regulation responsible for sex conversions and assess their potential contribution to gender gaps in persimmon. These insights may clarify the importance of the flexibility in epigenetic layers in a long-lived individual to overcome the genetic robustness of plants.

2.1.2 Materials and Methods

Plant materials

Flower buds and leaves on the OF, SMM and RF branches of the hexaploid *D. kaki* cv. Saijo trees grown in Shimane prefecture, Japan (35.47°N, 132.85°E), were harvested on April 4, 2018, which corresponds to the early developmental stage before organs differentiated (Akagi et al. 2016a). Normal male or female flower buds from Taishu and Saijo trees grown in the experimental orchard of Kyoto University were harvested during the same developmental stage (April 1–5, 2015 and 2016). Normal female flower buds and leaves from other branches in the mutated tree and other trees of the same cultivar, Saijo in the orchard, where the mutated branch was found, were harvested on April 9, 2019. The collected samples were frozen at −80°C until used for DNA/RNA extractions.

PCR expression analysis

Total RNA was extracted from the collected flower samples with the PureLink® Plant Reagent (Invitrogen, USA) and then purified by a phenol/chloroform extraction. First-strand cDNA was synthesized with the ReverTra Ace® qPCR RT Master Mix (Toyobo, Japan) for a subsequent analysis of *MeGI* expression by PCR. The constitutively expressed *DkActin* gene was used for standardizing the expression levels (Akagi et al. 2019a, Akagi et al. 2019b).

Preparation and sequencing of mRNA-Seq and smRNA-Seq libraries

To characterize the expression dynamics during the noncanonical sex conversions, the mRNA-seq and smRNA-seq data of OF, SMM, and RF flowers were compared with that for the canonical male–female conversion in some monoecious cultivars carrying the *OGI* gene (Akagi et al. 2016a, Yang et al. 2019). To construct mRNA-Seq libraries, total RNA was purified with the Dynabeads mRNA Purification Kit (Invitrogen, USA). Illumina sequencing libraries were prepared with the KAPA HyperPrep Kit (Roche, Switzerland) as previously

described (Yang et al. 2019). Thirteen cycles of PCR enrichment were completed, followed by a DNA cleanup step with AMPure XP beads (AMPure:reaction = 0.8:1; Beckman Coulter, USA). The constructed libraries were sequenced with the HiSeq 4000 system (50-bp single-end reads; Illumina).

To construct smRNA-Seq libraries, the smRNA fraction was isolated from total RNA and concentrated with the mirVana miRNA Isolation Kit (Life Technology, USA). Approximately 200 ng of concentrated smRNA was used to construct the libraries with the NEBNext Small RNA Library Prep Set (NEB, USA) as previously described (Akagi et al. 2016a). Fifteen cycles of PCR enrichment were completed, followed by a DNA cleanup step with AMPure (AMPure:reaction = 1.1:1) to remove self-ligated adapter dimers. The constructed libraries were sequenced with the HiSeq 4000 system (50-bp single-end reads).

The Illumina sequencing was conducted at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, USA. The raw sequencing reads were processed with custom Python scripts developed in the Comai laboratory as previously described (Akagi et al. 2014a). Briefly, reads were split based on the index information and trimmed for quality (Phred sequence quality > 20 over a 5-bp sliding window) and to eliminate adapter sequence contamination. Reads < 35 and < 19 bp were excluded from the mRNA-Seq and smRNA-Seq analyses, respectively.

Expression profiling

The mRNA-Seq and smRNA-Seq reads were aligned to the reference coding sequences of the diploid Caucasian persimmon species *D. lotus* (Akagi et al. 2019a, Akagi et al. 2019b) with the default parameters of the Burrows–Wheeler Aligner (version 0.7.12; https://github.com/lh3/bwa; Li and Durbin 2009). The read counts per coding sequence were

generated from the aligned SAM files with a custom R script (Akagi et al. 2016a) to calculate the normalized gene expression levels. For the mRNA-Seq data, the DEGs between OF/RF and SMM flowers were detected by a DESeq analysis (sharing mode = fit-only; Robinson et al. 2010). The DEGs were further filtered based on RPKM and FDR values (RPKM >1.0 and FDR <0.1). Putative functions of each gene were assigned according to a BLASTX search of the TAIR10 database (http://www.arabidopsis.org/index.jsp).

Bisulfite PCR-Seq analysis

We conducted a bisulfite PCR-Seq analysis of the 300-bp 5′ *MeGI* promoter region in the early-stage flower buds and leaves to determine the DNA methylation levels in the SMM and OF/RF (defined as 'mutated branch'), in the other branches in the same tree, and in other trees of the same cultivar, as well as in the canonical male-producing cultivars (cv. Taishu) carrying the *OGI* gene. Genomic DNA was extracted with the cetyltrimethylammonium bromide method (Akagi et al. 2014b) or DNeasy Plant Mini Kit (Qiagen, Netherlands), after which it was treated with the EZ DNA Methylation-Gold kit (Zymo Research, USA) to deaminate and convert non-methylated cytosine residues into uracil residues. Bisulfite-treated *MeGI* promoter sequences were amplified with Epi Taq HS (Takara Bio Inc, Japan) and specific sense- and antisense-direction primer sets. The amplicons were subjected to an end-repair step, after which an A-base overhang and an adapter were added as described in the mRNA Illumina libraries construction section. Five cycles of PCR enrichment were completed, followed by a DNA cleanup step with AMPure (AMPure:reaction = 0.7:1). The constructed libraries were sequenced with the HiSeq 4000 system (150-bp paired-end reads). The forward and reverse reads of the bisulfite PCR-Seq data were separately mapped to the 5′ promoter region of the *MeGI* gene as described and then analyzed with the Methylpy

program (https://github.com/yupenghe/methylpy) to determine the DNA methylation levels as described by Kawakatsu et al. (2016). Only residues with a coverage of > 20 were used for analyzing DNA methylation. The samples harvested in 2019 were conducted for the bisulfite sequencing of only sense sequence as the index. The significances of DNA methylation levels between samples were calculated with Fisher's exact test.

MethylC-Seq analysis

The MethylC-Seq libraries were prepared with genomic DNA extracted from flower buds as previously described (Urich et al. 2015). Briefly, DNA was fragmented to 350-bp segments, followed by an end-repair step and the addition of an adenosine tail and an adapter. The resulting adapter-ligated DNA was isolated by two rounds of purification with AMPure XP beads (200–600 bp) and then subjected to a bisulfite conversion with the EZ DNA Methylation-Gold kit. The bisulfite-converted DNA molecules were enriched by four cycles of PCR. The amplicons were purified using AMPure (AMPure:reaction = 1:1). The methylC-Seq libraries were sequenced with the HiSeq 4000 system (100-bp paired-end reads). The MethylC-Seq reads were divided into two groups (forward and reverse reads) and then mapped to the primary genomic scaffold sequences of diploid *D. lotus* (Akagi et al. 2019a, Akagi et al. 2019b) with the Methylpy program. Genome-wide DNA methylation levels were calculated as methylated residues/filtered read count with coverage > 10 within 500-kb bin. The DNA methylation was visualized using the circos plot program as described by Kawakatsu et al. (2016). The DNA methylation levels for gene bodies and the 1,000-bp regions flanking the coding sequences were visualized using the deepTools program (https://deeptools.readthredocs.io/en/develop/). The data in the read count files were filtered (coverage > 5) and separated based on the DNA methylation contexts (CGN, CHG and CHH).

The DNA methylation levels on gene bodies and flanking regions were calculated with Compute Matrix and visualized using the plotProfile program of deepTools as described by Kawakatsu et al. (2016).

Correlation between DNA methylation and expression levels

The genome-wide methylation patterns in the 5′ 1-kb promoters and gene bodies were calculated as previously described by Kawakatsu et al. (2016). The methylated cytosines of all three contexts (CGN, CHG and CHH) in the promoters and gene bodies were extracted from the read count file with a custom Python script and the generic feature format (gff) file of the diploid *D. lotus* (Akagi et al. 2019a, Akagi et al. 2019b). The methylation patterns were calculated as methylated residues/filtered read count with coverage > 50. The correlation between the methylation patterns and mRNA/smRNA expression levels was calculated based on the Pearson correlation analysis. The data for the methylation patterns were combined with the filtered mRNA and smRNA data (RPKM > 1). The combined data were standardized based on the Z-score and visualized with a heatmap. The relationships between the 300-bp promoter methylation bias in SMM/female flowers with the mRNA bias and smRNA bias in SMM/female flowers were visualized with plotted figures.

2.1.3 Results and Discussion

Production of noncanonical male flowers on SMM branches depends on regulated MeGI expression

The canonical male flower production system in individuals carrying *OGI* depends on the accumulation of smRNAs targeting *MeGI* transcripts (*smMeGI*), which substantially downregulates *MeGI* expression (Akagi et al. 2016a). The *MeGI* expression levels were

much lower in the SMM flowers than in the two types of female flowers (Fig. 2-2a). These results suggested that changes in *MeGI* expression affect this noncanonical sex conversion. The smRNAs, especially 21- and 22-nucleotide smRNAs, targeting *MeGI* accumulated specifically in SMM flowers in the noncanonical system, which may involve the noncanonical RNA-dependent DNA methylation pathway and the action of Pol II and RDR6 (Matzke and Mosher 2014, Matzke et al. 2015), and the accumulation pattern was consistent with that in the canonical male production system (Fig. 2-2b). A comparison of the genome-wide smRNA surveys of the SMM flowers and the two female flowers (OF and RF) indicated that relatively few genes exhibited SMM-specific up- or downregulated expression. Moreover, the smRNA bias in the SMM/OF ratio was highly correlated with that in the SMM/RF ratio ($r = 0.65$; Fig. 2-2c). Only 20 genes exhibited a significantly enriched accumulation of smRNAs in an SMM-specific manner (>5-fold higher in SMM flowers than in female flowers, RPKM > 10). These genes included *MeGI*, which was associated with the most biased smRNA accumulation in SMM flowers (SMM/OF = 67.7 and SMM/RF = 79.1; Fig. 2-2c). An analysis of the mRNA-Seq data for early-stage flowers revealed 12 genes that were differentially expressed between the SMM flowers and the female (OF and RF) flowers according to the DESeq analysis (RPKM >1, *P*-value < 0.1; Fig. 2-2d). Of these genes, the expression levels of *MeGI* and *AGL6*, which encode transcription factors, were significantly downregulated in SMM flowers (*P*-value = 0.01 and 0.05, respectively), which is consistent with their expression patterns in canonical male flowers (Yang et al. 2019). The smRNA-Seq and mRNA-Seq analyses suggested that the smRNA-dependent regulation of *MeGI* expression is responsible for the sex conversion between SMM and female (OF/RF) flowers in the noncanonical system as

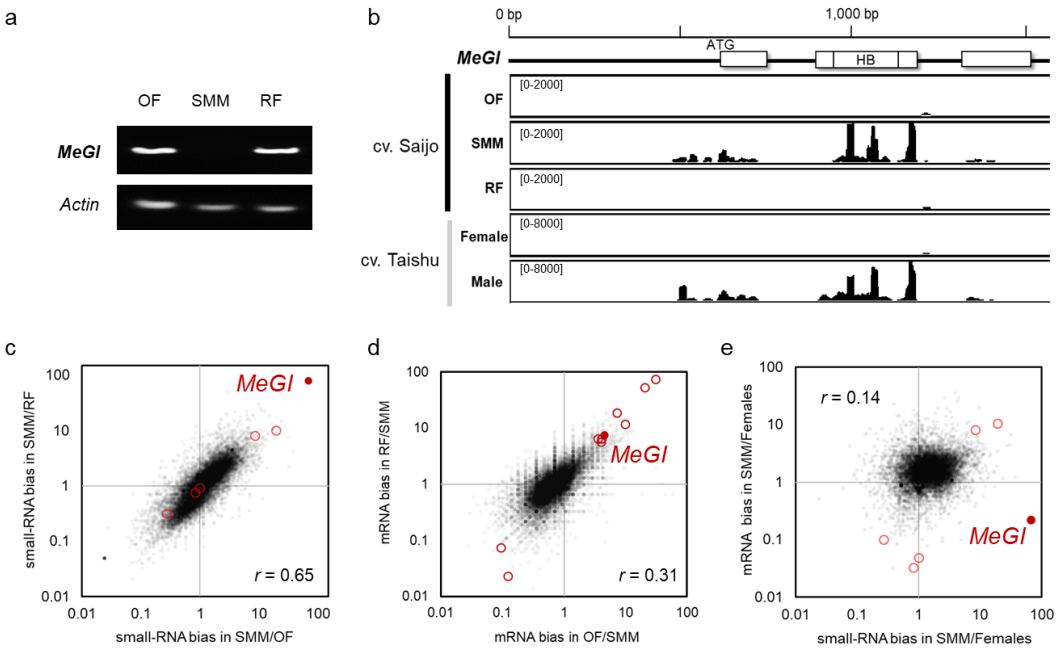

Figure 2-2. SMM-specific downregulation of *MeGI* expression via the production of the noncanonical smRNAs targeting *MeGI*.
a, The *MeGI* expression level in early-stage flowers (stage 1; Yang et al. 2019) was assessed by qRT-PCR. The *Actin* gene served as the reference control. **b**, Accumulation of smRNA across the *MeGI* gene and its promoter region in OF, SMM, and RF flowers (stage 1) from Saijo trees and in normal male and female flowers from monoecious (canonical male-producing) Taishu trees. A substantial accumulation of smRNAs was observed in the male flowers of both cultivars, with a similar accumulation pattern. **c-d**, Distribution of the smRNA and mRNA expression biases between male and female flowers during the somaclonal sex conversions. Correlation of the smRNA and mRNA expression biases between SMM/OF (X axis) and SMM/RF (Y axis). Five smRNA DEGs (RPKM >1) and twelve mRNA DEGs between SMM and OF/RF flowers are highlighted in red. **e,** Distribution of the relationship between smRNA accumulation and mRNA expression in the comparison of SMM and the OF/RF flowers. Six DEGs between SMM and OF/RF flowers are highlighted in red. Of these DEGs, *MeGI* specifically exhibited downregulated expression, with a considerable accumulation of smRNA.

well as in the canonical system. Furthermore, among the differentially expressed genes (DEGs) between SMM and OF/RF flowers, only *MeGI* was associated with a substantial accumulation of smRNAs (Fig. 2-2e), which is also consistent with the canonical sex-determination system (Akagi et al. 2016a). These observations suggested that the biased accumulation of trans-acting smRNAs has little effect on sex differentiation in canonical and noncanonical systems.

DNA methylation of the MeGI promoter in the noncanonical male production system

The weighted DNA methylation level in SMM flowers was 8.0%, which was significantly

higher than that of OF or RF flowers (0.60% and 0.50%, respectively, *P*-value < 2.2e−16 and < 2.2e−16 with Fisher's exact test, respectively; Fig. 2-3a). More precisely, the methylation levels of all three contexts (CGN, CHG and CHH) were substantially higher in SMM flowers than in the OF and RF flowers (Fig. 2-3a). These results were indicative of an association between the methylation levels in the *MeGI* promoter and the OF–SMM–RF sex conversions. Similar to the DNA methylation pattern in flower buds, leaves also showed higher DNA methylation in SMM-specific manner, although their methylation levels were quite lower than those in flowers (Fig. 2-3b; 0.60% for SMM, while 0.40% and 0.45% for OF and RF with *P*-value < 2.2e−16 and < 2.2e−16 for Fisher's exact test against SMM, respectively). The CGN and CHG contexts showed clear differences in comparison of SMM and OF/RF, while the CHH methylation showed only slight (but statistically significant) differences (*P*-

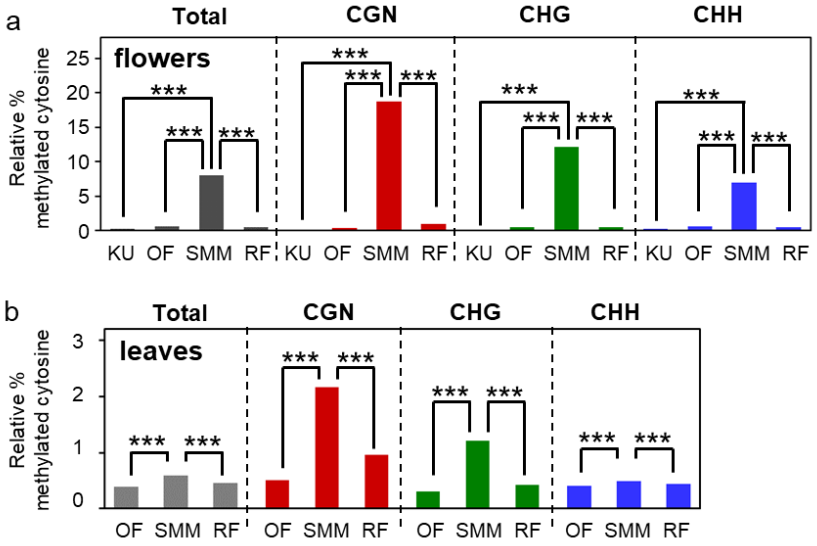

Figure 2-3. DNA methylation of the *MeGI* promoter in flowers and leaves during the noncanonical sex conversion.
The DNA methylation levels in the *MeGI* 5' promoter region (300 bp from the start codon) in the three contexts (CGN, CHG and CHH), in early flower developing stage (stage 1, Yang et al. 2019) (**a**) and in young leaves (**b**). The DNA methylation levels were substantially higher in SMM flowers on all the three contexts than those in OF/RF flowers from Saijo trees grown in Shimane prefecture. This tendency was consistent with the DNA methylation levels in leaves in comparison of OF–SMM–RF, in all three contexts, although the biases were fundamentally smaller than in flowers. In comparison of female flowers, the methylation levels in OF and RF were significantly higher than in female flowers from Saijo trees planted at Kyoto University (KU). This tendency was consistent with the DNA methylation levels in leaves in comparison of OF–SMM–RF, in all three contexts, although the biases were fundamentally smaller than in flowers. Bars indicate SE, ****P* < 0.001.

value = 6.5e−9 and 0.001, respectively; Fig. 2-3b). In monoecious cultivars, the DNA methylation of the *MeGI* promoter is critical for the maintenance of smRNA accumulation. The DNA methylation level for maintaining or releasing smRNAs varies between cultivars, with independent threshold levels, as revealed in a previous study involving treatments with DNA demethylation agents (Akagi et al. 2016a). Our results were suggestive of the possibility that the more extensive DNA methylation of the *MeGI* promoter in SMM than in OF and RF with branch unit resulted in an SMM-specific smRNA accumulation, consistent with the canonical system.

Genome-wide comparison of DNA methylation levels during somaclonal sex conversions

To assess the genome-wide tendencies of DNA methylations in OF, SMM and RF flowers, we analyzed the DNA methylomes by whole-genome bisulfite sequencing with DNA from early-stage OF, SMM and RF flowers as the template. Genome-wide analyses revealed 13,840 CGN, 31,037 CHG and 147,532 CHH contexts. To characterize the differentially methylated regions in the persimmon genome among SMM and OF/RF samples, we calculated the weighted methylation levels with 500-kb bin windows in the whole genome (Fig. 2-4a-f) and within the 1,000-bp segments flanking gene bodies (Fig. 2-4g-i). The genome-wide tendencies indicated that SMM-enriched CGN and CHG methylations were observed on both genes and repetitive regions (Fig. 2-4a,b). The CGN and CHG sequences exhibited SMM-enriched methylation levels relative to the OF (*P*-value = 3.1e−123 for CGN and *P*-value = 3.2e−20 for CHG, respectively) and RF levels (*P*-value = 1.5e−313 for CGN and *P*-value = 9.0e−61 for CHG, respectively) according to the paired t-test (Fig. 2-4d,e). In addition, the methylation levels were significantly higher in OF flowers than in RF flowers (*P*-value = 4.6e−216 for CGN and *P*-value = 2.4e−29 for CHG; Fig. 2-4d). These methylation-

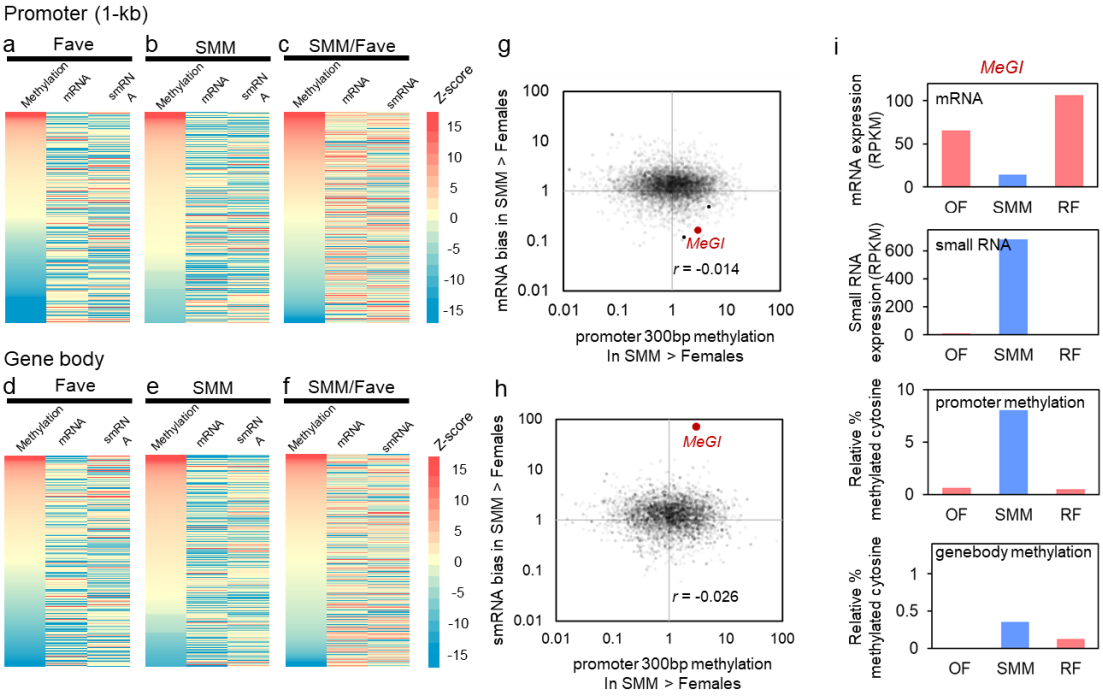

Figure 2-4. Changes to genome-wide DNA methylation during the noncanonical sex conversions. **a-c**, Distribution of DNA methylation levels in the three contexts [CGN (**a**), CHG (**b**) and CHH (**c**)]. From the outer layer, DNA methylation levels in 500-kb bin in OF (1), SMM (2) and RF (3) and the densities of genes (4) and repetitive regions (5) were given. DNA methylation levels were significantly higher in the SMM, on CGN and CHG in genome-wide, remarkably on repetitive regions. **d-f**, Average whole-genome DNA methylation levels in all three contexts [CGN (**d**), CHG (**e**), CHH (**f**)]. **g-i**, Each context was given in red, green and blue. Bars indicate SE. ***$P < 0.001$. DNA methylation levels in 50-bp bins of the gene bodies and in the 1,000-bp flanking regions in the CGN (**g**), CHG (**h**) and CHH (**i**) contexts. Data for the OF, SMM and RF flowers are presented in red, blue and gray, respectively. SS, start site (start codon); TS, termination site (stop codon).

level tendencies were consistent with the relative DNA methylation levels for the *MeGI* gene in the SMM, OF and RF flowers and leaves (Fig. 2-3). However, there were no significant differences in the genome-wide CHH methylation level in the three flower types (*P*-value = 0.65 for OF vs. SMM and 0.76 for RF vs. SMM; Fig. 2-4c,f). These results implied that the male-specific greater methylation in the CGN and CHG contexts might be associated with the noncanonical male production system.

Our bin analysis of methylation levels surrounding complete genes indicated that the CHG context exhibited an SMM-specific enrichment of methylation both in the gene body and in the flanking regions (Fig. 2-4h). In contrast, there were no significant differences in the methylation of the CGN context in the OF and SMM flowers, whereas the methylation level

decreased in the transition from SMM to RF flowers (Fig. 2-4g). The CHH context was similarly methylated in the three flower types (Fig. 2-4i). Thus, the DNA methylation levels of the gene bodies and flanking regions were consistent with the genome-wide methylation patterns, with only some inconsistencies (e.g. CGN context of SMM and OF flowers). The DNA methylation in *MeGI* was significantly changed in all three contexts, CGN, CHG and CHH, between SMM and females (OF and RF) (Fig. 2-3), whereas potentially coordinated genome-wide DNA methylation showed fundamental changes only in CGN or CHG. This difference in the methylated contexts between in *MeGI* and at whole-genome might be a clue for the mechanism of this somaclonal conversion to male.

We also assessed the correlation between relative methylation levels and mRNA/smRNA expression levels. The DNA methylation patterns for the 5′ 1-kb promoters and gene bodies were not associated with mRNA and smRNA dynamics (3,860 methylated genes; read coverage >50) in the female (OF and RF) and SMM flowers (Fig. 2-5a-c and d-f for the DNA methylation of the promoter regions and gene bodies, respectively). Furthermore, the DNA methylation biases in the promoter regions of the SMM and female (OF and RF) flowers were not significantly correlated with the mRNA (Fig. 2-5g, $r = -0.014$) or smRNA (Fig. 2-5h, $r = -0.026$) expression biases. The DNA methylation and mRNA or smRNA expression levels were consistent for only a part of genes, such as *MeGI* (Fig. 2-5g,h). Therefore, the methylation patterns clearly linked to sex determination may be limited to specific genes, such as *MeGI*.

Most of the genes in the genome were insensitive to the epigenetic variations between male and female flowers, although the epigenetic status of *MeGI* dramatically affected its expression pattern and the resulting sex of flowers. Our results imply that sex-specific epigenetic variations can regulate the expression of only some genes in persimmon flowers.

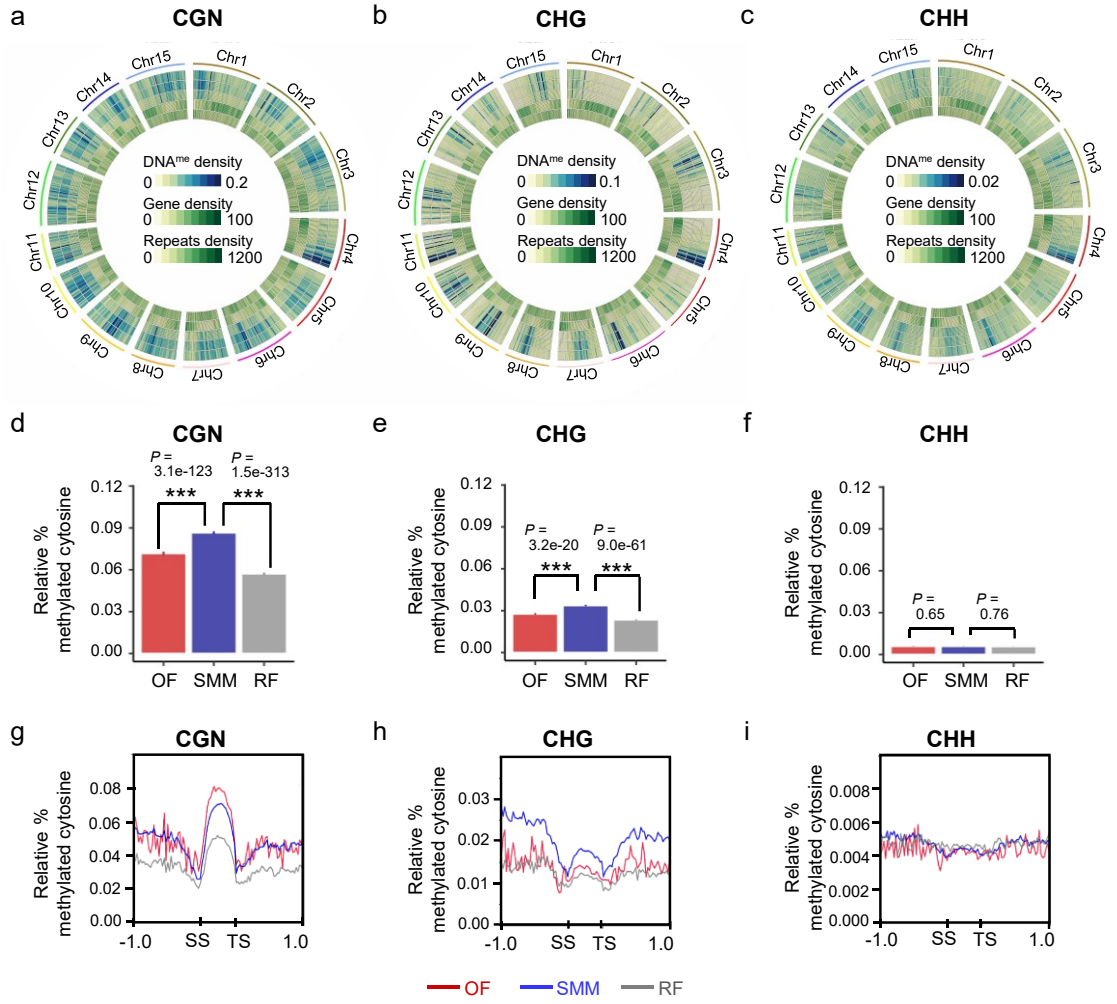

Figure 2-5. Correlation between DNA methylation and mRNA/smRNA expression levels.
a-f, Heatmap visualization of the correlations between mRNA/smRNA expression levels and relative DNA methylation levels of the 5′ 1-kb promoters (**a-c**) or in the gene bodies (**d-f**). Data are provided for the average OF/RF flower (Fave), SMM flower and the SMM/Fave ratio. The methylation and expression levels are indicated as standardized values across the genes. **g-h**, No clear correlations ($r < 0.1$) and tendencies were observed between the genome-wide DNA methylation and gene expression levels. Genome-wide distribution of the DNA methylation and mRNA (**g**) or smRNA (**h**) biases between Fave and SMM flowers. The DNA methylation biases in Fave and SMM flowers were correlated with the expression levels of specific genes (e.g. *MeGI*), but not with the genome-wide gene (or smRNA) expression levels. **i**, Summary of the MeGI mRNA and smRNA expression, and DNA methylation levels in the promoter and gene body in *MeGI*, in comparison of OF-SMM-RF.

Future investigations of the genome-wide DNA methylation patterns in various cultivars (or genotypes) may provide deeper insights into the importance of epigenetic variations in long-lived persimmon trees. From a viewpoint of environmental factors involved in this noncanonical male production, empirical knowledge concerning the cultivation of persimmon trees suggests that young and vigorous plants (or branches) tend to produce more female

flowers, which may support the possibility that a genome-wide epigenetic status in plants affects *MeGI* methylation. Considerable epigenetic flexibility in long-lived plants, or in trees, is often associated with long vegetative propagation (Bräutigam et al. 2013). In persimmon, the Saijo tree bearing SMM branch has been maintained for approximately 20 years after the grafting of young scions. On the other hand, the Saijo tree in Kyoto, which has been maintained for > 50 years, has so far not exhibited any sex conversions. This suggests that plant age alone cannot explain the epigenetic flexibility and resulting sex conversion in persimmon.

2.1.4 Summary

Our results suggested that the noncanonical male production system fundamentally shares a regulatory pathway with the canonical system with *OGI* allele, where both depend on the epigenetic regulation of *MeGI*, given in a model of Fig. 2-6. In the noncanonical system, the *MeGI* methylation level might be coordinated with the genome-wide fluctuation in DNA methylation, although there still might be a space of discussion regarding this potential link. The SMM-specific genome-wide hypermethylation, especially in the CHG context, was synchronized with the *MeGI* methylation in the mutated branch unit (Fig. 2-3). This might suggest unrevealed mechanism to regulate fluctuation of CHG methylation in *D. kaki*-specific manner, which would be a potential causal factor for this somaclonal sex conversion. These insights into establishment of noncanonical male production system would clarify the importance of the flexibility in epigenetic layers in a long-lived individual to overcome the genetic robustness of plants.

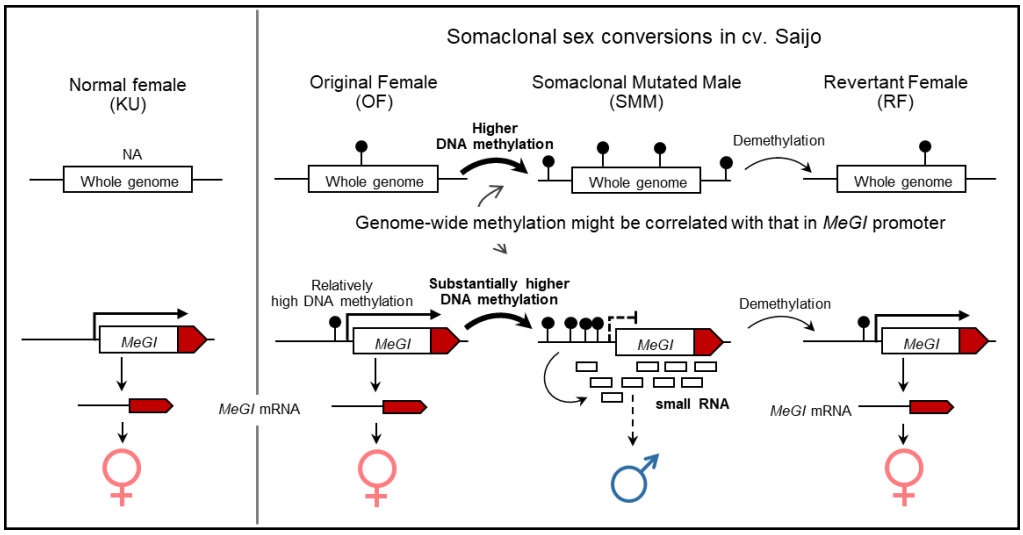

Figure 2-6. Model for the noncanonical sex conversion system in a genetically female persimmon tree. The basic methylation levels in Saijo trees in the region including the mutated tree might be slightly higher than those in normal Saijo tree. Noncanonical sex conversions are associated with the epigenetic changes to MeGI, which might be coordinated with the genome-wide DNA methylation tendencies during the transition to and from male flowers. This change in DNA methylation to derive male branch occurred in a branch unit.

Chapter 3 Reinvention of hermaphroditism via activation of a *RADIALIS*-like gene in persimmon

3.1.1 Introduction

In addition to the conversion from dioecy to monoecy, hexaploid *D. kaki* further exhibits conversion from male to hermaphrodite flowers, in response to natural environmental signals, or to artificial cytokinin treatment (Yonemori et al. 1993). Regarding the occasional production of hermaphrodite flowers, we would start from that phytohormones variously affect sex expressions in *D. kaki*, as well as in a number of other plant species. In *Cucurbitaceae*, endogenous ethylene pathway induced female differentiation (Boualem et al. 2015), and brassinosteroid pathway is mainly involved in sex determination in maize (Hartwig et al. 2011). Treatment of cytokinin induced male differentiation in hemp (*Cannabis sativa* L.), while female differentiation in spinach (*Spinacia oleracea*), grape (*Vitis vinifera*), and also kiwifruit (Grant et al. 1994, Akagi et al, 2018). Auxin induced male differentiation in asparagus (*Asparagus officinalis*) and hop (*Humulus lupulus*), while female differentiation in hemp (Grant e al. 1994, Negi and Olmo 1966). These results suggested many ways to potentially invent sex-determination system in plants. Regarding *D. kaki*, consistent with grape or kiwifruit, treatment with cytokinin to male flowers in the early flower developing stage result in the conversion into hermaphrodite flowers (Yonemori et al. 1990), which is consistent with the higher concentration of internal cytokinin in female and hermaphrodite flowers than in male (Sun et al. 2017; Li et al. 2021). Importantly, these cases of reversion to hermaphroditism have not been observed in diploid *Diospyros*. Still, the fundamental mechanisms that only highly polyploid *Diospyros* species can express hermaphroditism, would be in a mystery. In this study, we aimed to identify the molecular pathways associated

with polyploid-specific reversions to hermaphrodite flowers in hexaploid *D. kaki*.

3.1.2 Materials and Methods

Plant Materials

The developmental stages of Oriental persimmon (hexaploid *D. kaki* Thunb.) buds/flowers were defined in previous studies (Yang et al. 2019) and Caucasian persimmon (diploid *D. lotus*) buds/flowers were also defined based on observation of gynoecium and androecium development: stage 2 (organs are developing; 18 April 2018) and stage 3 (organs are maturing; 1 May 2018) (see Fig. 2-1). For the development of artificial hermaphrodite flowers, ten parts per million (ppm) of N-(2-chloro-4-pyridyl)-N'-phenylurea (synthetic CK) was applied to developing male flowers of hexaploid *D. kaki* and diploid *D. lotus* at stage 1.5 (12 April 2018), which were maintained at Kyoto University (Kyoto, Japan). To construct mRNA-Seq libraries, male, female, NH, and AH flowers from hexaploid *D. kaki* were sampled from 23 cultivars and male and CK-treated male flowers were samples from diploid *D. lotus*.

RNA isolation and Illumina sequencing

Total RNA was extracted from flowers as described previously (Yang et al. 2019) using the PureLink® Plant Reagent (Invitrogen, Carlsbad, CA, USA), after which the mRNA was isolated using the DynabeadsTM mRNA Purification kit (Ambion, Foster City, CA, USA). The mRNA-Seq libraries were prepared with the KAPA RNA HyperPrep Kit (Roche, Switzerland) as previously described (Yang et al. 2019), followed by a DNA cleanup step with AMPure XP beads (Beckman Coulter, USA; AMPure : reaction = 0.8 : 1). The purified libraries were quantified with a Qubit 2.0 fluorometer (Invitrogen) and then sequenced on the Illumina Hiseq 4000 system at the QB3 Genomic Sequencing Laboratory of UC Berkeley

(http://qb3.berkeley.edu/gsl). The resulting SR50 sequencing reads were analyzed at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley. Raw sequencing reads were processed using custom Python scripts developed in the Comai Laboratory as previously described (Akagi et al., 2014). All Illumina sequencing data have been deposited in the DDBJ database: Short Read Archives (SRA) database (SRA Submission ID: DRA013154, Run IDs: DRR332477–332618).

Expression profiling in the four flower types of hexaploid D.kaki

mRNA-Seq reads from the four flower types collected from 23 cultivars were aligned to the reference coding sequences (CDSs) of the diploid Caucasian persimmon, *D. lotus* (http://persimmon.kazusa.or.jp/index.html), using the default parameters of the Burrows‐Wheeler Aligner (version 0.7.12) (https://github.com/lh3/bwa) (Li and Durbin, 2009). These settings allow the mapping of variable allele sequences from hexaploid *D. kaki* (Yang et al. 2019). Read counts per CDS were determined from the aligned SAM files using a custom R script to calculate the RPKM for each gene. To examine the gene expression dynamics in female, male, NH, and AH flowers collected in stages 2 and 3, a principal component analysis (PCA) was conducted using the genes with RPKM > 1, using prcomp in R. Additionally, Pearson's product-moment correlation coefficients were calculated for each flower type, using the cor.test with the ‘pearson’ argument in R, with average gene expression levels as the parameters. Clustering of gene expression levels in each flower type was conducted using the hclust function in R, with average expression levels for each sexual type as variables, and visualized heatmap using the ComplexHeatmap package (Gu et al. 2016).

Identification of differentially expressed genes

Differentially expressed genes (DEGs) between NH, AH, and male flowers were detected with edgeR (Robinson et al. 2010; McCarthy et al. 2012), using the paired-test option, an in-house R script. NH, AH, and male flowers from the same cultivars were paired and used as biological replicates. The DEGs were filtered according to RPKM, bias and *P*-values (RPKM > 1, bias > 1.2-fold, and *P*-value < 0.1). Putative functions of each gene were determined with a BLASTX search of the TAIR10 database (https://www.arabidopsis.org/index.jsp).

Clustering the differentially expressed genes and construction of the co-expression networks. Clustering of the DEGs combined with comparison between NH, AH, and M flowers in stage 2 were conducted for each sexual type using the hclust in R, with average gene expression levels as parameters. Among 14 clusters, Clst 10, 4, and 5 were defined as NH-specific, AH-specific, and NH/AH-shared upregulation clusters, respectively. Three gene co-expression networks were constructed using the DEGs within Clst 10, 4, and 5, respectively, based on the WGCNA package which is a representative algorithm used for developing co-expression networks (Langfelder and Horvath 2008). The soft-thresholding powers for signed networks in Clst 10, 4, and 5 were set at 9, 9, and 10, respectively, to produce the scale-free topology of the networks. In the co-expression networks, genes were represented by nodes, and the correlation values (weight) between two genes were calculated by Pearson's correlation coefficient. The genes in the same module were visualized with the Cytoscape program (Shannon et al. 2003), and only lines with weights greater than 0.171, 0.05, and 0.06 were visualized in the three modules, respectively, to be dominated by a few highly connected genes and an increasing number of less-connected nodes. A GO enrichment analysis was performed on the genes in the three networks with agriGO (http://bioinfo.cau.edu.cn/agriGO). A threshold for the significance of enriched GO term was

set as a *P*-value < 0.1.

Treatment with ABA and transcriptomic analysis

To test the effect of ABA signaling on the development of hermaphrodite from male flowers, (+)-cis,trans-Abscisic acid (ABA) was applied to male flowers of hexaploid *D. kaki* cultivars and diploid *D. lotus* at stage 1.5 (15 April 2019 for *D. kaki* Taishu and 11 April 2019 for *D. kaki* Zenjimaru, *D. kaki* Shogatsu, and *D. lotus*). Ten and 100 mg/L ABA or water (control) including 0.01% Tween 20 was sprayed onto male flowers. The developing flowers were collected from two *D. kaki* cultivars, Taishu and Zenjimaru, and *D. lotus* for ABA-treated and control male flower at stage 2 (26 April 2019 for Taishu and 29 April 2019 for Zenjimaru, Shogatsu, and *D.lotus*). mRNA-Seq libraries were constructed and sequenced as described above. The Illumina reads were aligned to the reference CDSs of the diploid Caucasian persimmon, *D. lotus* (http://persimmon.kazusa.or.jp/index.html) using the default parameters of the Burrows-Wheeler Aligner. The read counts per CDSs were calculated based on the aligned SAM files using a custom R script. Differences in gene expression levels between the ABA-treated male and the control male flowers were assessed with the edgeR package in R (version 3.0.1).

Measurement of plant hormones content

To measure the endogenous concentration of ABA and other phytohormones, the developing male and hermaphrodite flowers at stage 2 (20 April 2019), and the young flowers of control and *DkRAD* transgenic *Nicotiana tabacum* lines (corolla about 2 cm long) were collected. Phytohormone contents were analyzed as previously described (Gupta et al. 2017) with minor modifications. Briefly, frozen samples were ground with extraction

solvent [1% acetic acid (AcOH), 80% acetonitrile (MeCN)] containing internal standards [D6-ABA, D4-SA, 13C6-IAA, D2-GA1, D2-GA4, D5-tZ, D6-iP (OlChemim, Czech Republic), D2-JA (Tokyo Kasei, Japan), and 13C6-JA-Ile (kindly gifted by Dr. Yusuke Jikumaru, Riken, Japan; present affiliation: Agilent Technologies Japan, Ltd)]. After sample purification by Oasis HLB and MCX columns (Waters Corporation, Milford, MA, USA), the acidic fraction was loaded onto an Oasis WAX column (Waters Corporation), and the main acidic fraction was eluted with 1% AcOH and 80% MeCN. A solution comprising 3% formic acid and 97% MeCN was then applied to the WAX column to elute SA. For phytohormone quantification, the Agilent 1260–6410 Triple Quad LC/MS system (Agilent Technologies Inc., Santa Clara, CA, USA) equipped with CAPCELLPAC ADME HR S2 column (OSAKASODA, Japan) and XDB-C8 Guard column (Agilent Technologies Inc.) were used.

Transformation and Phenotyping of Transgenic Plants

The full length of the RAD transcript sequence was PCR amplified using PrimeSTAR Max (TaKaRa, Japan) from cDNA cloned from flower buds from *D. kaki* cv. Zenjimaru. The amplicons were cloned into the pENTR-D-TOPO vector (Invitrogen, Grand Island, NY), and subsequently recombined into the Gateway compatible pFAST-G0255 downstream of the CaMV35S promoter by LR clonase. The resulting plasmid pFAST-RAD was used for plant transformation in *A. thaliana*. The vectors of pPLV26-RAD (Dlo_pri0016F.1_g05470.1 and Dlo_pri0120F.1_g01930.1) were constructed using the In-Fusion Cloning kit (TaKaRa, Japan) to connect the gene under the control of CaMV35S in pPLV26 vector (Shimada et al. 2010). The resulting plasmid pPLV26-RAD was used for plant transformation in *N. tabacum*. *A. thaliana* plants were transformed with pFAST-RAD as described by Akagi et al. (2018). Briefly, *A. thaliana* ecotype Columbia-0 plants were grown at 18℃ under white light (400-750

mm) with a 16-h light / 8-h dark photoperiod until transformation. The pFAST-RAD construct was introduced into *Agrobacterium tumefaciens* strain EHA105 by electroporation along with the helper vector pSOUP. Wild-type *A. thaliana* plants were transformed using a floral-dip method. The putative transgenic plants were screened using the Basta herbicide (phosphinothricin, glufosinate). Tobacco plants (*Nicotiana tabacum*) cv. Petit Havana SR1 was grown in vitro under white light with 16-h-light and 8-h-dark cycles at 23℃ until transformation. The binary construct was introduced into *Agrobacterium tumefaciens* strain EHA105 as described above. Young petioles and leaves of tobacco plants were transformed by the leaf disk method (Horsch et al. 1985). Transgenic plants were selected on Murashige and Skoog medium supplemented with 50 μg/mL kanamycin.

Transcriptomic analysis in the DkRAD-overexpressed A. thaliana flowers and D. kaki leaves

The mRNA-Seq analysis in *A. thaliana* developing flowers (stage 8; Smyth et al. 1990) was conducted with three independent *DkRAD*-overexpressed lines, which were transformed with pFAST-RAD, and the control lines. The mRNA-Seq analysis in *D. kaki* was conducted with the leaves transiently overexpressed with *DkRAD* and the control samples. Young leaves of cv. Zenjimaru were inoculated with 20 μl of Agrobacterium suspension (OD = 1.0) transformed by pPLV26-RAD and the control (pPLV26), respectively per leaf, with three biological replicates. After 3 days incubation at room temperature, total RNAs were extracted from the inoculated regions. The mRNA-Seq libraries were prepared and sequenced as described above. The Illumina reads were aligned to the reference CDSs of *A. thaliana* (Araport 11) BWA by using the default parameters. Differentially expressed genes were detected with the DESeq package in R (version 3.0.1) for Arabidopsis using 3 individuals per replicate, and with the paired edgeR package in R (version 3.0.1) for persimmon leaves using

3 pairs of leaves. The DEGs detected were compared to the DEGs identified between NH/AH and male flowers.

3.1.3 Results and Discussion

Morphological characterization of hermaphrodite flowers in two independent ways in hexaploid persimmon

In hexaploid *D. kaki*, monoecious individuals carrying the semi-inactivated *OGI* occasionally produced hermaphrodite flowers instead of male flowers in two independent ways, (i) naturally (natural hermaphrodite, NH), or (ii) following cytokinin (CK) treatment of male flowers (artificial hermaphrodite, AH) (Fig. 3-1a,b). In both cases, hermaphrodite flowers could be differentiated from male flowers, immediately after the androecium/gynoecium primordia differentiation (stages 2-3; defined in Yang et al. 2019). In both NH and AH, cell proliferation in carpels of developing male flowers was reactivated to eventually form an enlarged carpel (Fig. 3-1c-k), while male functions were not significantly affected (Fig. 3-1l-n). These results suggested that the development of hermaphrodite flower resulted from reactivation of gynoecium growth in male flower, which is normally repressed due to epigenetic silencing of *MeGI* expression (Akagi et al. 2016a). However, *MeGI* expressions in NH and AH flowers were not significantly different from that observed in developing male flowers (Fig. 3-1o, P = 0.82-0.97 at stage 2 and P = 0.96 at stage 3 with a one-sided Student's t-test), and substantially lower than in developing female flower (P = 0.003-0.004 at stage 2 and P = 5.0e-4 at stage 3 with a one-sided Student's t-test), in each of the 23 monoecious persimmon cultivars which frequently develop hermaphrodite flowers. Consistent with the *MeGI* expression pattern, the genes acting as the downstream of *MeGI*, including *SVP* and *KNOX*/*OFP*/*GRF* (Yang et al. 2019), did not exhibit any significant difference in expression

levels between male and NH or AH flowers, in stages 2 and 3 (P = 0.85–0.94 with a one-sided Student's t-test). These results suggested that the canonical genetic sex determining pathway, depending on *MeGI* expression, was not directly associated with the formation of NH and AH.

To identify the genes involved in the development of these hermaphrodite flowers, we

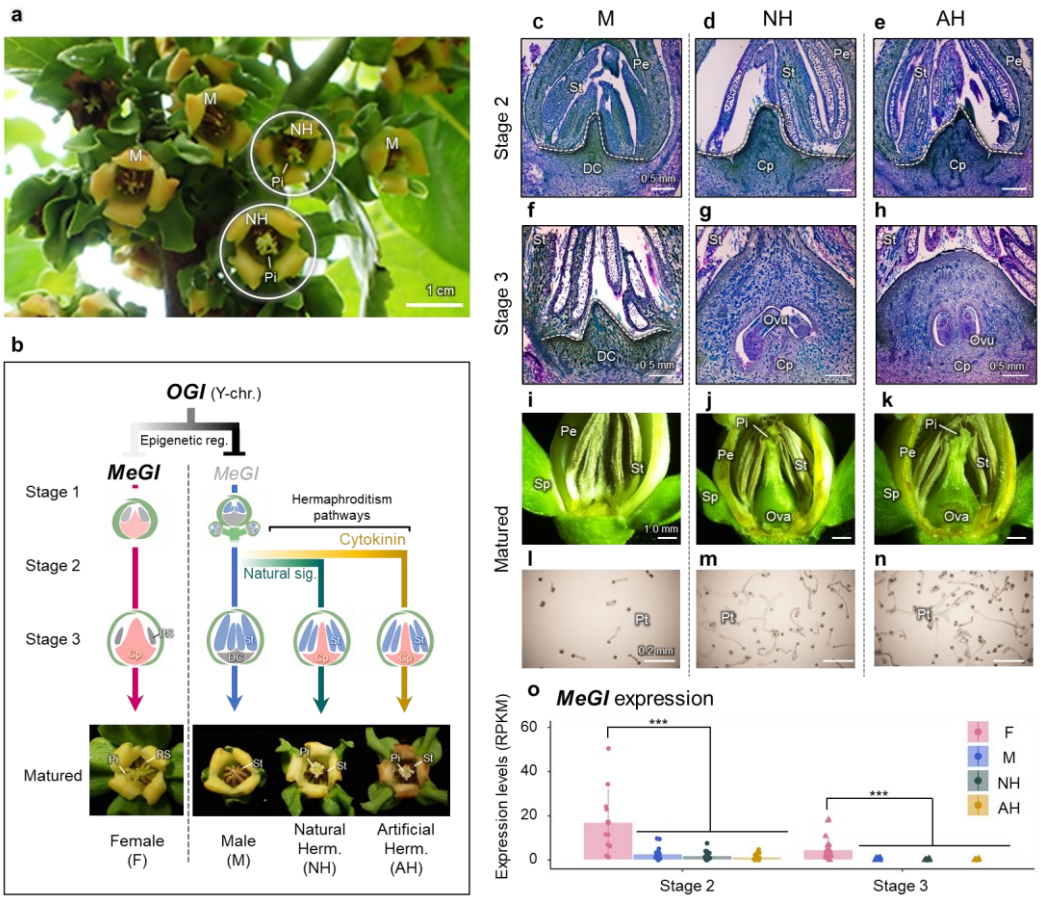

Figure 3-1. Structure and development of hermaphrodite flowers derived from male flowers in hexaploid *D. kaki*.
a, Production of hermaphrodite flowers in a male branch. In monoecious *D. kaki*, sexes (or male and female flowers) are located on separated branches (Akagi et al. 2016a). Hermaphrodite flowers are randomly produced on otherwise male branches. **b**, Schematic model for the development of flowers from male flowers, triggered by either natural environmental signals (natural hermaphrodite, NH) or cytokinin treatment (artificial hermaphrodite, AH). Cp, carpel; DC, defective carpel; Pi, pistil; RS, rudimentary stamen; St, stamen. Note that the browning of AH petals (in the panel **b**) is not a stable characteristic. **c-n**, Morphological characterization of male, NH, and AH flowers at various organ developmental stages. Flower cross-sections stained with toluidine blue at stage 2 (**c-e**) and stage 3 (**f-h**). Dissected flower (**i-k**) and pollen fertility (**l-n**) at the mature stage. DC, deficient carpel; Ova, Ovary; Ovu, ovule; Pe, Petal; Pi, Pistil; Pt, Pollen tube; Sp, Sepal. Scale bars: 0.5 mm for stages 2 and 3, 1.0 mm for the mature stage, and 0.2 mm for pollen germination. o, MeGI expressions in NH and AH flowers were not upregulated from male flowers (P = 0.82-0.97), and significantly lower than female flowers during the organ developmental processes (P < 0.005 at stage 2 and P < 8.0e-4 at stage 3 with a one-sided Student's t-test). *** indicate P < 0.001 for statistical significance.

produced Illumina mRNA-Seq libraries from NH, AH, male (M), and female (F) flowers, at stages 2 and 3 in each of the 23 monoecious persimmon cultivars. The overall expression patterns observed in the four flower types at stages 2 and 3 (RPKM > 1; 15,019 genes) highlighted distinct differences in expression dynamics between NH and AH flowers (Fig. 3-2a for Pearson correlation analysis). At stage 2, the transcriptome of male and NH flowers was moderately correlated (R = 0.56), while AH flowers were little correlated with male and NH flowers (R = -0.05 for vs male, R = 0.02 for vs NH flowers). On the other hand, at stage 3, NH and AH flowers showed a highly correlation (R = 0.76), and both these hermaphrodites were more correlated with female (R = 0.46-0.48) than with male (R = 0.33-0.43). These results suggested the possibility that the development of NH and AH flowers triggered gynoecium differentiation in distinct manners (stage 2), but shared activation of the gynoecium development pathway at the maturing stage (stage 3) (Fig. 3-2b).

To uncover the pathways specific to the reactivation of gynoecium development in hermaphrodite flowers, we focused on 1,315 upregulated and 3,914 downregulated differentially expressed genes (DEGs) when compared to stage 2 male flowers (FDR < 0.1, bias > 1.2-fold). A total of 5,163 upregulated and downregulated DEGs were grouped into 14 clusters, based on their expression patterns in the four flower types (Fig. 3-2b). Surprisingly, most of the NH flower upregulated genes including in Clst 9 and Clst10, were downregulated in AH flowers (Fig. 3-2b), suggesting opposite expression modifications in NH and AH flowers at stage 2, which was consistent with the described expression dynamics (Fig. 3-2a).Here, we focus on three of these clusters: (i) NH-specific upregulation (Clst 10), (ii) AH-specific upregulation (Clst 4), and (iii) upregulation shared between NH and AH (Clst 5) (Fig. 3-2c,d) to detect each hermaphrodite-specific pathway independent from feminization pathway in female flowers.

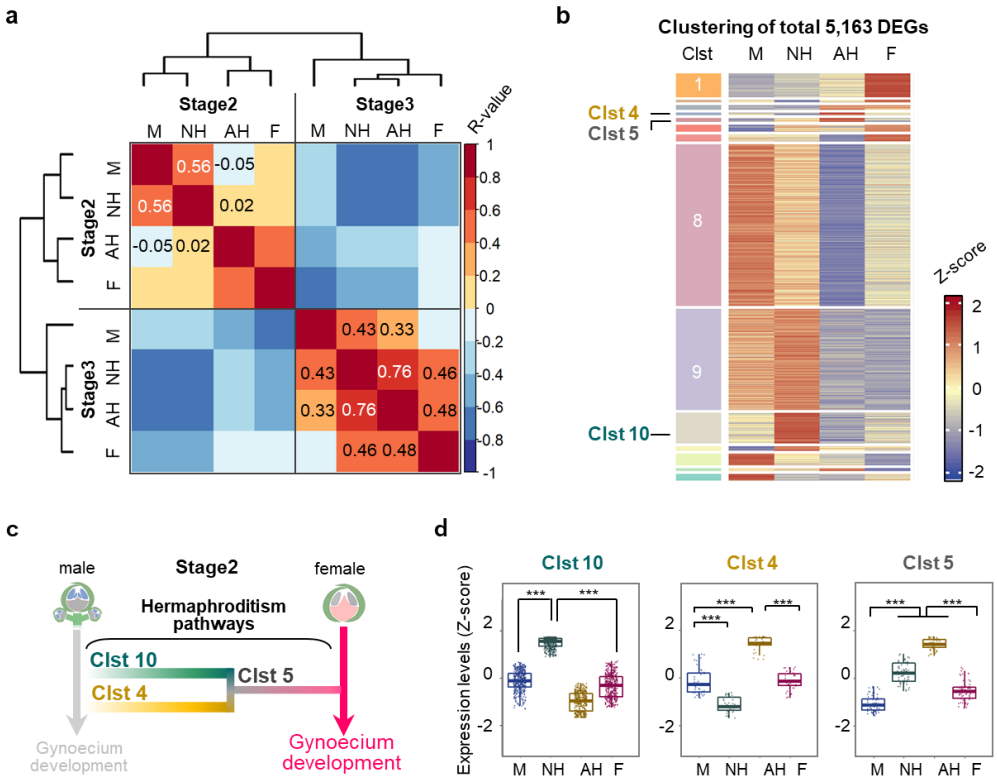

Figure 3-2. Transcriptomic profiles during the development of hermaphrodite flowers from male flowers in hexaploid *D. kaki*.

a, Pearson correlation matrix of average expression levels in the four flower types at stages 2 and 3 (15,019 genes, RPKM > 1). NH and AH flowers were nested into different clusters at stage 2, while a high correlation between gene expression levels in NH and AH flowers was observed in stage 3. **b**, Clustering of total 5,163 DEGs between male and NH/AH flowers at stage 2. **c**, Model of the distinct development pathways underlying NH and AH flowers developments, and leading to gynoecium development. **d**, Three clusters, Clst 10, 4, and 5, were defined as NH-specific up-regulation, AH-specific upregulation, and NH/AH-shared upregulation, respectively.

Co-expression network analysis (Fig. 3-3) with GO enrichment assessment suggested that genes involving drought/oxidative stresses and photosystems were highly enriched in Clst 10 (NH-specific upregulation) (P = 0.07 and 0.05, respectively), and that they can act as the hub genes connecting many genes in the network. These functions are reminiscent of ABA signaling/response, and, consistently, a number of ABA-responsive genes were included in this cluster. On the other hand, Clst 4 (AH-specific upregulation) included mainly cytokinin signal-related genes, and was enriched in genes with cell-cycle and adenine phosphoribosyl transferase GO terms (P = 0.02 and 3.1e-5, respectively), as well as stress-responsive genes. These results strongly suggest that the key signals involved in the development of NH and

AH flowers are distinct from each other. Clst 5 (NH/AH-shared upregulation), which hypothetically integrates the signals from NH- and AH-specific clusters, mainly contained stress-responsive and cell-cycle/cytokinin related genes, with functions that overlap with those found in Clst 10 and 4. Importantly, four transcription factors connected many genes nested to Clst 5. These four transcription factors might therefore play an important role in linking the NH- and AH- specific signals to the path that ultimately leads to hermaphroditism.

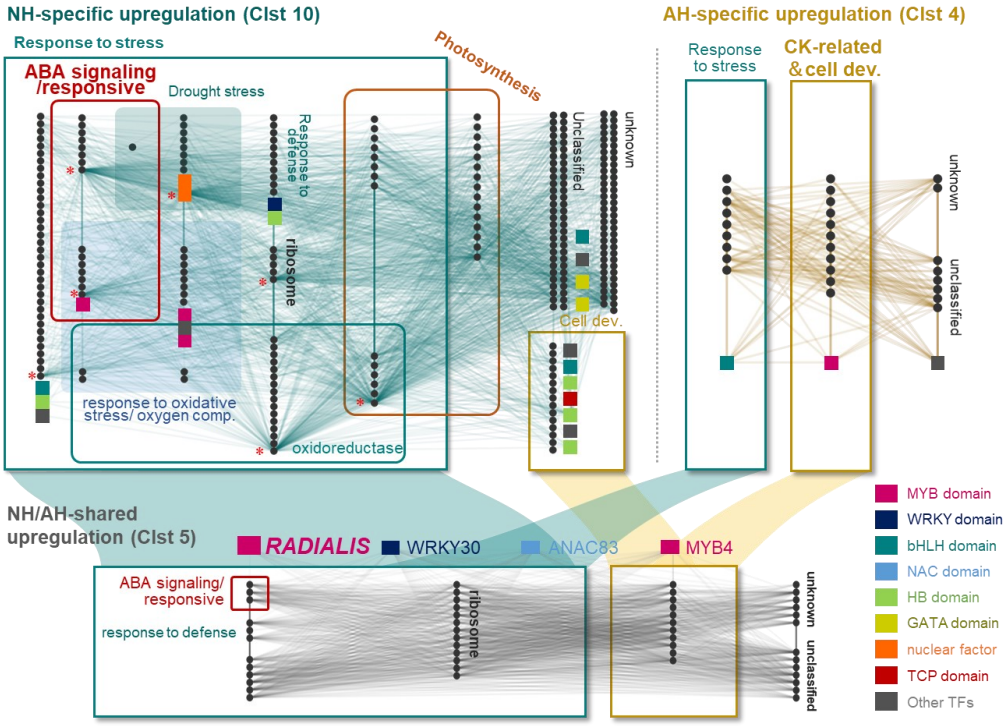

Figure 3-3. Visualization of the gene networks in Clst 10, 4, and 5.
The genes related to response to stress and cell division are represented by teal and gold squares, respectively. The genes related to photosynthesis (orange frame) and ABA signaling/response (red frame) accumulated only in the NH-specific and NH/AH-shared upregulated clusters. Transcription factors are represented by filled squares, with the colors corresponding to each TF domain. The core genes connected to more than 35 genes are indicated with red asterisks.

To verify the involvement of ABA-mediated stress signaling in the reactivation of gynoecium growth in NH flowers, male flowers from hexaploid *D. kaki* were treated with ABA at stage 1.5. In these flowers, ABA application restored gynoecium growth depending on ABA concentration (Fig. 3-4a,b, $P = 0.06$ and 0.03 for 10 and 100 mg/L ABA treatment, respectively,

one-sided paired Student's *t*-test), resulting in production of almost perfect hermaphrodite flowers (Fig. 3-4c). LC-MS analysis in male and NH flowers at stage 2, however, showed no significant differences in concentration of the endogenous plant hormones, including ABA, between male and NH flowers (Fig. 3-4d, P = 0.32 for ABA concentration, one-sided Student's *t*-test), except for trans-Zeatin (tZ, one of the major forms of cytokinin). Trans-Zeatin (tZ) levels were slightly higher in NH flower (P = 0.02), which is consistent with the described observations that NH and AH (CK treatment) flowers exhibited similar expression patterns at stage 3 (Fig. 3-2a). When comparing expression patterns between ABA-treated male and

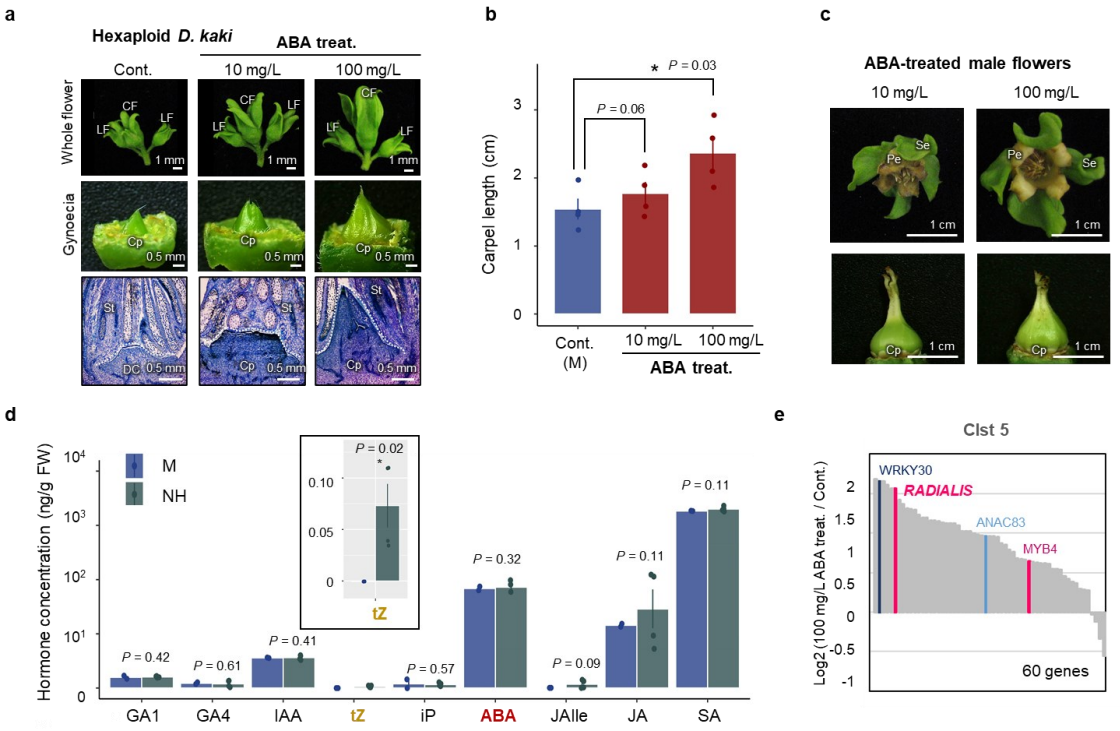

Figure 3-4. ABA treatment induces hermaphroditism in hexaploid *D. kaki*.
a, Treatment of male flowers with 10 or 100 mg/L of ABA. ABA-treated male flowers showed enlarged sepals and flower buds, and restoration of gynoecium development, depending on ABA concentrations. **b**, In comparison to male, flowers treated with higher concentration of ABA produced significantly longer carpels (n = 4 biological replicates). *P*-values were obtained by comparing mean carpel length values using the paired Student's t-test. **c**, Gynoecium development was restored in male flowers treated with 10 and 100 mg/L ABA. Cp, carpel; Pe, petal; Se, sepal.**d**, Endogenous plant hormone (gibberellin A1, gibberellin A4, indole-3-acetic acid, trans-zeatin, N6 - ($\triangle^2$-isopentenyl) adenine, ABA, jasmonoyl isoleucine, jasmonic acid, and salicylic acid) concentrations in M and NH flowers at stage 2. No significant differences were detected for any of analyzed hormones, including ABA (P > 0.09), except for trans-Zeatin (tZ) (P = 0.02, a one-sided t-test). **e**, Almost all genes in Clst 5 were upregulated in male flowers treated with 100 mg/L ABA, in comparison to the control male flowers (56 of 60 genes, bias > 1.2-fold). The four transcription factors identified in Clst 5 based on the presence of a TF domain were highlighted (see Fig. 3-3).

control (male) flowers, most genes included into the NH-specific co-expression Clst 10 (Fig. 3-3) were upregulated in the ABA-treated male. A total of 2,263 DEGs were identified between the ABA-treated male and the control male flowers (FDR < 0.1, bias >1.2-fold, by paired-edgeR), and these DEGs were significantly enriched in GO terms associated with response to wounding, cell division, carpel development, and photosynthesis. This was fully consistent with the gene functions represented in Clst 10. Furthermore, the ABA-treated flowers exhibited significant upregulations of most genes in Clst 5 (NH/AH-shared upregulation) (Fig. 3-4e). Taken together, these results are consistent with the hypothesis that ABA signaling, rather than ABA content, could activate the NH specific pathway and the putative downstream NH/AH-shared network.

Neither ABA nor CK treatments resulted in restoration of gynoecium development (or conversion to hermaphroditism) in male flowers of diploid dioecious *D. lotus*, which is one of the closest wild relatives of hexaploid *D. kaki* (Yonemori et al. 2008; Akagi et al. 2020) (Fig. 3-5a-c). On the other hand, treatment of male flowers of hexaploid *D. kaki* resulted in carpel elongation (Fig. 3-4a-c, 3-5d). Thus, the polyploid-specific ability to restore gynoecium might be related to the acquisition of a factor responsive to both the ABA and CK signaling pathways. To identify this potential factor, we investigated the four transcription factors upregulated in the NH and AH shared cluster (Clst 5, see Fig. 3-3) and ABA-treated male flowers (Fig. 3-4e). Only one of them, a small-Myb *RADIALIS*-like gene (*DkRAD*; Dlo_pri0016F.1_g05470.1 in the *D. lotus* genome) exhibited an expression pattern consistent with this function: it was upregulated in developing hermaphrodite flowers of hexaploid *D. kaki* (NH, AH, and ABA-treatment), while no significant expression could be detected in diploid *D. lotus* developing flowers (Fig. 3-5e). These data identify *DkRAD* as a good candidate.

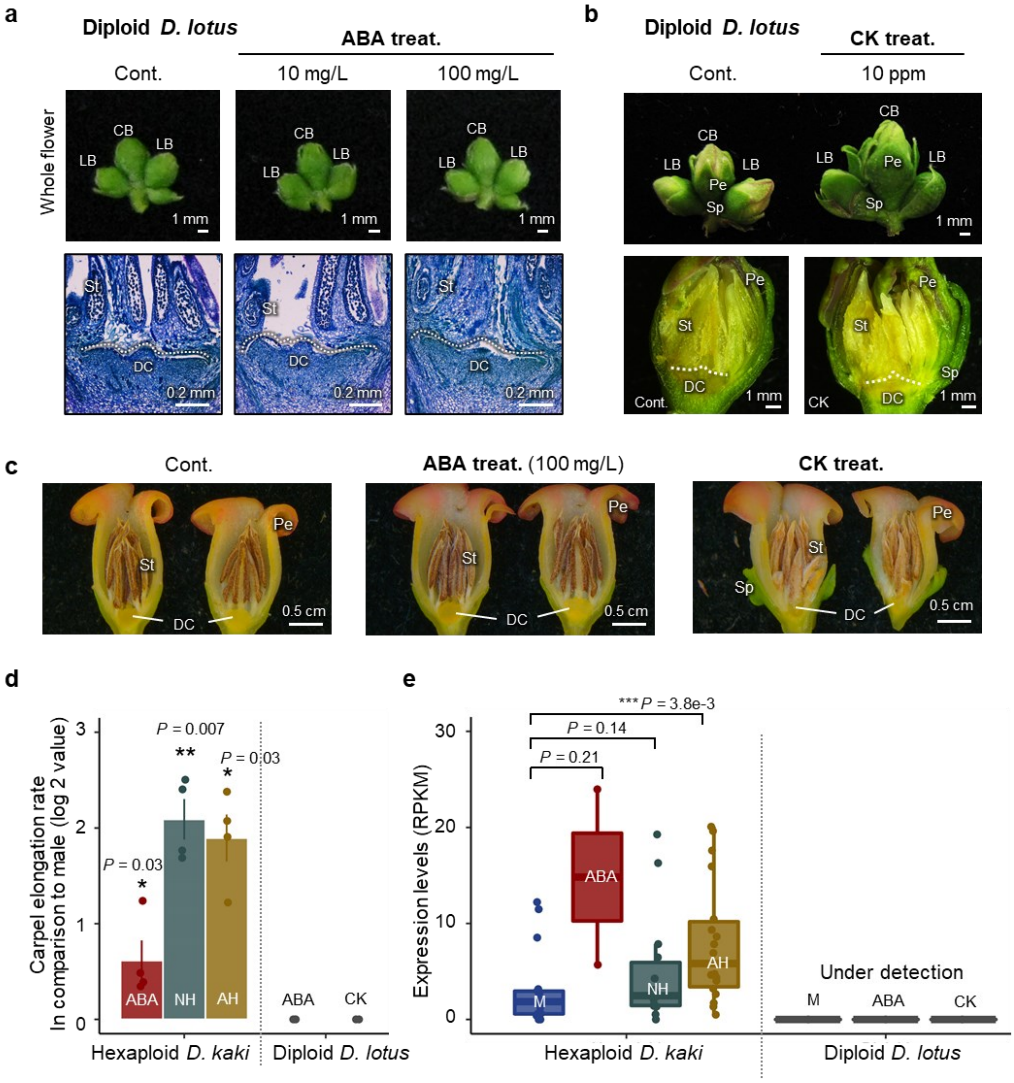

Figure 3-5. No significant differences of synthetic ABA and CK treatment on diploid *D. kaki* male flowers and *DkRAD* expression specifically in polyploid persimmon.
a-c, *D. lotus* male flowers (cv. Kunsenshi-male and its F1 population) treated with 10 or 100 mg/L ABA and 10 ppm CK did not restore gynoecium development in both developing (**a**,**b**) and flowering stage (**c**). CB, central bud; Cp, carpel; DC, defective carpel; LB, lateral bud; Pe; Petal, Sp; Sepal, St, Stamen. **d**, Comparison of the gynoecium development in male flowers of hexaploid *D. kaki* [ABA treated, NH and AH (CK treatments)] and diploid *D. lotus* (ABA and CK treatments). Gynoecium restoration rates were indicated in comparison to the averaged (deficient) gynoecium length in male flowers (n = 4 biological replicates). **e**, Hexaploid-specific up-regulation of *DkRAD* in all three hermaphrodite flowers types (NH, AH, or ABA treatment) in comparison to male flowers. In diploid *D. lotus*, expression of the *DkRAD* ortholog was undetectable even after ABA or CK treatments.

DkRAD is highly homologous to *RADIALIS* from *Antirrhinum majus*, which was originally identified as a key gene involved in floral asymmetry, petal morphogenesis (Corley et al. 2005; Costa et al. 2005; Lucibelli et al. 2020), or floral transition (Hamaguchi et al. 2008). Since assessment of flower traits in transgenic *Diospyros* plants is challenging due to a long juvenile phase (> 6-years), the function of *DkRAD* was assessed in transgenic lines of two distantly related model plants, *Arabidopsis thaliana* and *Nicotiana tabacum*, as previously performed when characterizing the functions of candidate sex determinants in persimmon and kiwifruit (Akagi et al. 2014; Akagi et al. 2018). Ectopic over-expression of *DkRAD* in *A.*

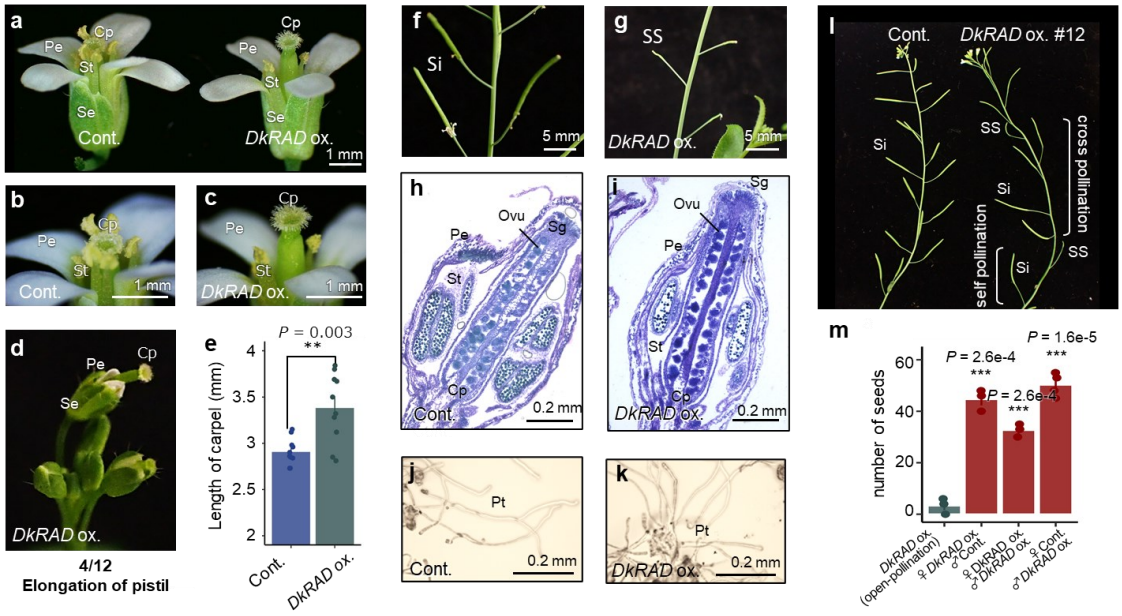

Figure 3-6. Functional validation of *DkRAD* in *Arabidopsis thaliana*.
a-i, Ectopic expression of *DkRAD* (DkRAD ox.) in *A. thaliana* resulted in carpel overgrowth. Flower morphology of the control and *DkRAD ox.* lines (**a-c**). The overgrowing carpel protruded before anthesis (**d**). Comparison of carpel length in the control and *DkRAD ox.* line (n = 10 for biological replicates) (**e**). **f-i**, Siliques were normally developed after self-pollination in the control (**f**), while *DkRAD ox.* lines were self-sterile (**g**). The control and *DkRAD* ox. lines showed no substantial differences in ovule and anther development (**h** and **i** for immediately before the anthesis). Self-sterility of the *DkRAD* ox. lines could be due to the physical distance between stigma and anthers, caused by the hyper-growth of carpel. **j-k**, Pollen tube growth in the control (**j**) and *DkRAD* ox. lines (**k**). Cp, carpel; Pe, petal; Pt, pollen tube; Se, sepal; Sg, stigma; Si, silique; SS, sterile silique; St, stamen. **l**, Cross-pollination with control lines could produce fertile silique, suggesting that female organs are functional in the *DkRAD* transgenic lines. **m**, Comparison of the fertile seed numbers between open-pollination and artificial self-pollination or reciprocal crossing with the control in *DkRAD* ox. #12. All the artificial self-pollination and reciprocal cross-pollination exhibited recovery of the seed fertility, supporting that physical distance between anthers and stigma would be the cause of sterility in the *DkRAD* ox. lines. P-values were detected in comparison to *DkRAD ox.* #12.

thaliana, under the control of the CaMV35S promoter, induced gynoecium overgrowth (Fig. 3-6a-e; *P* = 0.003 in comparison to the controls see Fig. 3-6b, one-sided Student's t-test), resulting in failure of self-fertilization in open pollination (Fig. 3-6f-g), while artificial pollination could produce normal seeds (Fig. 3-6l-m). On the other hand, their androecium (male) functions were not affected (Fig. 3-6h-k). Similar results were observed in transgenic *N. tabacum* with ectopic over-expression of *DkRAD*, and also of the close paralog *DkRAD2* (Dlo_pri0120F.1_g01930.1 in the D. lotus genome). Gynoecium overgrowth in the transgenic *N. tabacum* lines (Fig. 3-7a-c) was visible and possibly associated with increased cell proliferation (Fig. 3-7d-f), while androecium (male) functions were not affected (Fig. 3-7g-i).

These morphological effects of *DkRAD* overexpression in two model plants were inconsistent with the phenotype of overexpression of feminizing factor *MeGI*, which produced sterile androecium (Akagi et al. 2014, 2019, and 2020). These results suggest that *MeGI* and *DkRAD* enable feminization via two different molecular pathways. Taken together, these results suggest that *DkRAD* can act as a feminization factor which might be able to overwrite the existing sex determination systems.

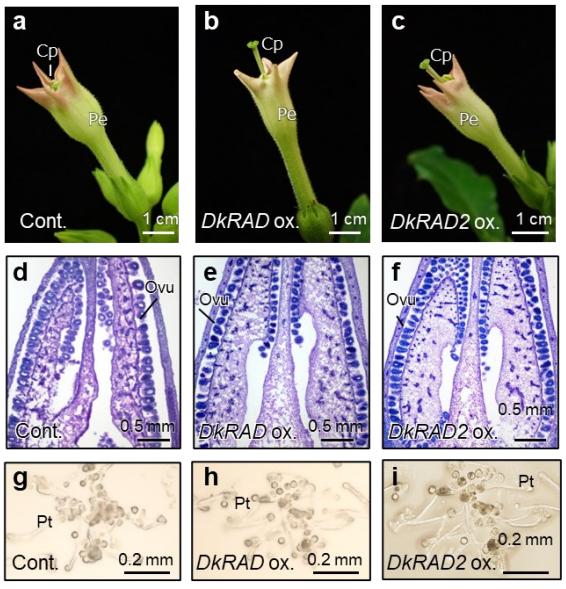

Figure 3-7. Functional validation of *DkRAD* in *Nicotiana tabacum*.
a-c, In *N. tabacum*, ectopic expression of *DkRAD* (*DkRAD ox.*) also resulted in carpel overgrowth. Ectopic over-expression of a close paralog of *DkRAD*, *DkRAD2 ox.*, resulted in a similar phenotype. **d-i**, Consistent with the results in Arabidopsis, ovule development and pollen fertility were not affected in *DkRAD* ox. and *DkRAD2* ox. transgenic lines (**d-f**-w for cross-sections immediately before anthesis, stained with toluidine blue, and **g-i** for pollen germination).

Transcriptome analysis in flowers from *DkRAD*-overexpressing Arabidopsis lines suggested that *DkRAD* could induce the expression of genes associated with auxin- and defense/stress-signals, as well as cell cycle or circadian rhythms (FDR < 0.1, *N* = 517). Specifically, the DEGs involving these pathways were *REVEILLE 1/8* (*RVE1/8*), *LATE ELONGATED HYPOCOTYL* (*LHY*) (Rawat et al. 2009, 2011), *TEOSINTE BRANCHED 1, CYCLOIDEA AND PCF TRANSCRIPTION FACTOR 3* (*TCP3*) (Li et al. 2013), *MYB73* (Yang et al. 2020), or *NGATHA* (*NGA*) (Martínez-Fernández et al. 2014) (Fig. 3-8a). Importantly, *MYB73*, which can regulate auxin and salt stress responses (Yang et al. 2020; Wang et al. 2021) was also upregulated in NH and AH persimmon flowers (Fig. 3-8b,c). Auxin often involves gynoecium growth in *A. thaliana*, partially via activation of *NGA* (Marsch- Martínez et al. 2016). These results suggest that *DkRAD* might initiate or participate in a novel pathway leading to the restoration of gynoecium via auxin signaling, and potentially *MYB73* activation

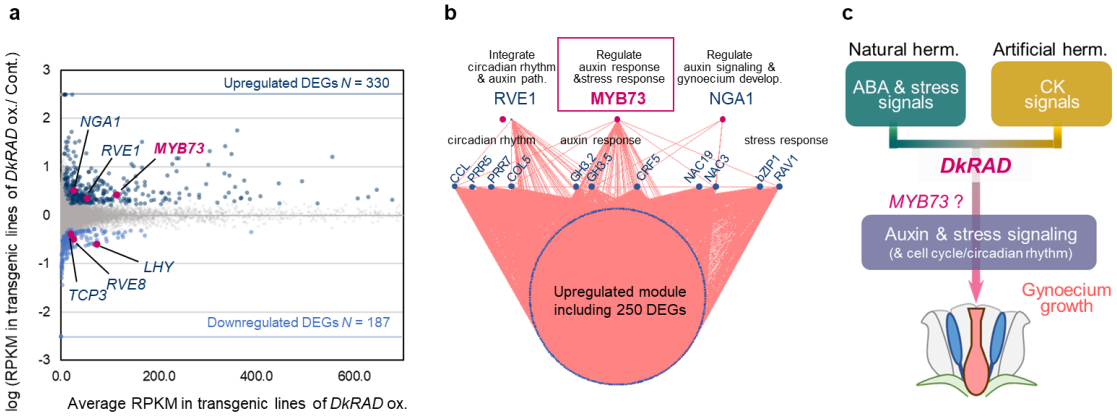

Figure 3-8. Comparison of gene expression patterns in *DkRAD* overexpressing and control *A. thaliana* flower.
a, Detection of differentially expressed genes (DEGs) between the transgenic lines of *DkRAD* ox. and control developing flowers in *A. thaliana* (stage 8-10). Distribution of the expression patterns of the DEGs. The X and Y axes correspond to the normalized expression level (RPKM) and *DkRAD* ox./control flowers ratio, respectively. The 330 upregulated and 187 downregulated DEGs are highlighted in blue circles (FDR < 0.1). The DEGs with known auxin- and defense/stress-signals functions, or those involved in cell cycle or circadian rhythm regulation functions are indicated with pink circles. **b**, Visualization of the co-expression network exhibiting upregulation in *DKRAD* ox. Transcription factors are aligned at the top. **c**, Hypothetical regulatory pathway to produce hermaphrodite flowers from male flowers in persimmon. Activation of *DkRAD* expression, which integrates NH and AH signals, would lead to upregulation of auxin and stress responsive signaling, potentially via *MYB73* expression, resulting in promotion of gynoecium growth.

(Fig. 3-8d). The difference in expression patterns between diploid *D. lotus* and hexaploid *D. kaki* further suggests that *DkRAD* has evolved the ability to develop hermaphrodite flowers through a change in expression pattern in the gynoecium (gain of *cis*-regulation), potentially highlighting the importance of novel *cis*-functions (while the protein and *trans*-regulation remain conserved) for evolution of new sex determination systems, consistent with a Cucumis sex determinant, *WIP1* (Roldan et al. 2020), and the kiwifruit Y-encoded *Shy Girl* factor (Akagi et al. 2018).

3.1.4 Summary

Our results suggest that the formation of hermaphrodite flowers in *D. kaki* would not be associated with a reversion and loss of existing pathways, but with the establishment of a novel pathway via activation of *DkRAD* in developing flowers (or “reinvention” of the hermaphroditism). The establishment of this novel function for *DkRAD* specific to polyploid (or whole-genome duplicated) *D. kaki* may share the concept with the neofunctionalizations in *MeGI* and *OGI* to act as the sex determinants, which were driven by lineage-specific whole/partial genome duplications (Akagi et al. 2020). Although we have no direct evidence for the hypothesis that the recent hexaploidization to derive *D. kaki* resulted in the acquisition of the novel function of *DkRAD*, it has been theoretically suggested that polyploidizations, or genome/gene duplication events, in general, could act both for establishment of, and escape from dioecy (Ashman et al. 2013). This study provides a new molecular clue for the transition

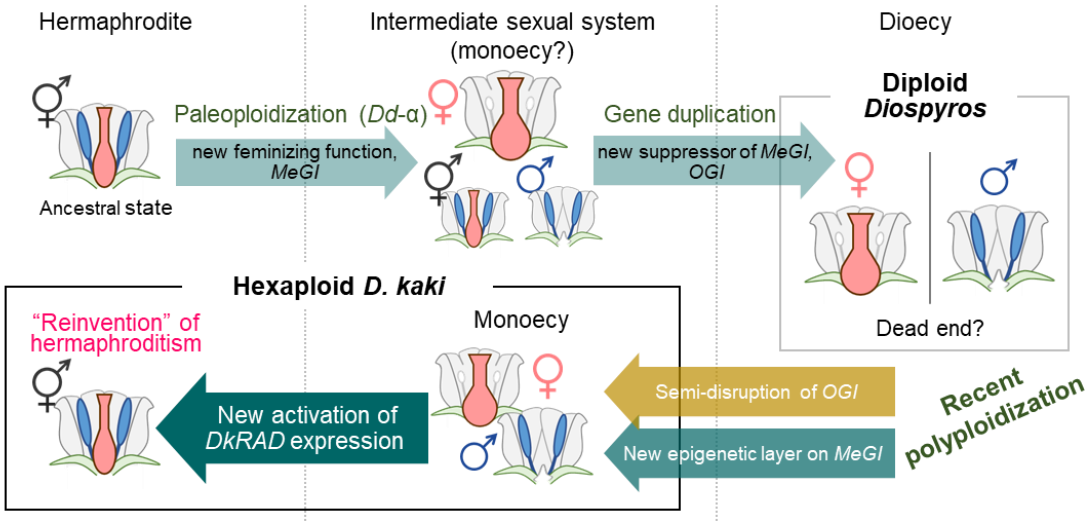

Figure 3-9. Model of the contributions of polyploidization or genome/gene duplication events to the various transitions between sexual systems in the genus *Diospyros*.
Loss-of-function and gain-of-function events are indicated in gold and deep green arrows, respectively. In the evolution of dioecious *Diospyros* species, from putatively ancestral hermaphroditism, transitions into unisexuality (or possibly monoecy), and further into dioecy were triggered by two duplication events: paleo-polyploidization *Dd*-α, and a small segmental duplication, respectively (Akagi et al. 2020). Additionally, hexaploidization in *D. kaki* facilitated silencing of *OGI* via a transposon insertion and the evolution of a layer of epigenetic regulation on *MeGI*, which resulted in establishment of monoecy (Akagi et al. 2016). Further, acquisition of *DkRAD* expression during flower development triggered a full reversion to the production of hermaphrodite flowers.

“out of” dioecy via polyploidization (Fig. 3-9). Plants have undergone frequent lineage-specific polyploidizations (Van de Peer et al. 2021), which might provide good opportunities for novel functions/pathways to trigger establishment of diverse sexual systems.

Conclusion

Polyploidization has occurred repeatedly in plant evolution to play many important roles, including acquisition of trait plasticity adapting to environmental fluctuations. In this study, we focused on the molecular mechanisms of various plastic sex expressions in hexaploid Oriental persimmon (*Diospyros kaki*), which have evolved monoecy and hermaphroditism from dioecy in diploid progenitors. In other words, we aimed to understand the process of sexual transitions to degrade original dioecious system depending on the sex chromosome including the sex determining gene *OGI*. Our results suggested that polyploidization have constructed novel responsive pathways against various environmental conditions, which allowed flexible sex expressions in hexaploid persimmon.

In Chapter 1, we focused on physiological and genetic mechanisms to determine male and female balances in a tree, in monoecious cultivars. cv. Kumemaru, which exhibits unsilenced male flower production, is a good material to unveil the physiological mechanism. We assessed on the *OGI* regulation, and co-expression networks, in comparison of cv. Kumemaru and female cultivars. The *OGI* expression in flower primordia was slightly released in cv. Kumemaru, which was thought to result in stable male flower production. Co-expression network analysis on the differentially expressed genes between cv. Kumemaru and the others, suggested potential involvement of stress-related signaling, particularly in the JA-related pathway, for the stable activation of *OGI* expression in cv. Kumemaru. From a genetic viewpoint, we also conducted genome-wide association analysis (GWAS) with a segregating *D. kaki* population exhibiting substantial biases in male and female flower production. To deal with the autohexasomic inheritance in *D. kaki*, we developed new methods associating or correlating quantitative genotypes with the phenotypes. Importantly, Y-linked regions including *OGI*, exhibited substantial association to male flower production

rate, in allele dosage-dependent manner. We additionally identified 3 genetic regions associated to male flower production rate, including some candidate genes potentially involved in stress signaling. Considering all of the results together, the plasticity in flower sexuality in a tree, is regulated by *cis*-elemental status of the *OGI* gene and potentially by autosomal *trans*-acting factors via stress-signaling pathways.

Genetically female cultivars with no *OGI* allele (or Y-chromosome) in the genome (6A+6X for the karyotype), can very occasionally produce male flowers. In Chapter 2, to characterize the mechanism underlying this “non-canonical” male flower production independent of *OGI*, we focus on a female cultivar, Saijo, which often exhibited reversible somaclonal conversions from female to male (and vice-versa) in a specific area, with application of multi-omics approaches. This somaclonal male conversion was caused by small RNA production in *MeGI* gene resulting in reduction in expression of *MeGI* and its downstream feminizing pathways. Somaclonally converted male flowers specifically showed higher cytosine methylation in the *MeGI* promoter regions than original and revertant female flowers, which was consistent with the canonical male producing system with *OGI* allele. This transition of epigenetic status between male and female were synchronized with genome-wide fluctuation in cytosine methylation levels, especially in the CHG context. These results suggested that flexibility in epigenetic layers can overwrite on genetic sex expression, in hexaploid *D. kaki*.

In addition to the conversion from dioecy to monoecy, hexaploid *D. kaki* further exhibits conversion from male to hermaphrodite flowers, in response to natural environmental signals, or to artificial cytokinin treatment. Importantly, these conversion into hermaphrodite flowers have not been observed in diploid *Diospyros* species. In Chapter 3, to identify the molecular pathways integrating the polyploid-specific production of hermaphrodite flowers,

we detected co-expression gene networks associated with natural hermaphrodite (NH) or cytokinin-treated hermaphrodite (or artificial hermaphrodite: AH) formation with hundreds of transcriptomic data in flower developing stages. Both NH and AH showed no change in expression patterns of the genetic sex determinants, *OGI* and *MeGI*. NH and AH exhibited quite antagonistic expression behaviors, relating to ABA and cytokinin signaling for NH and AH, respectively. Only a small-Myb *RADIALIS*-like gene, named *DkRAD*, was identified as the most likely candidate to integrate these two distinct expression pathways to be hermaphrodite. *DkRAD* expression was consistently upregulated in NH and AH, and also with ABA treatment, in hexaploid *D. kaki*-specific manner. Ectopic over-expression of *DkRAD* in two model plants, *Arabidopsis thaliana* and *Nicotiana tabacum*, resulted in hyper-growth of gynoecium in wild types, and restoration of defected stigma, namely restoration of hermaphroditism from a male mutant. These all results suggested that hexaploid-specific activation of *DkRAD* via multiple environmental signals, is a key to formation of hermaphrodite flowers in *D. kaki*. This also proposed that, in *D. kaki*, the conversion into hermaphroditism, which is the ancestral state of flower sexuality in plants, is not a loss-of-function in the existing sex determination pathway (or reversion into the ancestors), but a reinvention of new pathways.

Our results so far suggested that polyploidization have constructed novel epigenetic layers or signaling pathways responsive to various environmental conditions resulting in flexible sex expressions in hexaploid persimmon, which shed new light on the potential significance of polyploidization in plant sex evolution.

Literature cited

Ainsworth, C. 2000. Boys and girls come out to play: the molecular biology of dioecious plants. *Ann. Bot.* 86, 211-221.

Akagi, T., Henry, I. M., Tao, R., and Comai, L. 2014. A Y-chromosome-encoded small RNA acts as a sex determination in persimmons. *Science*. 346, 646-650.

Akagi, T., Shirasawa, K., Nagasaki, H., Hirakawa, H., Tao, R., Comai, L., et al. 2020. The persimmon genome reveals clues to the evolution of a lineage-specific sex determination system in plants. *PLoS Genet.* 16, e1008566.

Akagi, T., Henry, I., M., Kawai, T., Comai, L., and Tao, R. 2016a. Epigenetic regulation of the sex determination gene *MeGI* in polyploid persimmon. *Plant Cell*. 28, 2905-2915.

Akagi, T., Kawai, T. and Tao, R. 2016b. A male determinant gene in diploid dioecious *Diospyros*, *OGI*, is required for male flower production in monoecious individuals of Oriental persimmon (*D. kaki*). *Sci. Hort.* 213, 243–251.

Akagi, T., Ikegami, A., Tsujimoto, T., Kobayashi, S., Sato, A., Kono, A., and Yonemori, K. 2009. DkMyb4 is a Myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. *Plant Physiol.* 151, 2028-2045.

Akagi, T., Henry, I. M., Ohtani, H., Beppu, K., Kataoka, I., and Tao, R. 2018. A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit. *Plant Cell.* 30, 780-795.

Akagi, T., Pilkington, S. M., Varkonyi-Gasic, E., Henry, I. M., Sugano, S. S., Sonoda, M., et al. 2019. Two Y-chromosome-encoded genes determine sex in kiwifruit. *Nat. Plants.* 5, 801-809.

Akagi, T., Shirasawa, K., Nagasaki, H., Hirakawa, H., Tao, R., Comai, L., et al. 2020. The persimmon genome reveals clues to the evolution of a lineage-specific sex determination system in plants. *PLoS Genet.* 16, e1008566.

Akagi, T., Tao, R., Tsujimoto, T. Kono, A., and Yonemori, K. 2012. Fine genotyping of a highly polymorphic *ASTRINGENCY*-linked locus reveals variable hexasomic inheritance in persimmon (*Diospyros kaki* Thunb.) cultivars. *Tree Genet. Genomes* 8, 195-204.

Akagi, T., Kajita K., Kibe T., Morimura H., Tsujimoto T., Nishiyama S., et al. 2014. Development of molecular markers associated with sexuality in Diospyros lotus L. and their Application in *D. kaki* Thunb. *J. Jpn. Soc. Hortic. Sci.* 83, 214–221

Ashman, T. L., Kwok, A., and Husband, B. C. 2013. Revisiting the dioecy-polyploidy association: alternate pathways and research opportunities. *Cytogenet. Genome Res.* 140, 241-255.

Ballester, P., Navarrete-Gómez, M., Carbonero, P., Oñate-Sánchez, L., and Ferrándiz, C.

2015. Leaf expansion in Arabidopsis is controlled by a TCP-NGA regulatory module likely conserved in distantly related species. *Physiol. Plantarum.* 155, 21-32.

Barrett, S. C. 2002. The evolution of plant sexual diversity. *Nat. Rev. Genet*. 3, 274-284.

Bawa, K. S. 1980. Evolution of dioecy in flowering plants. *Annu. Rev. Ecol. Syst.* 11, 15-39.

Bhaskar, R. V., Mohanty, B., Verma, V., Wijaya, E., and Kumar, P. P. 2015. A hormone-responsive C1-domain-containing protein At5g17960 mediates stress response in *Arabidopsis thaliana*. *PLoS One* 10, e0115418.

Boualem, A., Troadec, C., Camps, C., Lemhemdi, A., Morin, H., Sari, M. A., et al. 2015. A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. *Science* 350, 688-691.

Bräutigam, K., Vining K. J., Lafon-Placette C., Fossdal C. G., Mirouze M., Marcos J. G., et al. 2013. Epigenetic regulation of adaptive responses of forest tree species to the environment. *Ecol. Evol.* 3, 399-415.

Charlesworth, B., and Charlesworth, D. 1978. A model for the evolution of dioecy and gynodioecy. *Am. Nat.* 112, 975-997.

Chat, J., Jáuregui, B., Petit, R. J., and Nadot, S. 2004. Reticulate evolution in kiwifruit (*Actinidia*, Actinidiaceae) identified by comparing their maternal and paternal phylogenies. *Ame. J. Bot*. *91*, 736-747.

Chen, X., Hu, Y., and Zhou, D. 2011. Epigenetic gene regulation by plant Jumonji group of histone demethylase. *Biochem. Bioohys. Acta.* 1809, 421-426.

Cho, H. S., and Pai, H. S. 2000. Cloning and characterization of ntTMK1 gene encoding a TMK1-homologous receptor-like kinase in tobacco. *Mol. Cells.* 10, 317-324.

Comai, L. 2005. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* 6, 836-846.

Corley, S. B., Carpenter, R., Copsey, L., and Coen, E. 2005. Floral asymmetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5068-5073.

Costa, M. M. R., Fox, S., Hanna, A. I., Baxter, C., and Coen, E. 2005. Evolution of regulatory interactions controlling floral asymmetry. *Development* 132, 5093-5101.

Dai, N., Wang, W., Patterson, S. E., and Bleecker, A. B. 2013. The TMK subfamily of receptor-like kinases in Arabidopsis display an essential role in growth and a reduced sensitivity to auxin. *PLoS ONE* 8, e60990.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo M. A. et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158.

Duangjai, S., Wallnöfer, B., Samuel, R., Munzinger, J., Chase, M. W. 2006. Generic delimitation and relationships in Ebenaceae sensu lato: evidence from six plastid

DNA regions. *Am. J. Bot.* 93, 1808-1827

Dufresne, F., Stift, M., Vergilino, R., and Mable, B. K. 2013. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of‐the‐art molecular and statistical tools. *Mol. Ecol.* 23, 40-69.

Du, J., Johnson, L. M., Jacobsen, S. E., and Patel, D. 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nat. Rev. Mol. Cell Biol.* 16, 519–532.

Devoto, A., and Turner, J. G. 2004. Jasmonate-regulated Arabidopsis stress signalling network. *Physiologia Plantarum.* 123, 161-172.

Egert, A., Felix, K., and Shaun, P. 2013. Abiotic stress-induced accumulation of raffinose in Arabidopsis leaves is mediated by a single raffinose synthase (*RS5*, At5g40390). *BMC Plant Bio.* 13, 218.

Endelman, J. B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4, 250-255.

Esumi, T., Watanabe, A., Kosugi, Y., Ohata, K., and Itamura, H. 2015. Staminate flowers on 'Saijo' persimmon (*Diospyros kaki* Thunb.). *Bull Fac Life Env Sci Shimane Univ* 20, 3-8 (In Japanese with English summary)

Ferrão, L. F. V., Benevenuto, J., Oliveira, I. B., Cellon, C., Olmstead, J., Kirst, M., et al. 2018. Insights Into the Genetic Basis of Blueberry Fruit-Related Traits Using Diploid and Polyploid Models in a GWAS Context. *Front. Ecol. Evol.* 6, 1-16.

Fridborg, I., Kuusk, S., Robertson, M., and Sundberg, E. 2001. The Arabidopsis protein SHI represses gibberellin responses in Arabidopsis and barley. *Plant Physiol.* 127, 937-948.

Fresnedo-Ramírez, J., Chan, H. M., Parfitt, D. E., Crisosto, C. H., and Gradziel, T. M. 2017. Genome-wide DNA-(de)methylation is associated with Noninfectious Bud-failure exhibition in Almond (*Prunus dulcis* [Mill.] D.A.Webb). *Sci. Rep.* 7, 1-12.

Fu, X., Richards, D. E., Ait-Ali, T., Hynes, L. W., Ougham, H., Peng, J., et al. 2002. Gibberellin-mediated proteasome-dependent degradation of the barley DELLA protein SLN1 repressor. *Plant Cell* 14, 3191–3200.

Garcia, A. A. F., Mollinari, M., Marconi, T. G., Serang, O. R., Silva, R. R., Vieira M. L. C., et al. 2013. SNP genotyping allows an in-depth characterization of the genome of sugarcane and other complex autopolyploids. *Sci. Rep.* 3, 1-10.

Ghelis, T., Bolbach, G., Clodic, G., Habricot, Y., Miginiac, E., Sotta, B., et al. 2008. Protein tyrosine kinases and protein tyrosine phosphatases are involved in abscisic acid-dependent processes in Arabidopsis seeds and suspension cells. *Plant Physiol.* 148, 1668–1680.

Glick, L., Sabath, N., Ashman, T. L., Goldberg, E., and Mayrose, I. 2016. Polyploidy and

sexual system in angiosperms: Is there an association?. *Am. J. Bot. 103*, 1223-1235.

Goldberg, E. E., Otto, S. P., Vamosi, J. C., Mayrose, I., Sabath, N., Ming, R., et al. 2017. Macroevolutionary synthesis of flowering plant sexual systems. *Evolution*. 71, 898–912.

Grandke, F., Ranganathan, S., Czech, A., Haan, J. R., and Metzler, D. 2014. Bioinformatic tools for polyploid crops. *J. Agri. Sci. Tech.* 593-601.

Grandke, F., Singh, P., Heuven, H. C. M., Haan, J. R., and Metzler, D. 2016. Advantages of continuous genotype values over genotype classes for GWAS in higher polyploids: a comparative study in hexaploidy chrysanthemum. *BMC Genom.* 17, 1-9.

Grant, S., Houben, A., Vyskot, B., Siroky, J., Pan, W. H., Macas, J., et al. 1994. Genetics of sex determination in flowering plants. *Dev. Genet.* 15, 214-230.

Gupta, A., Hisano, H., Hojo, Y., Matsuura, T., Ikeda, Y., Mori, I. C., et al. 2017. Global profiling of phytohormone dynamics during combined drought and pathogen stress in *Arabidopsis thaliana* reveals ABA and JA as major regulators. *Sci. Rep.* 7, 1-13.

Gu, Z., Eils, R., and Schlesner, M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32, 2847-2849.

Hamaguchi, A., Yamashino, T., Koizumi, N., Kiba, T., Kojima, M., Sakakibara, H., et al. 2008. A small subfamily of Arabidopsis *RADIALIS-LIKE SANT/MYB* genes: a link to HOOKLESS1-mediated signal transduction during early morphogenesis. *Biosci. Biotechnol. Biochem.* 72, 2687-2696.

Hariganeya, N., Kikuchi, A., Kamada, H. 2009. SET domain-containing protein genes are involved in *Arabidopsis thaliana* embryogenesis. *Plant Biotech.* 26, 293-300.

Harkess, A., Zhou, J., Xu, C., Bowers, J. E., Van Der Hulst, R., Ayyampalayam, S., et al. 2017. The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nat. Commun.* 8.

Harkess, A. Huang, K., Hulst, R., Tissen, B., Caplan, J. L., Koppula, A., et al. 2020. Sex determination by two Y-linked genes in garden asparagus. *Plant Cell* 32*,* 1790-1796.

Hartwig, T., Chuck, G. S., Fujioka, S., Klempien, A., Weizbauer, R., Potluri, D. P. V., et al. 2011. Brassinosteroid control of sex determination in maize. *Proc. Natl. Acad. Sci. U. S. A.* 108, 19814-19819

Hasegawa, K. 1995. Effects of girdling and strapping of lateral branches on male and female flowering in persimmon cv. Nishimurawase. *Res. Rep. Kochi Univ*. 44, 11-18.

Hasegawa, K., Fukuta, T., Kitahima, A., and Ogata, T. 2004. Effect of permanent and temporary strapping of 2-year old branches on male and female flower bud formation and return bloom in Japanese persimmon. *Res. Rep. Kochi Univ*. 53, 42-52.

Heilbuth, J. C. 2000. Lower species richness in dioecious clades. *Am. Nat.* 156, 221-241.

Heitz, T., Widemann, E., Lugan, R., Miesch, L., Ullmann, P., Désaubry, L., et al. 2011. Cytochromes P450 CYP94C1 and CYP94B3 catalyze two successive oxidation steps of plant hormone jasmonoyl-isoleucine for catabolic turnover. *J. Biol. Chem.* 287, 6296-6306.

Horsch, R. B., Rogers, S. G., and Fraley, R. T. 1985. Transgenic plants. *Cold Spring Harb. Symp. Quant. Biol.* 50, 433-437.

Hu, Y., Jiang, Y., Han, X., Wang, H., Pan, J., and Yu, D. 2017. Jasmonate regulates leaf senescence and tolerance to cold stress: crosstalk with other phytohormones. *J. Exp. Bot.* 68, 1361-1369.

Ito, K., Barnes, P., and Adcock, I. M. 2000. Glucocorticoid Receptor Recruitment of Histone Deacetylase 2 Inhibits Interleukin-1*β*-Induced Histone H4 Acetylation on Lysines 8 and 12. *Mol. Cell. Biol.* 20, 6891-6903.

Käfer, J., Marais, G. A., and Pannell J. R. 2017. On the rarity of dioecy in flowering plants. *Mol. Ecol.* 26, 1225-1241.

Kazama, Y., Ishii, K., Aonuma, W., Ikeda, T., Kawamoto, H., Koizumi, A., et al. 2016. A new physical mapping approach refines the sex-determining gene positions on the *Silene latifolia* Y-chromosome. *Sci. Rep.* 6, 1-10.

Kawakatsu, T., Huang, S. shan C., Jupe, F., Sasaki, E., Schmitz, R. J. J., Urich, M. A. A., et al. 2016. Epigenomic Diversity in a Global Collection of *Arabidopsis thaliana* Accessions. *Cell*. 166, 492–506.

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., et al. 2012. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing, *Genome Res.* 22, 568-576.

Krasovec, M., Chester, M., Ridout, K., and Filatov, D. A. 2018. The mutation rate and the age of the sex chromosomes in *Silene latifolia*. *Curr. Biol.* 28, 1832-1838.

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

Law, J. A., Vashisht, A. A., Wohlschlegel, J. A., and Jacobsen, S. E. 2011. SHH1, a Homeodomain Protein Required for DNA Methylation, As Well As RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA Polymerase IV. *PLoS Genet.* 7, 1-10.

Lee, W. Y., Lee, D., Chung W., and Kwon, C. S. 2009. Arabidopsis ING and Alfin1-like protein families localize to the nucleus and bind to H3K4me3/2 via plant homeodomain fingers. *Plant J.* 58, 511-524.

Lei, J., Jayaprakasha, G. K., Singh, J., Uckoo, R., Borrego, E. J., Finlayson, S., et al. 2019. Circadian clock-associated1 controls resistance to aphids by altering indole

glucosinolate protection. *Plant Physiol.* 181, 1344-1359.

Li, H. and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25, 1754– 1760.

Li, H., Wang, L., Mai, Y., Han, W., Suo, Y., Diao, S., et al. 2021. Phytohormone and integrated mRNA and miRNA transcriptome analyses and differentiation of male between hermaphroditic floral buds of andromonoecious *Diospyros kaki* Thunb. *BMC Genome* 22, 1-19.

Li, S., and Zachgo, S. 2013. TCP3 interacts with R2R3-MYB proteins, promotes flavonoid biosynthesis and negatively regulates the auxin response in *Arabidopsis thaliana*. *Plant J.* 76, 901-913.

Liu, Z., Moore, P. H., Ma, H., Ackerman, C. K., Ragiba, M., Yu, Q., et al. 2004. A primitive Y chromosome in papaya marks incipient sex chromosome evolution. *Nature* 427, 348-352.

Lozano-Duran, R., Rosas-Diaz, T., Gusmaroli, G., Luna, P., Taconnat, L., Deng, X. W., et al. 2011. Geminiviruses subvert ubiquitination by altering CSN-mediated derubylation of SCF E3 ligase complexes and inhibit jasmonate signaling in *Arabidopsis thaliana*. *Plant Cell* 23, 1014-1032.

Lucibelli, F., Valoroso, M. C., and Aceto, S. 2020. Radial or Bilateral? The Molecular Basis of Floral Symmetry. *Genes* 11, 395.

Martínez-Fernández, I., Sanchis, S., Marini, N., Balanza, V., Ballester, P., Navarrete-Gomez, M., et al. 2014. The effect of NGATHA altered activity on auxin signaling pathways within the Arabidopsis gynoecium. *Front. Plant Sci.* 5, 210.

Marsch-Martínez, N., and de Folter, S. 2016. Hormonal control of the development of the gynoecium. *Curr. Opin. Plant Biol.* 29, 104-114.

Massonnet, M., Cochetel, N., Minio, A., Vondras, A. M., Lin, J., Muyle, A., et al. 2020. The genetic basis of sex determination in grapes. *Nat. Commun.* 11, 1-12.

Matzke, M. A., and Mosher, R. A. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15, 394–408.

Matzke, M. A., Kanno, T., and Matzke A. J. 2015. RNA-directed DNA methylation: the evolution of a complex epigenetic pathway in flowering plants. *Annu. Rev. Plant Biol.* 66, 243–267.

McCarthy, D. J., Chen, Y., and Smyth, G. K. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288– 4297.

Mellor, J. 2006. It takes a PHD to read the histone code. *Cell* 126, 22–24.

Miller, J. S., and Venable, D. L. 2000. Polyploidy and the evolution of gender dimorphism

in plants. *Science 289*, 2335-2338.

Miyamoto, T., Uemura, T., Nemoto, K., Daito, M., Nozawa, A., Sawasaki, T., et al. 2019. Tyrosine kinase-dependent defense responses against herbivory in Arabidopsis. *Front. Plant Sci.* 10, 1-9.

Müller, N. A., Kersten, B., Montalvao, A. P. L., Mahler, N., Bernhardsson, C., Brautigam, K., et al. 2020. A single gene underlies the dynamic evolution of poplar sex determination. *Nat. Plants* 6, 630-637.

Murase, K., Shigenobu, S., Fujii, S., Ueda, K., Murata, T., Sakamoto, A., et al. 2016. MYB transcription factor gene involved in sex determination in *Asparagus officinalis. Genes Cells.* 22, 115–123.

Negi, S. S., and Olmo, H. P. 1966. Sex conversion in a male *Vitis vinifera* L. by a kinin. *Science* 152, 1624-1624.

Ong-Abdullah, M., Ordway, J. M., Jiang, N., Ooi, S. E., Kok, S. Y., Sarpan, N., et al. 2015. Loss of *Karma* transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature*. 525, 533–537.

Osborn, T. C., Pires, J. C., Birchler, J. A., Auger, D. L., Chen, Z. J., Lee, H., et al. 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends Genet.* 19, 141-147.

Pannell, J. R., Obbard, D. J., and Buggs, R. J. 2004. Polyploidy and the sexual system: what can we learn from *Mercurialis annua*? *Biol. J. Linn. Soc. 82*, 547-560.

Parker, J. S., and Clark, M. S. 1991. Dosage sex-chromosome systems in plants. *Plant Sci.* 80, 79-92.

Peña, P., Davrazou, F., Shi, X., Walter, K. L., Verkhusha, V. V., Gozani, O., et al. 2006. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* 442, 100–103.

Picq, S. Santoni, S., Lacombe, T., Latreille, M., Weber, A., Ardisson, M., et al. 2014. A small XY chromosomal region explains sex determination in wild dioecious *V. vinifera* and the reversal to hermaphroditism in domesticated grapevines. *BMC Plant Biol.* 14, 1-17.

Qian, W., Miki D., Zhang H., Liu Y., Tang K., Kan Y., et al. 2012. A histone acetyltransferase regulates active DNA demethylation in *Arabidopsis*. *Science* 336, 1445–1448.

Rawat, R., Schwatz, J., Jones, M. A., Sairanen, I., Cheng, Y., Andersson, C. R., et al. 2009. REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16883-16888.

Rawat, R., Takahashi, N., Hsu, P. Y., Jones, M. A., Schwartz, J., Salemi, M. R., et al. 2011. REVEILLE8 and PSEUDO-REPONSE REGULATOR5 form a negative feedback loop

within the Arabidopsis circadian clock. *PLoS Genet.* **7**, e1001350.

Renner S. S. 2014. The relative and absolute frequencies of angiosperm sexual systems: dioecy, monoecy, gynodioecy, and an updated online database. *Am. J. Bot.* 101, 1588-1596.

Renner, S. S., and Ricklefs, R. E. 1995. Dioecy and its correlates in the flowering plants. *Am. J. Bot*. 82, 596-606.

Ritsema, T., Brodmann, D., Diks, S. H., Bos, C. L., Nagaraj, V., Pieterse, C. M. J., et al. 2009. Are small GTPases signal hubs in sugar-mediated induction of fructan biosynthesis? *PLoS ONE* 4, e6605.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 26, 139– 140.

Roldan, M. V. G., Izhaq, F., Verdenaud, M., Eleblu, J., Haraghi, A., Sommard, V., et al. 2020. Integrative genome-wide analysis reveals the role of WIP proteins in inhibition of growth and development. *Commun. Biol.* 3, 1-12.

Rosyara, U. M., Jong, W. S., Douches, D. S., and Endelman, J. V. 2016. Software for genome-wide association studies in autopolyploids and its application to potato, *Plant Genome* 9, 1-10.

Salinas-Mondragón, R. E., Garcidueñas-Piña, C., and Guzmán, P. 1999. Early elicitor induction in members of a novel multigene family coding for highly related RING-H2 proteins in *Arabidopsis thaliana*. *Plant Mol. Biol*. 40, 579–590.

Sangwan, V., Foulds, I., Singh, J., and Dhindsa, R. S. 2001. Cold-activation of *Brassica napus BN115* promoter is mediated by structural changes in membranes and cytoskeleton, and requires Ca2+ influx. *Plant J.* 27, 1–12.

Schommer, C., Palatnik, J. F., Aggarwal, P., Chételat, A., Cubas, P., Farmer, E. E., et al. 2008. Control of jasmonate biosynthesis and senescence by miR319 targets. *PLoS Biol.* 6, e230.

Serrano, M., and Guzmán, P. 2004. Isolation and gene expression analysis of *Arabidopsis thaliana* mutants with constitutive expression of *ATL2*, an early elicitor-response RING-H2 zinc-finger gene. *Genetics* 167, 919-929.

Shafiq, S., Berr, A., and Shen, W. 2013. Combinatorial functions of diverse histone methylations in *Arabidopsis thaliana* flowering time regulation. *New Phytol.* 201, 312-322.

Shaked, H., Avivi-Ragolsky, N., and Levy, A. A. 2006. Involvement of the *Arabidopsis SWI2/SNF2* Chromatin Remodeling Gene Family in DNA Damage Response and Recombination. *Genetics* 173, 985-994.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504.

Shimada, T. L., Shimada, T., and Hara-Nishimura, I. 2010. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *Plant J.* 61, 519-528.

Shirasawa, K., Hirakawa, H., and Isobe, S. 2016. Analytical workflow of double-digest restriction site-associated DNA sequencing based on empirical and *in silico* optimization in tomato, *DNA Res.* 23, 145-153.

Smyth, D. R., Bowman, J. L., and Meyerowitz, E. M. 1990. Early flower development in Arabidopsis. *Plant Cell* 2, 755-767.

Sun, P., Li, J., Du, G., Han, W., Fu, J., Diao, S., et al. 2017. Endogenous phytohormone profiles in male and female floral buds of the persimmons (*Diospyros kaki* Thunb.) during development. *Sci. Hortic.* 218, 213-221.

Suo, Y., Sun, P., Cheng, H., Han, W., Diao, S., Li, H., et al. 2020. A high-quality chromosomal genome assembly of *Diospyros oleifera* Cheng. *GigaScience* 9, giz164

Torres, M. F., Mohamoud, Y. A., Younuskunju, S., Suhre, K., and Malek, J. A. 2021. Evidence of recombination suppression blocks on the Y chromosome of date palm (*Phoenix dactylifera*). *Front. Plant Sci.* 12, 697.

Tsugama, D., Matsuyama, K., Ide, M. Hayashi, M., Fujino, K., and Masuda, K. 2017. A putative *MYB35* ortholog is a candidate for the sex- determining genes in *Asparagus officinalis*. *Sci. Rep.* 7, 41497

Urich, M. A., Nery J. R., Lister R., Schmitz R. J., and Ecker J. R. 2015. MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. *Nat. Protoc.* 10, 475–483.

VanBuren, R., Wai, C. M., Zhang, J., Han, J., Arro, J., Lin, Z., et al. 2016. Extremely low nucleotide diversity in the X-linked region of papaya caused by a strong selective sweep. *Genome Biol.* 17, 1-11.

VanBuren, R., Zeng, F., Chen, C., Zhang, J., Wai, C. M., Han, J., et al. 2015. Origin and domestication of papaya Y^h chromosome. *Genome Res. 25*, 524-533.

Van de Peer, Y., Mizrachi, E., and Marchal, K. 2017. The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18, 411-424.

Volz, S. M., and Renner, S. S. 2008. Hybridization, polyploidy, and evolutionary transitions between monoecy and dioecy in Bryonia (Cucurbitaceae). *Am. J. Bot.* 95, 1297-1306.

Wang B., Yang X., Wang Y., Xie Y., and Zhou X. 2018. Tomato yellow leaf curl virus V2 interacts with host histone deacetylase 6 to suppress methylation-mediated

transcriptional gene silencing in plants. *J. Virol.* 92, e00036-18.

Wang, J., Na. J., Yu, Q., Gschwend, A. R., Han, J., Zeng, F., et al. 2012. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proc, Natl, Acad. Sci. U.S.A.* 109, 13710-13715.

Wang, L., Qiu, T., Yue, J., Guo, N., He, Y., Han, X., et al. 2000. Arabidopsis ADF1 Regulated by MYB73 is Involved in Response to Salt Stress via Affecting Actin Filaments Organization. *Plant Cell Physiol.* 62, 1387-1395.

Wendel, J. F. 2000. Genome evolution in polyploids. *Plant Mol. Evol.* 225-249.

Westergaard, M. 1958. The mechanism of sex determination in dioecious flowering plants. *Adv. Genet.* 9, 217-281.

Wickham, H. 2009. Ggplot2: elegant graphics for data analysis. *Springer‐Verlag, New York.* ISBN 978-3-319‐24277‐4.

Xu, Q., Fu, H. H., Gupta, R., and Luan, S. 1998. Molecular characterization of a tyrosine-specific protein phosphatase encoded by a stress-responsive gene in Arabidopsis. *Plant Cell* 10, 849-857.

Xu, T., Dai, N., Chen, J., Nagawa, S., Cao, M., Li, H., et al. 2014. Cell surface ABP1-TMK auxin-sensing complex activates ROP GTPase signaling. *Science* 28, 1025-1028.

Yakushiji, H., and Nakatsuka, A. 2007. Recent persimmon research in Japan. *Jpn. J. Plant Sci.* 1, 42–62.

Yang H., Akagi T., Kawakatsu T., and Tao R. 2019. Gene networks orchestrated by *MeGI*: a single-factor mechanism underlying sex determination in persimmon. *Plant J.* 98, 97–111.

Yang, Y., Zhang, L., Chen, P., Liang, T., Li, X., and Liu, H. 2020. UV-B photoreceptor UVR8 interacts with MYB73/MYB77 to regulate auxin responses and lateral root development. *EMBO J.* 39, e101928.

Yonemori, K., Sugiura, A., Tanaka, K., and Kameda, K. 1993. Floral ontogeny and sex determination in monoecious-type persimmons, *J. Amer. Soc. Hort. Sci.*, 118, 293-297.

Yonemori, K., Honsho, C., Kanzaki, S., Ino, H., Ikegami, A., Kitajima, A., et al. 2008. Sequence analyses of the ITS regions and the matK gene for determining phylogenetic relationships of *Diospyros kaki* (persimmon) with other wild *Diospyros* (Ebenaceae) species. *Tree Genet. Genomes* 4, 149-158.

Yonemori K., Yomo Y., and Sugiura, A. 1990. Sexuality in Japanese persimmons. 2. Induction of sex conversion in male flower by cytokinin treatment. *J. Japan Soc. Hort. Sci.* 59, 230-231.

Zhang H., Lang Z., and Zhu J. 2018. Dynamics and function of DNA methylation in plants.

Nat. Rev. Mol. Cell. Biol. 19, 489–506.

Zhou, C., Zhang, L., Duan, J., Miki, B., and Wu, K. 2005. *HISTONE DEACETYLASE19* Is Involved in Jasmonic Acid and Ethylene Signaling of Pathogen Response in Arabidopsis. *Plant Cell* 17, 1196-1204.

Acknowledgements

I wish to express my sincere appreciation to Dr. Takashi Akagi, Research Professor of Okayama University, for his valuable comments and critical reading, and Dr. Yasutaka Kubo, Professor of Okayama University, Dr. Koichiro Ushijima, Associate Professor of Okayama University, Dr. Naoko Fujita, Project Assistant Professor of Okayama University for their valuable discussion and encouragements.

I am deeply thank to Dr. Isabelle M Henry, Project Scientist of University of California, Davis, Dr. Luca Comai, Professor of University of California, Davis, Dr. Ryutaro Tao, Professor of Kyoto University, Dr. Hisayo Yamane, Associate Professor of Kyoto University, Dr. Yoko Ikeda, Associate Professor of Okayama University Institute of Plant Science and Resources, Matsuura Takakazu, Technical Staff of Okayama University Institute of Plant Science and Resources, Dr. Taiji Kawakatsu at National Agriculture and Food Research Organization, Dr. Eiji Yamamoto, Specifically Appointed Lecturer of Meiji University, Dr. Kenta Shirasawa at Kazusa DNA Research Institute, Dr. Toshihiro Saito, Dr. Atsushi Kono, Dr. Noriyuki Onoeu, Ryusuke Matsuzaki, at Institute of Fruit Tree Science, National Agriculture and Food Research Organization, and Dr. Tomoya Esumi of Shimane University, Dr. Seiichi Uchida, Professor of Kyushu University, Dr. Takeshi Niikawa, Dr. Tetsuya Suzuki, Mayu Sugiura at Gifu Prefectural Agricultural Technology Center, for providing insightful comments and continuous encouragements for this study.

Also, I wish to express my thanks to Yang Ho-Weng, a member of Lab. Pomology (2017-2018 in Kyoto University) and Maria Suzuki, Eriko Kuwada, Ayano Horiuchi, Natsumi Enoki, Rika Uchida, members of Lab. Postharvest (2019-2022 in Okayama University), for helping my experiments and insightful discussions.

Finally, I sincerely thank to my respectful parents, Masaru Masuda and Hideko Masuda, my brothers, Satoru Masuda and Toru Masuda, for their every support.

Kanae Masuda