Population genetic studies on the evolutionary history of Japanese apricot (Prunus mume)

Numaguchi, Koji

(Degree)
博士（農学）

(Date of Degree)
2020-09-25

(Date of Publication)
2022-09-25

(Resource Type)
doctoral thesis

(Report Number)
甲第7887号

(URL)
https://hdl.handle.net/20.500.14094/D1007887

Doctoral Dissertation

Population genetic studies on the evolutionary history of Japanese apricot (*Prunus mume*)

ウメの進化に関する集団遺伝学的研究

Koji Numaguchi

沼口 孝司

Graduate School of Agricultural Science
Kobe University
July, 2020

Contents

Chapter 1. General introduction

1.1 Domestication and cultivar differentiation in fruit tree crops 1

1.2 Origin and history of use in Japanese apricot 2

1.3 Empirical classification of *P. mume* cultivars 3

1.4 Molecular phylogenetic analysis of *P. mume* cultivars 4

1.5 Current situation of genetic vulnerability and importance of diversity in *P. mume* 5

1.6 The goal of the present study 7

Chapter 2. Microsatellite marker development and population structure analysis in Japanese apricot (*P. mume*)

2.1 Introduction 8

2.2 Materials and methods 9

2.3 Results 12

2.4 Discussion 19

2.5 Summary 21

Chapter 3. Estimation of demographic history of Japanese and Taiwanese populations in Japanese apricot based on microsatellite marker genotypes

3.1 Introduction 22

3.2 Materials and methods 22

3.3 Results and discussion 24

3.4 Summary 27

Chapter 4. Population structure analyses for the East Asian cultivars of Japanese apricot based on exon capture resequencing

4.1 Introduction 28
4.2 Materials and methods 28
4.3 Results and discussion 36
4.4 Summary 43

Chapter 5. Genomic signatures for natural selection and interspecific introgression in East Asian cultivars of Japanese apricot

5.1 Introduction 44
5.2 Detection of selective sweeps potentially involved in the establishment or population differentiation of Japanese apricot
5.2.1 Materials and methods 45
5.2.2 Results and discussion 47
5.3 Estimation of interspecific fragment introgressions between Japanese apricot and related *Prunus* species
5.3.1 Materials and methods 51
5.3.2 Results and discussion 52
5.4 Summary 58

Chapter 6. General discussion

6.1 Genetic diversity in East Asian cultivars of Japanese apricot 59
6.2 Roles of interspecific introgression and natural selection in Japanese apricot 60
6.3 Inferences for the evolution of Japanese apricot 62

Conclusion **64**

References **66**

Appendix

Appendix 1 Supplementary figures 79
Appendix 2 Supplementary tables 84

Acknowledgements **106**

Chapter 1. General introduction

1.1 Domestication and cultivar differentiation in fruit tree crops

Crop domestication is a process to generate new cultivated species by human beings (Doebley et al., 2006). In major annual crops, such as rice, wheat and maize, domestication started about 10,000–12,000 years ago, and the life of human beings gradually shifted from hunter-gatherer to agrarian societies (Purugganan and Fuller, 2009). Compared with the wild forms, these crop species have accumulated many phenotypic changes desirable for human beings. For example, rice has traits of closed panicle (Ishii et al., 2013), elect plant growth (Jin et al., 2008) and non-seed shattering (Konishi et al., 2006; Li et al., 2006; Htun et al., 2014). Wheat keeps non-dehiscent spike (Watanabe et al., 2002) and free-threshing grain (Dvorak et al., 2012), and maize has softer glume (Wang et al., 2005) and reduced axillary branching (Doebley et al., 1997). Genetic factors responsible for these traits have been identified, and they also have shown to go through natural (or artificial) selections (Stitzer and Ross-Ibarra, 2018; Chen et al., 2019; Haas et al., 2019).

Perennial fruit tree crops are also important worldwide. They include apple, grapes, citrus, peach, olive and nuts, and their total cultivation area is equivalent to about one-eighth of the world food production area (McClure et al., 2014). Compared to the annual crops, much less facts have been elucidated on fruit tree domestication and improvement (Miller and Gross, 2011). This may be due to the fact of ecological difference between annuals and perennials (Gaut et al., 2015). Fig. 1 shows a model of domestication and differentiation of tree crops by Gaut et al. (2015). Most perennial (tree) crops have outcrossing behavior, and cultivars are usually maintained through vegetative propagation, rather than seed propagation (Miller and Gross, 2011). Therefore, most of the genetic loci are maintained in heterozygous forms (Gaut et al., 2015; Akagi et al., 2016). As these genotypes are maintained for many years, somatic mutations may accumulate in perennial plants. Then, outcross may take place among the plants. These cause the difficulty to reveal the genetic structure. Additionally, long-lived perennials have extended juvenile phase, and number of generations after domestication is much smaller than that of annuals (Gaut et al., 2015). Thus, analytical methods developed from the study of annual crops cannot be applied easily for perennial crops. However, recently, researchers have noticed that the characteristics of perennial crops are similar to those of human and other animals, and they could identify candidate genes for domestication and

improvement through scanning genomic signatures of natural (or artificial) selection in perennial fruit tree crops, such as peach (Akagi et al., 2016). Therefore, studies on fruit tree domestication and cultivar differentiation history are expected to be accelerated by applying the recent advanced methodologies for genome-wide analysis.

In this study, the author investigated Japanese apricot, one of the popular fruit tree crops in the East Asia (especially in Japan), and tried to clarify the pathway for domestication and cultivar differentiation through a series of population genetic surveys.

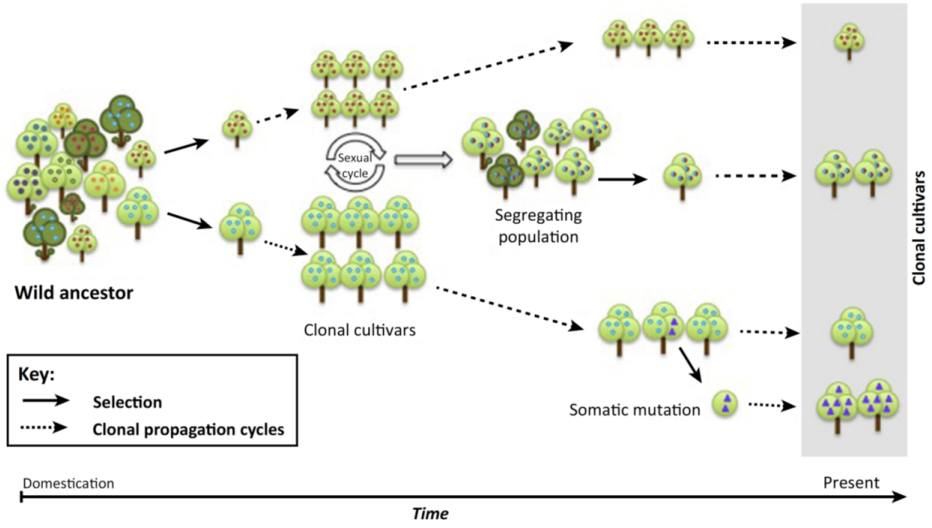

Fig. 1. A model of domestication and differentiation in tree crops by Gaut et al. (2015).

1.2 Origin and history of use in Japanese apricot

Japanese apricot (*Prunus mume* Sieb. et Zucc.), which is also called Chinese mei or mume, is a kind of deciduous fruit tree species. It is popular in the East Asia (China, Japan and Taiwan). It has long been used for ornamental or medicinal purposes although fresh fruit is not edible due to the strong acidity. Now, its processed fruits are familiar with Japanese people, such as liquor, syrup and pickles called “umeboshi”. The origin of *P. mume* is considered to be mountainous areas of Yunnan, Sichuan and Tibet in China (Mega et al., 1988; Horiuchi et al., 1996; Chen, 2017). China also have origins of other major stone fruit species such as apricot and peach, and is thought to be their diversity center (Faust et al., 2011). *P. mume* in China is distributed throughout the south side of the Yangtze River and is thought to have been cultivated from at least 3,000 years ago, as

fruit remains (stones) were excavated from ruins in Yin Period (1400–1200 BC) (Horiuchi et al., 1996; Chen, 2017). In China, many varieties of Japanese apricot have been cultivated mainly for ornamental purposes and thus researchers have been focusing on ornamental traits (Zhang et al., 2015, 2018). In Taiwan, which have geographical proximity to south part of China, cultivars generally have very weak chilling requirement for endodormancy release to adapt subtropical climates (Horiuchi et al., 1996; Yamane, 2014). The use of *P. mume* in Japan started at least 2,000 years ago, because the oldest remains were discovered from ruins in Yayoi Period (1000–300 BC) (Horiuchi et al., 1996). As *P. mume* remains are not found before the Yayoi Period, most researchers consider that *P. mume* cultivars in Japan were introduced from China by people (Mega et al., 1988; Horiuchi et al., 1996). However, there is a hypothesis that wild *P. mume* populations were originally distributed in south part of Japan (and Taiwan) long time ago (Yoshida, 1984). In fact, some wild *P. mume* samples were collected in Miyazaki Prefecture in Japan (https://agriknowledge.affrc.go.jp/RN/3030041889). These wild individuals might derive from the feral cultivars, however, it cannot be concluded that all the Japanese cultivars were introduced from China through human activity (Mega et al., 1988; Horiuchi et al., 1996). Thus, the origin of Japanese cultivars is still not clear.

1.3 Empirical classification of *P. mume* cultivars

Conventionally, *P. mume* is classified into ornamental and fruit cultivars (Mega et al., 1988; Horiuchi et al., 1996). Ornamental group consists of more cultivars than fruit group because *P. mume* was originally used for ornamental purposes (Mega et al., 1988; Horiuchi et al., 1996; Chen, 2017). Therefore, *P. mume* may have relatively short history as “fruit tree”. In Japan, for example, only 50 cultivars were confirmed in 1936, and modern breeding was started after that (Horiuchi et al., 1996). This indicates that *P. mume* fruit may have gradually been utilized with the development of fruit processing techniques (Yoshida, 1984). Probably, Japanese fruit cultivars may be derived from the ornamental cultivars. Further, the fruit cultivars are divided into small, middle and large fruit size groups (Yoshida and Yamanishi, 1988; Horiuchi et al., 1996; Yaegaki et al., 2003). Small fruit group has very small-sized fruits with less than 10 g and the harvesting time is very early. Most of the small fruit cultivars have self-compatible crossing behavior. Middle and large fruit cultivars are difficult to be distinguished because of continuous distribution of the fruit size. In addition, there are putative interspecific hybrids between *P. mume* and *P. armeniaca* (apricot) and between *P. mume* and *P. salicina* (Japanese plum), called “Anzu-ume” and “Sumomo-ume”, respectively. They have extremely large-sized

fruits (more than 50 g) and the harvesting time is very late. Since their leaf and stone morphologies are similar to those of *P. armeniaca* and *P. salicina*, they are easily to be distinguished from the other *P. mume* cultivars (Yoshida and Yamanishi, 1988; Tzonev and Yamaguchi, 1999).

Generally, flowering and leafing periods of *P. mume* are earlier than those of *P. armeniaca* and *P. salicina* (Hijikata, 1984). Among *P. mume* cultivars, Taiwanese cultivars shows very early flowering and leafing in Japan (Horiuchi et al., 1996; Yamane, 2014). Since most of *P. mume* cultivars show best performance on flowering and leafing in temperate climate, they may have been adjusted to the climate conditions. In cold climate, early flowering has disadvantage on low activity of pollinators and frost damage of fruits. In Japan, late flowering (leafing) large fruit (e.g., 'Shirokaga') or anzu-ume (e.g., 'Bungo') cultivars were selected and cultivated in the cold area (Yoshida and Yamanishi, 1988; Tzonev and Yamaguchi, 1999). This suggests that introgressions from related *Prunus* species may have taken important roles for the adaptation of *P. mume* to cold climates.

1.4 Molecular phylogenetic analysis of *P. mume* cultivars

P. mume belongs to the same subgenus *Prunus* as closely related two species, *P. armeniaca* (apricot) and *P. salicina* (Japanese plum) (Yoshida, 1984; Bortiri et al., 2001). Since they are partially cross-compatible to each other (Yamaguchi et al., 2018; Morimoto et al., 2019), interspecific hybrids can be produced if they flower at the same period. There are some *P. mume* cultivar groups having characteristics of *P. armeniaca* or *P. salicina* (Yoshida, 1984). Therefore, subgenus *Prunus* species may have evolved through interspecific crosses. Phylogenetic classification in *P. mume* was first attempted using isozyme analysis. Hijikata (1984) described that specific markers for *P. mume*, *P. armeniaca* and *P. salicina* were respectively identified in isozyme analyses and they could detect interspecific hybrids. Then, 'Bungo', 'Fushida', 'Seiyobai', 'Shirokaga' and 'Gyokuei' were found to have *P. armeniaca* specific band, and 'Sumomoume' and 'PM1-1' were shown to have *P. salicina* specific band (Hijikata, 1984). After the invention of PCR, RAPD (Random Amplified polymorphic DNA) analysis was applied for the *P. mume* cultivars. Shimada et al. (1994) confirmed that putative anzu-ume hybrids such as 'Bungo' and 'Takadaume' are descendants of *P. armeniaca*, and found that Taiwanese cultivars were classified into different clade from Japanese cultivars. Based on the chloroplast markers in *trnL-trnF* region, maternal parents of the anzu-ume cultivars, such as 'Seiyobai', 'Bungo' and 'Taihei', were revealed to be *P. armeniaca* (Ohta et al., 2006).

Afterwards, Hayashi et al. (2008) carried out a phylogenetic analysis on 127 *P. mume* cultivars using 14 microsatellite markers designed from the *P. persica* (peach) and *P. armeniaca* genome sequences. They could confirm the previous results, however, population structure among Japanese cultivars were not well investigated. Using REMAP (retrotransposon-microsatellite amplified polymorphism) and IRAP (inter-retrotransposon amplified polymorphism) markers, another group reported that ornamental and fruit cultivars were classified into distinct clades using 84 Chinese and Japanese *P. mume* cultivars (Yuying et al., 2011). Since the differentiation of the fruit cultivars is quite recent (Yoshida, 1984), more precise analysis is expected to reveal population structure of *P. mume*.

1.5 Current situation of genetic vulnerability and importance of diversity in *P. mume*

Currently, *P. mume* is mainly cultivated in Wakayama Prefecture. The cultivation area corresponds to more than 60% of total production in Japan (MAFF, 2019), and the most prominent variety of 'Nanko' occupies more than 80% of the total production area in Wakayama Prefecture (MAFF, 2017). However, the production depending largely on a single cultivar is facing vulnerability against new pests.

Since the 1980s, graft-transmissible symptoms have been reported on *P. mume* 'Nanko' in Wakayama Prefecture. These include incomplete flower development (small flowers, sometimes lacking the stigma), interveinal chlorosis (the precursor of leaf-edge necrosis, another symptom), and early defoliation (called the "chagasu" syndrome by local farmers, which means "tea grounds" in Japanese) (Otsubo et al., 1991; Iemura et al., 1995) (Fig. 2). These symptoms are highly problematic because fruit yield is also markedly reduced in 'Nanko' trees having this disease (Kansako et al., 2000). Previously, double infection with cucumber mosaic virus (CMV) and prunus necrotic ringspot virus (PNRSV) was shown to be the cause of these symptoms, and this disease was named leaf-edge necrosis (Ohtubo et al., 2002; Kurihara et al., 2006). However, CMV and PNRSV were not detected by molecular biological methods in 'Nanko' trees cultivated in the midwestern part of Wakayama Prefecture (Nakaune et al., 2018). Instead (and for the first time), plum bark necrosis stem pitting-associated virus (PBNSPaV) and little cherry virus 2 (LChV-2), both of which belong to the genus *Ampelovirus*, were identified. Therefore, a large-scale survey of the two viruses were previously conducted by RT-PCR detection in midwestern area (Numaguchi et al., 2019). Among 208 surveyed 'Nanko' trees, two viral infections were significantly associated with incomplete flower development, low fruit bearing rate, and interveinal chlorosis. And LChV-2 infection was affected Nanko

fruit and stone sizes. Based on the results, two viruses were found to be widely distributed throughout the main cultivation area in Wakayama Prefecture (Fig. 3).

The major variety, 'Nanko' is largely affected by a viral disease, mume leaf-edge necrosis. This strongly indicates that the crop production should not rely on the limited cultivars. To overcome these problems, we need to utilize the genetic resources having various characteristics. Therefore, it is important to understand the current population structure and genetic diversity of the broadly collected resources of *P. mume*.

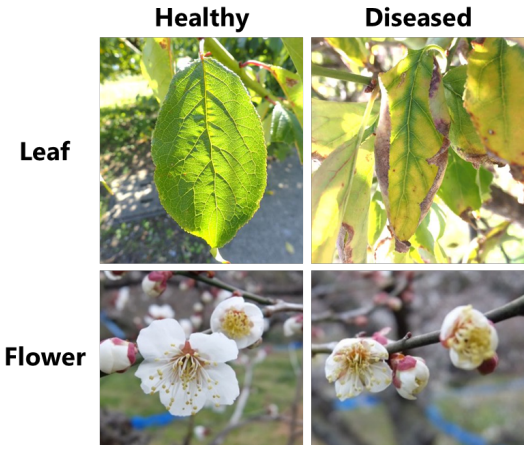

Fig. 2. Typical symptoms for mume leaf-edge necrosis in leaf (upper) and flower (lower).

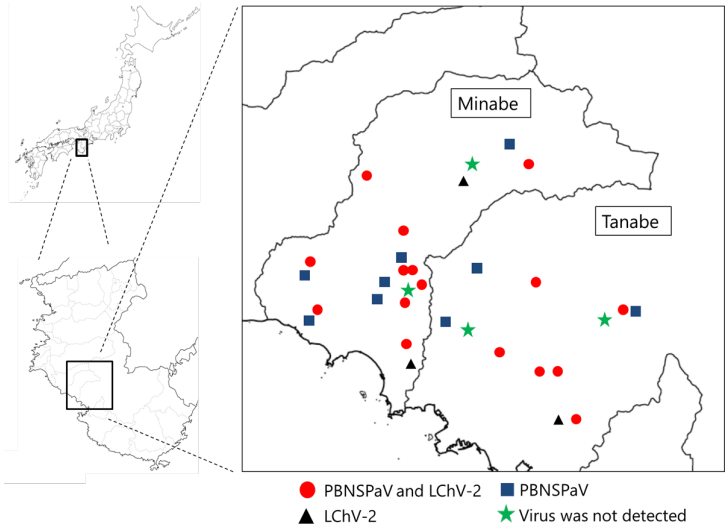

Fig. 3. Geographic distribution of PBNSPaV and LChV-2 in the midwestern districts (Minabe and Tanabe) of Wakayama Prefecture. Circles indicate the presence of both viruses. Squares and triangles represent the presence of PBNSPaV and LChV-2, respectively. Stars indicate no viruses detected.

1.6 The goal of the present study

As described above, domestication and cultivar differentiation of *P. mume* have little been elucidated despite its importance in the East Asia. In this study, genetic diversity of *P. mume* was investigated using the latest molecular and population genetic analysis to reveal the real evolutional process in *P. mume* as a fruit tree. The contents of this dissertation are as follows. In Chapter 2, microsatellite markers were newly designed based on the reference genome by Zhang et al. (2012) and examined the genetic structure of Japanese and Taiwanese populations of *P. mume*. In Chapter 3, an approximate Bayesian computation (ABC) analysis was carried out using 20 microsatellite markers to clarify the differentiation history among Japanese and Taiwanese cultivars. In Chapter 4, using target capture method (Gnirke et al., 2009) by next generation sequencing (NGS), the population structure of current *P. mume* cultivars in the East Asia was estimated based on the SNPs in 15,000 targeted exons. In addition, species differentiation was verified among *Prunus* species (*P. persica*, *P. armeniaca*, *P. salicina* and *P. mume*). In Chapter 5, on the basis of the population structure analyses, the genomic regions associated with geographic isolation (China, Japan and Taiwan) and human usage (for fruit or ornamental purposes in Japan) were estimated. Furthermore, the interspecific introgressions from *P. armeniaca* and *P. salicina* in the *P. mume* were estimated. In Chapter 6, the discussion on the evolution of *P. mume* was shown based on the overall results in the present study.

Chapter 2. Microsatellite marker development and population structure analysis in Japanese apricot (*P. mume*)

2.1 Introduction

Japanese cultivars of *P. mume* have been divided into two major groups based on their usage: fruit and flower ornamental cultivars (Mega et al., 1988; Horiuchi et al., 1996). The two groups differ in a variety of morphological traits, tree architecture, petal number, flower color, and fruit size, but clear genetic differences have yet to be reported (Shimada et al., 1994; Hayashi et al., 2008). In addition, there are some minor groups with characteristic traits. For example, cultivars in the small-fruit group literally bear very small-sized fruits (approximately less than 10 g), and mostly exhibit self-compatibility (Yaegaki et al., 2003). Taiwanese cultivars generally show a very weak chilling requirement for endodormancy release, which is a suitable survival strategy in subtropical climates (Kitamura et al., 2018). Interspecific hybrid groups derived from the crosses with apricot or Japanese plum also have very specific morphological and physiological characteristics (Yoshida and Yamanishi, 1988; Mehlenbacher et al., 1991; Yaegaki et al., 2012).

The direction of domestication is influenced by human preference (Zeder et al., 2006). In peach, for example, a distinct genetic differentiation was detected between fruit and ornamental cultivars based on genome-wide SNPs analysis (Akagi et al., 2016). However, there are not enough molecular markers for genetic analysis in Japanese apricot. Previously, microsatellite markers were designed using the peach and cherry nucleotide sequences (Gao et al., 2004), but most of them do not give informative band patterns. Hayashi et al. (2008) examined genetic diversity in Japanese apricot using microsatellite markers based on the peach and apricot genomic sequences, but the marker numbers and their genomic information are limited. Recently, Zhang et al. (2012) reported the genome sequences of eight linkage groups from a Chinese *P. mume* accession.

In this chapter, new microsatellite markers were first designed using the *P. mume* reference genome. And the selected highly polymorphic markers were used to fingerprint 124 *P. mume* accessions (mainly Japanese cultivars), as well as four other *Prunus* species (*P. armeniaca*, *P. salicina*, *P. persica*, and *P. dulcis*). The resulting genotype data were used to evaluate the genetic differentiation of Japanese apricot cultivars.

2.2 Materials and Methods

2.2.1 Plant materials

The plant materials used in this chapter are listed in Tables 1 and S1. A total of 124 *P. mume* accessions were used, including 46 fruit (F) cultivars, 10 small-fruit (FS) cultivars, 49 ornamental (O) cultivars, five Taiwanese (T) cultivars, 10 putative *P. armeniaca* × *P. mume* (AM) hybrids, and four putative *P. salicina* × *P. mume* (SM) hybrids. In addition, one accession each of apricot (Pa), Japanese plum (Ps), peach (Pp), and almond (Pd) were used as outgroup species. Of these, 61 accessions had multiple entries of different trees, and 79 accessions were the same as previously used by Hayashi et al. (2008) (Table S1). All the plant materials were maintained in the experimental orchards of the Japanese Apricot Laboratory, Wakayama Fruit Tree Experiment Station, Minabe, Wakayama, Japan. Genomic DNA was extracted from leaves using a DNeasy plant mini kit (Qiagen, Hilden, Germany).

Table 1. Number of the *Prunus* accessions and clones used in this study.

Species	Code	Group description	No. accessions	No. trees[z]
P. mume	F	Cultivars for fruits	46	95
	FS	Cultivars for small fruits (called "Ko-ume")	10	18
	O	Ornamental cultivars	49	50
	T	Taiwanese varieties	5	7
	AM	Putative hybrids (*P. mume* × *P. armeniaca*)	10	15
	SM	Putative hybrids (*P. mume* × *P. salicina*)	4	8
P. armeniaca	Pa	Apricot	1	2
P. salicina	Ps	Japanese plum	1	2
P. persica	Pp	Peach	1	1
P. dulcis	Pd	Almond	1	1
Total			128	199

[z] A total of 61 accessions had multiple tree entries.

2.2.2 Microsatellite primer design

Microsatellite primers were designed using the *P. mume* reference genome (Zhang et al., 2012). The nucleotide sequences of the eight linkage groups (NC_024126.1 for LG1, NC_024127.1 for LG2, NC_024128.1 for LG3, NC_024129.1 for LG4,

NC_024130.1 for LG5, NC_024131.1 for LG6, NC_024132.1 for LG7, and NC_024133.1 for LG8) were downloaded from the National Center for Biotechnology Information website (www.ncbi.nlm.nih.gov/genome/13911), and surveyed for microsatellite regions using a Tandem repeats finder (Benson, 1999). Microsatellite regions with ~20 GA/CT or AG/TC repeats were selected at each 1-Mb interval (a total of 201 regions). Primer pairs were then designed using Primer 3 (Untergasser et al., 2012) with the following parameters: 17–22 nucleotides long, Tm approximately 60°C, and product size in the range of 100–250 bp.

2.2.3 Microsatellite marker screening

The 201 microsatellite markers (designated as JAM, Table S2) were subjected to the following three screenings for the marker availability: 1) amplification ability with a standard cultivar of 'Nanko', 2) a wide range amplification check using eight *Prunus* accessions (three F cultivars and one each of FS, O, T, AM, and Pa), 3) polymorphism examination with 16 *P. mume* cultivars (10 F, three FS, and three O cultivars). For all screenings, PCR was performed in a 20 μL volume using Ex Taq (Takara Bio, Shiga, Japan) according to the manufacturer's instructions. The PCR conditions were as follows: an initial denaturation of 5 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 55°C, 1 min at 72°C, and 7 min at 72°C for final extension. PCR products were separated by 4% polyacrylamide gel electrophoresis and visualized by a silver staining method (Panaud et al., 1996).

2.2.4 Microsatellite marker analysis

Using a selected set of 20 highly polymorphic microsatellite markers, a total of 128 *Prunus* accessions (199 tree entries) were genotyped. In addition, 11 microsatellite markers previously reported by Hayashi et al. (2008) were also employed. To precisely determine the lengths of amplified fragments, PCR was performed using the post-labeling method described by Schuelke (2000) with minor modifications. To generate fluorescent PCR products, the U-19 universal primer (5'-GTTTTCCCAGTCACGACGT-3') was labeled with four kinds of fluorescent molecules (6-FAM, VIC, NED, and PET), as well as with 2-bp barcodes at the 3' ends (TG for 6-FAM-labeled U-19, AC for VIC-labeled U-19, CA for NED-labeled U-19, and GT for PET-labeled U-19). Non-labeled forward primers, with 7-bp 5' pig-tails (5'-GTTTCTT-3'), and U-19-fused reverse primers were also synthesized and used to perform multiplex PCR. The 20 μL PCR mixtures were

prepared using a Type-it Microsatellite PCR kit (Qiagen), with 50–100 ng template DNA, 0.2 μM each labeled U-19 primer, 0.2 μM each pig-tailed forward primer, and 0.04 μM each U-19-fused reverse primer. The PCR conditions were as follows: initial denaturation of 5 min at 95°C; followed by 10 cycles of 30 s at 95°C, 90 s at 60°C with a decrease of 0.5°C in each cycle, and 30 s at 72°C; followed by 22 cycles of 30 s at 95°C, 90 s at 55°C, and 30 s at 72°C, and a final extension step of 30 min at 60°C. Finally, the fragments were analyzed using an ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) using LIZ-600 size standard (Applied Biosystems). Alleles were defined based on fragment size (in nucleotides) and scored using Genemapper 4.1 (Applied Biosystems).

2.2.5 Genetic diversity and population structure analysis

Statistical analyses were performed after excluding two triploid cultivars (‘Horyukaku’ and ‘Takasago’) and the four accessions of other *Prunus* species (*P. armeniaca* ‘Heiwa’ was employed as an outgroup only for phylogenetic tree construction). Different genotypes found in the same accessions were independently examined in the analyses. For accessions that had identical genotypes at all loci, only one representative accession was used in the analyses, except for phylogenetic tree construction. Marker scores with null or more than three alleles were treated as missing data. However, only for the phylogenetic analysis, the score sets of loci with missing data were excluded. The number of alleles, observed heterozygosity (Ho), expected heterozygosity (He), number of effective alleles (Ne), fixation index (F_{IS}), Nei’s genetic distance (Nei et al., 1983), and G'_{ST} (Hedrick, 2005) were calculated using GenAlEx v. 6.502 (Peakall and Smouse, 2012). Here, G'_{ST} is defined as a standardized genetic differentiation measure of G_{ST} (Nei, 1972) for highly polymorphic markers (Hedrick, 2005). Polymorphism information content (PIC) was calculated using the POLYSAT package (Clark and Jasieniuk, 2011) in R (R Development Core Team, 2008). Power of discrimination (PD) (Kloosterman et al., 1993) was computed manually using the following formula: $PD = 1 - \Sigma gi^2$, where gi is the frequency of the ith genotype. Meanwhile, within-group genetic variation was assessed using the gene diversity value (D) (Nei, 1987). Statistical significance for F_{IS} and G'_{ST} was tested using GENEPOP 4.2 (Raymond and Rousset, 1995; Rousset, 2008) and GenAlEx v. 6.502, respectively. The levels of linkage disequilibrium (LD) were calculated as the squared allele-frequency correlation (r^2) values using MIDAS (Gaunt et al., 2006). The r^2 values were calculated between the most frequent alleles at each two markers, and among all combinations of alleles at each two markers and summed up with weight of allele frequencies, according to Iwata et al. (2013). The population structure

of the *P. mume* accessions was analyzed using three methods, principal coordinate analysis using GenALEx v. 6.502, neighbor-joining (NJ) tree construction using Poptree 2 with 1,000 bootstrap replicates (Takezaki et al., 2010), and individual-based Bayesian clustering using STRUCTURE 2.3.4 (Pritchard et al., 2000). The NJ tree was visualized using MEGA6 (Tamura et al., 2013). In the STRUCTURE analysis, the AM and SM hybrids were excluded in order to simplify the simulations. In addition, the number of genetic clusters (K) was set to 1–15, and 200,000 simulations were performed in each analysis after a burn-in period of 200,000 iterations. Ten runs were performed for each K value, and the optimal value was determined using L(K), |L"(K)|, and ΔK (Evanno et al., 2005), which were calculated using Structure Harvester (Earl and vonHoldt, 2012). A representative bar plot for each K was selected based on LnP(D) values generated by STRUCTURE, and visualized using Structure Plot v2.0 (Ramasamy et al., 2014).

2.3 Results

2.3.1 Selection of highly polymorphic microsatellite markers

Among the 201 microsatellite markers designed, 188 (93.5%) were successfully amplified using the template DNA of 'Nanko' (Table S2). Of these, 128 markers were randomly selected for second amplification screening. As a result, 59 markers gave clear bands for all eight *Prunus* accessions. They were preliminarily checked for the level of polymorphism using 16 *P. mume* cultivars. Their PIC values were compared and a total of 20 highly polymorphic markers were selected: at least two markers were selected from those giving the highest PIC values in each linkage group. They were renamed as PMKS markers (Table S2).

2.3.2 Allele combinations at 20 polymorphic microsatellite loci among Prunus accessions

In order to examine the allelic diversity, a total of 128 *Prunus* accessions (199 tree entries) were genotyped with a set of 20 highly polymorphic markers (Table S3). Identical genotypes for all the loci were observed among the trees in 56 out of 61 accessions having multiple entries, whereas different allele combinations were observed in the other five accessions: 'Fudono' (F4), 'Naniwa' (F24), 'PM1-1' (SM1), 'Tsuyuakane' (SM4), and 'Heiwa' (Pa) (Table S3).

Among the 124 *P. mume* accessions, a total of 107 allele combinations were

observed, with 92 of the combinations being unique to single accessions and 15 being shared by multiple accessions. Thirteen of the redundant genotypes were shared by two accessions each, and genotypes of 'B' and 'J' were observed for seven and three accessions, respectively (Table S3). In addition to the triploid cultivars ('Horyukaku' and 'Takasago'), several other accessions also gave more than three alleles at certain loci.

2.3.3 Efficacy of microsatellite markers in P. mume

The efficacy (i.e., polymorphism detection ability) of the 20 microsatellite markers was assessed with the allele combinations found in *P. mume* (Table 2). The PIC values ranged from 0.63 to 0.90 (mean = 0.79) and the PD values ranged from 0.84 to 0.97 (mean = 0.93). Using the same *P. mume* accessions, 11 microsatellite markers previously reported by Hayashi et al. (2008) were also evaluated (Tables 2 and S2). The average values of all the polymorphism indices (number of alleles, Ho, He, PIC, and PD) for the newly developed marker set were greater than those for previously reported, thereby reflecting the powerful detection ability of the new marker set. For example, 'Hakuo' (FS2), 'Koshu Saisho' (FS4), and 'Purple Queen' (FS8) shared a single genotype when using the previously developed set, whereas the newly developed markers were able to distinguish 'Hakuo' from the other two. Similarly, the new marker set was also able to distinguish 'Ikuyonezame' (O14) and 'Kinko' (O19). The F_{IS} values ranged from 0.01 to 0.46 (mean = 0.13) and 16 of 20 loci showed statistical significance at the 5% level (Table 2). The levels of LD were low among all combinations of 20 loci (Table S4).

2.3.4 Genetic variation within cultivar groups

Genetic diversity indices were calculated to evaluate variation within the six cultivar groups (F, FS, O, T, AM, and SM; Table 3), using the same set of allele combinations without duplicates. Although the Ho and gene diversity (D) values of the six *P. mume* groups were similar, a relatively high number of effective allele (Ne) values were observed for the F and O groups (Table 3).

2.3.5 Genetic differentiation among cultivar groups

Genetic differentiation among the *P. mume* cultivar groups was first evaluated using principal coordinate analysis with individual genotype data. The first principal coordinate axis (PC1) seemed to separate the F and O cultivars (Fig. 4), whereas the

Table 2. Polymorphism indices calculated for the present (A) and the previous marker sets (B).

(A)	Name	Linkage group[z]	N[y]	No. alleles	Ho[x]	He[x]	F_{IS}[x]	PIC[x]	PD[x]	Reference[w]
	PMKS15	LG1	105	17	0.76	0.86	0.11*[v]	0.84	0.96	1, 2
	PMKS21	LG1	105	20	0.84	0.87	0.04	0.86	0.96	1, 2
	PMKS49	LG2	105	15	0.70	0.84	0.16**	0.82	0.94	1, 2
	PMKS59	LG2	105	21	0.80	0.81	0.01	0.80	0.95	1, 2
	PMKS68	LG3	105	14	0.80	0.85	0.05*	0.83	0.95	1, 2
	PMKS75	LG3	105	23	0.89	0.91	0.02**	0.90	0.97	1, 2
	PMKS99	LG4	105	12	0.70	0.85	0.18**	0.83	0.94	1, 2
	PMKS113	LG4	105	12	0.76	0.81	0.06*	0.79	0.93	1, 2
	PMKS121	LG5	105	9	0.67	0.78	0.15**	0.76	0.92	1, 2
	PMKS131	LG5	98	15	0.46	0.85	0.46**	0.84	0.93	1, 2
	PMKS133	LG5	105	18	0.76	0.90	0.15**	0.89	0.96	1, 2
	PMKS149	LG6	105	17	0.70	0.75	0.08	0.73	0.91	1, 2
	PMKS164	LG6	105	8	0.60	0.72	0.17*	0.68	0.88	1, 2
	PMKS175	LG7	89	15	0.68	0.79	0.13**	0.77	0.93	1, 2
	PMKS179	LG7	105	14	0.74	0.78	0.05**	0.77	0.93	1, 2
	PMKS187	LG8	101	12	0.56	0.76	0.26**	0.74	0.88	1, 2
	PMKS191	LG8	105	14	0.64	0.74	0.14**	0.72	0.88	1, 2
	PMKS193	LG8	104	15	0.78	0.85	0.09	0.84	0.96	1, 2
	PMKS197	LG8	105	12	0.68	0.78	0.13*	0.76	0.92	1, 2
	PMKS201	LG8	105	10	0.56	0.65	0.14**	0.63	0.84	1, 2
	Mean			14.7	0.70	0.81	0.13	0.79	0.93	
(B)	Name	Linkage group[z]	N[y]	No. alleles	Ho[x]	He[x]	F_{IS}[x]	PIC[x]	PD[x]	Reference[w]
	UDP96-001	LG1	105	7	0.69	0.71	0.04	0.66	0.87	3, 7
	pchgms3	LG2	105	13	0.71	0.76	0.06	0.73	0.91	4, 7
	MA007a	LG5	105	14	0.80	0.84	0.05**	0.82	0.94	5, 7
	MA017a	Unknown	104	11	0.44	0.79	0.44**	0.77	0.90	5, 7
	MA040a	LG1	105	6	0.40	0.50	0.20	0.45	0.69	5, 7
	M6a	LG6	103	12	0.39	0.50	0.22**	0.48	0.67	5, 7
	M7a	LG6	105	10	0.38	0.71	0.47**	0.67	0.84	5, 7
	PaCITA4	LG4	105	17	0.87	0.89	0.02**	0.88	0.96	6, 7
	PaCITA7	LG2	104	19	0.88	0.91	0.03	0.91	0.98	6, 7
	PaCITA19	LG5	105	11	0.65	0.78	0.17**	0.75	0.92	6, 7
	PaCITA21	LG7	103	12	0.37	0.65	0.43**	0.62	0.79	6, 7
	Mean			12.0	0.60	0.73	0.19	0.70	0.86	

[z] After *P. mume* reference genome sequences by Zhang et al. (2012).

[y] Number of genotypes examined. Genotypes giving null or more than three alleles were excluded.

[x] Ho: observed heterozygosity, He: expected heterozygosity, F_{IS}: fixation index, PIC: polymorphism information content, PD: power of discrimination.

[w] 1: The present study, 2: Ishio et al., patent pending, 3: Testolin et al. (2000), 4: Sosinski et al. (2000), 5: Yamamoto et al. (2002), 6: Lopes et al. (2002), 7: Hayashi et al. (2008).

[v] * and **: significant at 5% and 1% level, respectively.

second axis (PC2) distinguished the T accessions from the others, and also seemed to explain within-group variation. However, even though the analysis revealed genetic differentiation among the *P. mume* accessions, the percentages of variation explained by PC1 and PC2 were relatively low (9.0% and 6.7%, respectively).

Table 3. Average genetic diversity indices for six cultivar groups in *P. mume* .

Group[z]	N[y]	No. alleles	Ho[x]	D[x]	Ne[x]
F	40	8.8	0.68	0.74	4.25
FS	9	5.3	0.71	0.67	3.49
O	40	9.0	0.71	0.78	4.91
T	5	4.5	0.64	0.65	3.35
AM	8	5.6	0.79	0.71	3.75
SM	5	3.6	0.70	0.61	3.12

[z] Group description is shown in Table 1.

[y] Number of genotypes excluding the duplicates within each group.

[x] Ho: observed heterozygosity, D: gene diversity, Ne: Number of effective alleles.

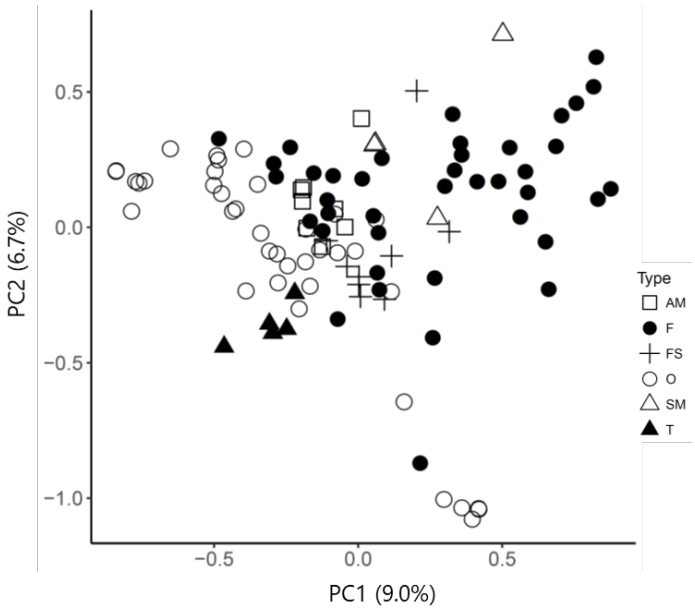

Fig. 4. Principal coordinate analysis of the six Japanese apricot cultivar groups. Values in parentheses indicate the proportion of the total variance explained by each principal coordinate. The abbreviations AM, F, FS, O, SM, and T indicate *P. armeniaca* × *P. mume* hybrids, fruit, small-fruit, ornamental cultivars, *P. salisina* × *P. mume* hybrids, and Taiwanese cultivars, respectively.

We next calculated pairwise Nei's genetic distance and G'_{st} values for the six cultivar groups (Table 4). Significant ($P < 0.01$) G'_{st} values were obtained for all pairs. Among the six groups, the T cultivars showed relatively high genetic distance values (1.053–2.042), followed by the two putative hybrid groups (SM and AM). However, in contrast to the results of the principal coordinate analysis, the genetic distance between the F and O cultivars was quite low (0.238). A similar low value (0.275) was also observed between the F and FS groups, although they have major phenotypic differences in terms of fruit size and harvesting time.

Table 4. Pairwise genetic differentiation measure, G'_{st} (below diagonal) and Nei's genetic distance value (above diagonal) between six cultivar groups.

Group[z]	F	FS	O	T	AM	SM
F		0.275	0.238	1.336	0.624	0.487
FS	0.193**[y]		0.265	1.111	0.583	0.872
O	0.201**	0.184**		1.053	0.568	0.860
T	0.730**	0.642**	0.631**		1.688	2.042
AM	0.438**	0.388**	0.393**	0.800**		1.022
SM	0.334**	0.546**	0.553**	0.859**	0.606**	

[z] Group description is shown in Table 1.

[y] **: Significant at 1% level.

2.3.6 Phylogenetic and STRUCTURE analyses

Phylogenetic analysis was performed with all the *P. mume* accessions with no missing data and *P. armeniaca* 'Heiwa' as an outgroup (Fig. 5). Three of the clades were strongly supported (bootstrap value > 70). One, which was clearly separate from the others, included the AM hybrids and the *P. armeniaca* 'Heiwa', whereas the two others contained T (T clade) and O cultivars, including 'China mume' (O-1 clade). The rest of the accessions belonged to two wider clusters that mainly included F and O cultivars.

STRUCTURE analysis was performed using 94 genotypes from the four *P. mume* cultivar groups (F, FS, O, and T). The highest ΔK value was obtained at $K = 2$, followed by $K = 7$ (Fig. 6a). In the bar plot for $K = 2$, all T cultivars showed a single orange-colored cluster (Fig. 6b), whereas the other three groups gave a mixture of two clusters. Yellow and orange clusters seemed to be dominant in the F and O cultivar groups, respectively. On the other hand, for $K = 7$, T cultivars could be clearly separated, and most FS cultivars showed a single red cluster. A purple cluster seemed to be dominant in the F group. O cultivars consisted of many admixture types, but cultivars belonging to the O-1 clade in the phylogenetic tree shared the same navy-colored cluster.

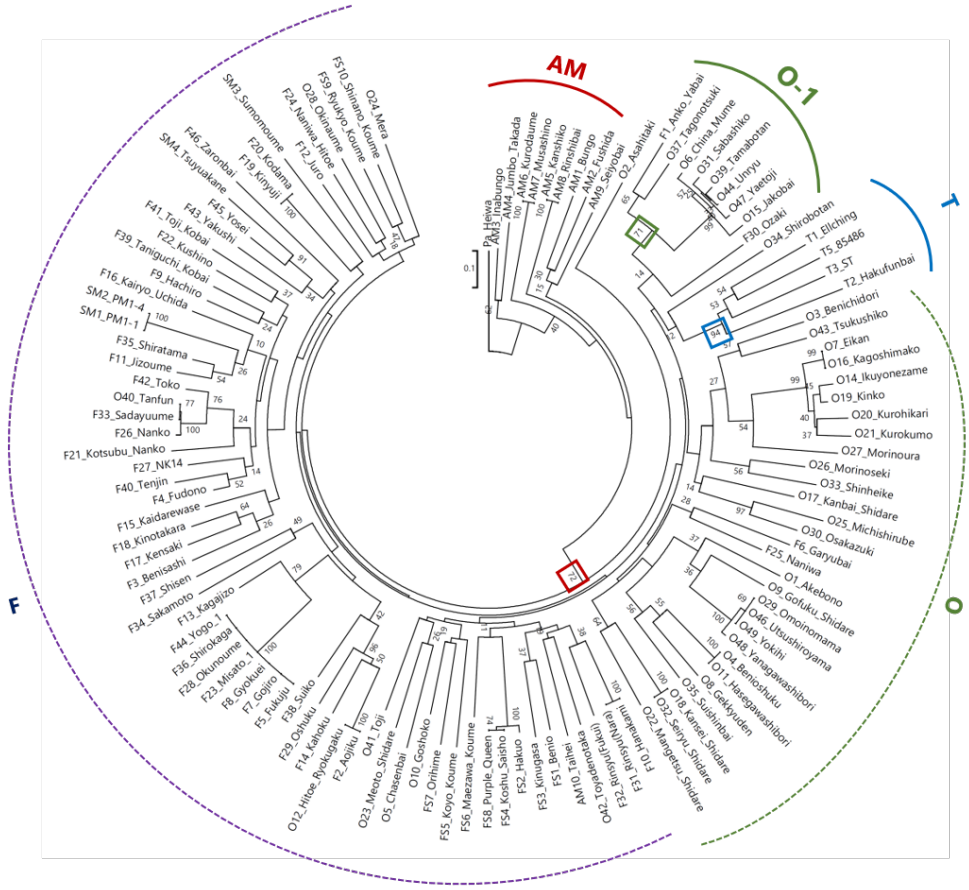

Fig. 5. Neighbor joining tree of *P. mume* accessions based on the allele frequency of 20 polymorphic microsatellite markers. *P. armeniaca* 'Heiwa' was used as an outgroup. Solid lines indicate clades (red, putative *P. armeniaca* × *P. mume* hybrids; blue, Taiwanese cultivars; and green, ornamental cultivars). Dotted lines indicate groups that mainly consist of fruit cultivars (purple) and ornamental cultivars (light green).

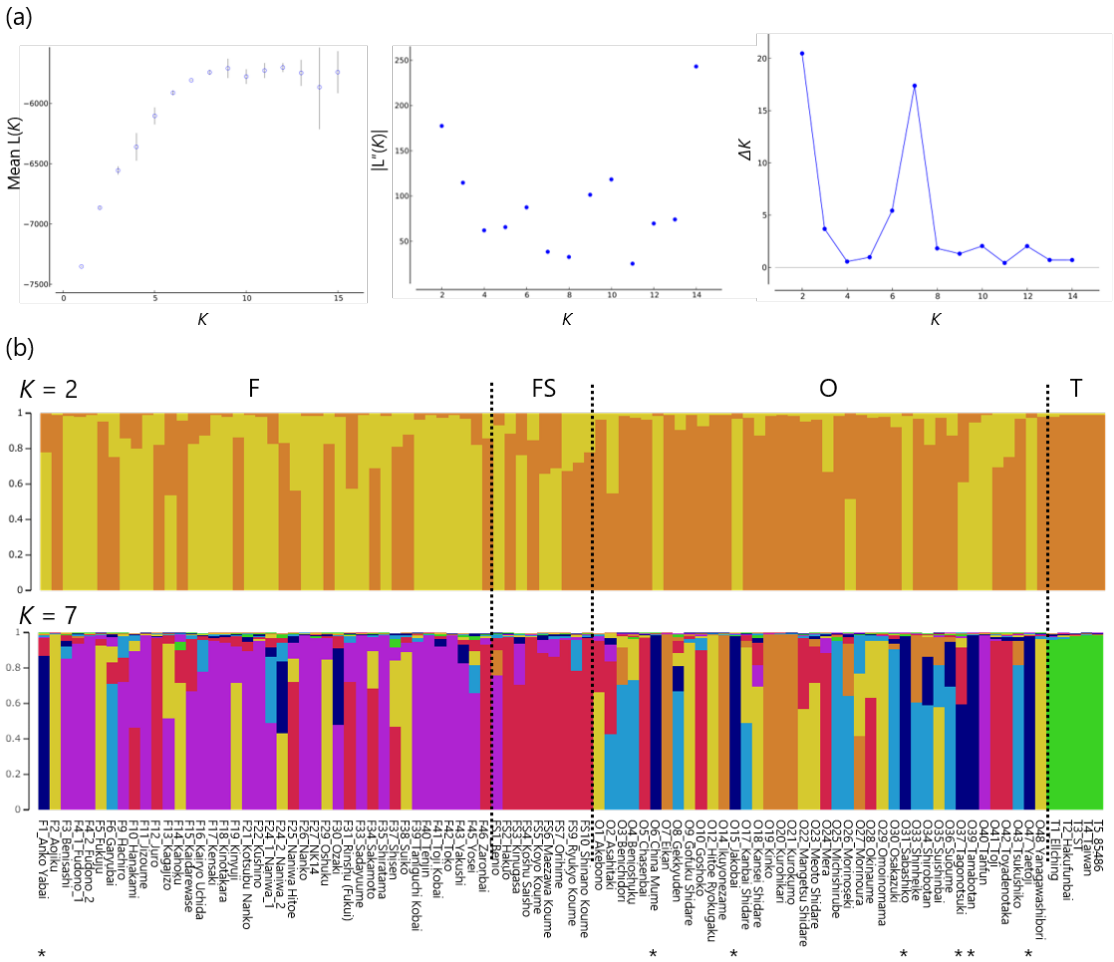

Fig. 6. STRUCTURE analysis of *P. mume* cultivars. Fruit (F), small-fruit (FS), flower ornamental (O), and Taiwanese (T) cultivars were employed for the simulations. (a) Values of mean L(*K*), |L"(*K*)|, and Δ*K*. The highest value was obtained at *K* = 2, implying the existence of two clusters. The Δ*K* value was also high at *K* = 7, indicating the existence of 7 clusters. (b) Bar plots for each value of *K*, at *K* = 2 and 7. Representative bar plots were selected based on LnP(D) values generated by STRUCTURE software. Cultivars belonging to the O-1 clade (Fig. 6) are indicated by asterisks.

2.4 Discussion

2.4.1 A new marker set of highly polymorphic microsatellites

Studies on genetic diversity and clone identification in *P. mume* have been relatively limited owing to a lack of effective molecular tools. Therefore, a total of 201 new microsatellite markers were designed using the *P. mume* reference genome reported by Zhang et al. (2012), and selected 20 highly polymorphic markers for further analysis. As shown in Tables 2 and S2, the present marker set exhibited greater polymorphism detection ability than those previously developed using the peach and apricot genomes (Hayashi et al., 2008) and was successfully used to fingerprint most of the *Prunus* cultivars (128 accessions, 199 trees). However, 15 redundant genotypes (designated as A to O in Table S3) were shared by two or more accessions. Hayashi et al. (2008) also reported that identical genotypes were detected among cultivars having very similar morphological characteristics, namely, among the three cultivars of 'Shirokaga' (F36), 'Gojiro' (F7), and 'Gyokuei' (F8), two of 'Kodama' (F20) and 'Kinyuji' (F19), and two of 'Hanakami' (F10) and 'Rinshu (Nara)' (F32). Since these accessions may be recent somaclonal variants, resulting from bud sports with slight mutations, it is quite difficult to distinguish them using microsatellite markers. Therefore, next generation sequencing will likely be needed to detect cultivar-specific SNPs.

On the other hand, different allele combinations were observed among tree entries in the five accessions: 'Fudono' (F4), 'Naniwa' (F24), 'PM1-1' (SM1), 'Tsuyakane' (SM4), and 'Heiwa' (Pa) (Table S3). Two different genotypes between trees were observed in 'Fudono' and 'Naniwa'. Since they are old local varieties, some chance seedlings showing similar phenotypes may have been cultivated as the same varieties. For the modern varieties 'PM1-1' and 'Tsuyuakane', length mutations were observed between two tree entries only at the PMKS113 locus, suggesting that they were generated by recent somaclonal variation. Two 'Heiwa' trees were identical at 15 of the 20 loci. 'Heiwa' is quite an old variety. Probably, the difference was also caused by somaclonal variation.

In recent years, Japanese fruit crops (strawberry, sweet cherry, and grapevine) have been taken overseas and cultivated. To control such dissemination, domestic fruit cultivars should be protected by international patent or registration with fingerprinting data. The highly polymorphic marker set developed here may be suitable for identifying cultivars of Japanese apricot, and moreover, 201 markers designed to cover whole genome linkage groups in *P. mume* (Table S2) can be useful for linkage and QTL analyses

to improve Japanese apricot.

Core collections have been reported for many crops, such as rice (Ebana et al., 2008), soybean (Kaga et al., 2012), and strawberry (Wada et al., 2017). They allow us to maintain a minimum number of genetic resources with maximum genetic variation. The present marker set can be utilized to develop a core collection of Japanese apricot to enhance efficient breeding.

2.4.2 Population structure in P. mume accessions

Among the *P. mume* accessions, the average observed heterozygosity of 20 PMKS markers was 0.70, which was lower than that of expected heterozygosity (0.81, Table 2), and most of the F_{IS} values showed significance at the 5% level. These results are due to the fact that the present group of Japanese accessions is not a natural population with random mating. Probably, Japanese accessions went through strong artificial selection, introgression and specific breeding. The levels of LD were low (Table S4), suggesting no association of alleles at 20 loci.

Generally, DNA markers with non-significant F_{IS} values are suitable for population structure analyses. In this study, all the 20 PMKS markers were preliminarily used to clarify the present genetic differentiation of *P. mume* accessions through three analyses (principal coordinate analysis, phylogenetic cluster analysis, and STRUCTURE analysis).

Among six cultivar groups, putative Taiwanese (T) cultivars were clearly distinguished from others in both phylogenetic and STRUCTURE analyses (Figs. 5 and 6b). The results are consistent with those of Hayashi et al. (2008) using SSRs derived from peach and apricot, and Shimada et al. (1994) using RAPD markers. The T cultivars are uniquely characterized by their weak bud dormancy, which enables them to adapt in subtropical climates (Kitamura et al., 2018). The AM cultivars except ‘Taihei’ (AM10) were clearly distinguished in the phylogenetic analysis (Fig. 5). They were derived from interspecific hybrids between *P. mume* and *P. armeniaca* (Yoshida and Yamanishi, 1988; Mehlenbacher et al., 1991). Although ‘Taihei’ (AM10) showed evidence for inheritance of *P. armeniaca* characteristics (Hayashi et al., 2008; Hayashi, 2009), it was included in the huge cluster of Japanese apricot. Probably, ‘Taihei’ went through a subsequent cross with Japanese apricot cultivars. Other F, FS, and O cultivars were not clearly divided into the respective cultivar groups. They formed a complex genetic structure in the Japanese population.

In this study, the genetic differentiation among Japanese apricot was examined

by three analyses. However, the exact sequence of the differentiation process was not revealed because Japanese apricot is thought to have originated in China and many cultivars may have been occasionally introduced to Japan. Many factors, such as human preference, geographical separation, introgression, and local breeding, may be involved to form the complex genetic structure in Japanese apricot. More comprehensive analytical methods (e.g., genome-wide SNP survey with wider genetic resources including Chinese cultivars and wild relatives) will shed light on the details of this species' evolution, domestication, and improvement history.

2.5 Summary

Japanese apricot (*Prunus mume* Sieb. et Zucc.) is one of the major fruit tree crops in Japan. However, a paucity of molecular tools has limited studies on the species' genetic diversity and clone identification. Therefore, a total of 201 microsatellite markers were newly designed using the *P. mume* reference genome and selected 20 highly polymorphic markers. The markers showed higher polymorphism detectability than those previously developed using peach and apricot genomes. They were used successfully for fingerprinting most of the *Prunus* cultivars examined (124 *P. mume* accessions and one accession each of *P. armeniaca*, *P. salicina*, *P. persica*, and *P. dulcis*), and the resulting genotype data were used to examine the genetic differentiation of six Japanese apricot cultivar groups, including those producing normal fruit, small-fruit, and ornamental flowers, as well as Taiwanese cultivars, putative hybrids of *P. armeniaca* and *P. mume*, and putative hybrids of *P. salicina* and *P. mume*. Phylogenetic cluster analysis showed three clades with high support values; one clade comprised the putative *P. armeniaca* × *P. mume* hybrids, and the two others included Taiwanese and ornamental cultivars. The rest of the accessions were clustered into two wide clusters, but not clearly divided into the respective cultivar groups. These complex relationships were supported by the principal coordinate and STRUCTURE analyses. Since Japanese apricot is thought to have originated in China, many factors such as human preference, geographical separation, introgression, and local breeding, may have been involved to form the present complex genetic structure in Japanese apricot.

Chapter 3. Estimation of demographic history of Japanese and Taiwanese populations in Japanese apricot based on microsatellite marker genotypes

3.1 Introduction

The origin of *P. mume* is not clearly determined, although it is believed to be originated from the mid-mountainous region in China (Mega et al., 1988). More than 300 *P. mume* cultivars are distributed mainly in East Asia, and they usually have self-incompatibility and require a certain amount of low temperature to break bud dormancy (Yamane 2014). On the other hand, Taiwanese cultivars adapted to subtropical climates generally show a weak chilling-requirement for dormancy release, and they were reported to be genetically different from Japanese cultivars (Shimada et al., 1994; Hayashi et al., 2008; Kitamura et al., 2018; Chapter 2). Japanese cultivars are thought to be introduced from China about 2,000 years ago for flower ornamental purposes (Mega et al., 1988). Then, fruit cultivars (including small-fruit cultivars) may have been generated in Japan. Interestingly, small-fruit cultivars mostly bearing small-sized fruits (~10 g) and showing self-compatibility are also preferred in some parts of Japan. As shown above, three major groups of ornamental, fruit and small-fruit cultivars are common in Japan.

Population structure among Japanese and Taiwanese cultivars of *P. mume* have been analyzed using random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) markers (Shimada et al., 1994; Hayashi et al., 2008; Chapter 2). However, these analyses could show only the present population structure, but not infer the demographic history such as change of effective population size and divergence time of subpopulations. Approximate Bayesian computation (ABC) (Beaumont et al., 2002) is a computational method which estimates the best fitting demographic model by comparing simulated datasets of assumed models and observed genotypes. Using datasets of simple DNA markers (e.g., SSR markers), this method has been successfully applied for various woody plant species, such as apple and olive (Cornille et al., 2012; Diez et al., 2015). In this chapter, the demographic history among Japanese and Taiwanese populations of *P. mume* was analyzed based on the ABC using SSR genotype data obtained in Chapter 2.

3.2 Materials and methods

3.2.1 Plant materials and datasets

A total of 124 *P. mume* accessions were subjected to the STRUCTURE analysis (Pritchard et al., 2000) using 20 SSR markers (PMKS15, 21, 49, 59, 68, 75, 99, 113, 121, 131, 133, 149, 164, 175, 179, 187, 191, 193, 197, 201) in Chapter 2. Of these, the representative cultivars forming mode clusters (or core clusters) in the STRUCTURE analysis in Chapter 2 were selected for four subpopulations of Pop1 (fruit cultivars), Pop2 (small-fruit cultivars), Pop3 (ornamental cultivars) and Pop4 (Taiwanese cultivars). For Pop1 and Pop3, the top 20 cultivars having high values of dominant clusters at $K = 2$ were selected, respectively. Similarly, eight and five cultivars sharing common clusters at $K = 7$ were selected for Pop2 and Pop4, respectively. The variety names of these 53 cultivars are listed in Table 5. All the plant materials were maintained in Japanese Apricot Laboratory, Wakayama Fruit Experiment Station, Japan.

ABC analysis was carried out using genotype datasets of the 53 cultivars based on 20 SSR markers. These datasets are available in Table S3.

Table 5. Plant materials used for the analysis.

Population[z]	Cultivar or accession name				
Pop1: Fruit (20)	Benisashi	Fudono_1	Fudono_2	Hachiro	Jizoume
	Kairyouchida	Kensaki	Kinotakara	Kotsubunanko	Kushino
	Nanko	NK14	Ozaki	Sadayuume	Shiratama
	Taniguchikobai	Tenjin	Tojikobai	Toko	Yakushi
Pop2: Small-fruit (8)	Hakuo	Kinugasa	Koshusaisho	Koyokoume	Maezawakoume
	Orihime	Ryukyokoume	Shinanokoume		
Pop3: Ornamental (20)	Benichidori	Benioshuku	Chasenbai	Eikan	Gofukushidare
	Hitoeryokugaku	Ikuyonezame	Kanbaishidare	Kinko	Kurohikari
	Kurokumo	Meotoshidare	Michishirube	Morinoura	Okinaume
	Omoinomama	Shinheike	Suishinbai	Tsukushiko	Yanagawashibori
Pop4: Taiwanese (5)	Ellching	Hakufunbai	ST	Taiwan	85486

[z] Parenthesis indicates number of cultivars in the population.

3.2.2 Estimation of demographic model

To make approximate Bayesian computation, a software of DIYABC v2.0 (Cornuet et al., 2014) was used to infer the demographic history. According to the results on genetic distances among four subpopulations in Chapter 2, Taiwanese group was inferred to diverge first. Therefore, the six following scenarios with three divergence times measured in number of generations (t1–t3) were assumed (Fig. 7). Scenario 1: all populations (Pop1–4) were simultaneously differentiated at t3. Scenario 2: Japanese (Pop1–3) and Taiwanese (Pop4) populations were first differentiated at t3, followed by

the ornamental cultivars (Pop3) at t2, and fruit (Pop1) and small-fruit cultivars (Pop2) at t1. Scenarios 3 and 4: After the differentiation of Taiwanese population at t3, fruit cultivars (Pop1) or small-fruit cultivars (Pop2) were next differentiated at t2, followed by the other two. Scenarios 5 and 6: After the Taiwanese separation at t3, two populations were differentiated at t2, followed by the emergence of introgressed small-fruit (Pop2) or fruit cultivars (Pop1) at t1. Under each assumed model, a computation was carried out with 1×10^6 simulated datasets using DIYABC. Here, generalized stepwise model with single nucleotide insertion or deletion was employed as the SSR mutation model. Posterior probabilities indicating the fitness of the assumed scenarios were calculated based on polychotomous logistic regression, according to Cornille et al. (2012) and Diez et al. (2015).

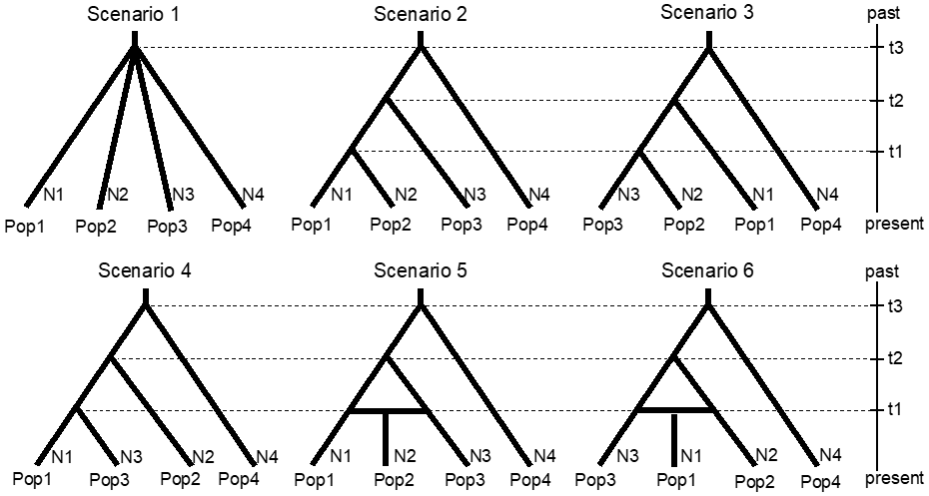

Fig. 7. Demographic models of six scenarios compared in approximate Bayesian computations.
Pop1: Fruit cultivars. Pop2: Small-fruit cultivars. Pop3: Ornamental cultivars. Pop4: Taiwanese cultivars. N1 - N4: Effective population sizes for Pop1 – 4. t1– t3: Divergence times.

3.3 Results and discussion

Population structures revealed by STRUCTURE analyses may often be affected by recent admixture in the populations (Pritchard et al., 2000; Anderson and Thompson, 2002; Excoffier et al., 2005). Therefore, in this chapter, the demographic history among four major subpopulations was estimated in *P. mume* by ABC method.

Based on the results on genetic distances among four subpopulations in Chapter 2, six scenarios with divergence times are assumed. Of these, DIYABC analysis supports Scenario 2 with highest posterior probability of 0.501 (95% confidence interval: 0.482–0.521) (Table 6). In this scenario, Japanese (Pop1–3) and Taiwanese (Pop4) populations had first diverged at t3, followed by differentiation between fruit (Pop1–2) and ornamental (Pop3) cultivars at t2, and finally fruit (Pop1) and small-fruit (Pop2) cultivars separated at t1 (Fig. 7). The 95% confidence interval (CI) of Scenario 2 did not overlap with those for the other scenarios explaining the separation of fruit and small-fruit cultivars was not recent event (Scenarios 1, 3, 4) or the admixture event occurred among Japanese populations (Scenarios 5, 6) (Table 6). Scenario 2 seems to be consistent with the suggestion in the previous studies, i.e., Japanese cultivars were originally introduced from China for ornamental purpose, and fruit and small-fruit cultivars were subsequently produced in Japan (Mega et al., 1988; Shimada et al., 1994; Hayashi et al., 2008).

Table 6. Relative posterior probabilities for the six demographic models based on DIYABC analysis.

Scenario	p [z]	CI (0.025) [y]	CI (0.975) [y]
1	0.148	0.126	0.170
2	0.501	0.482	0.521
3	0.013	0.000	0.034
4	0.003	0.000	0.025
5	0.261	0.234	0.289
6	0.073	0.054	0.092

[z] Posterior probabilities.

[y] Boundaries of the 95% confidence intervals (CI).

For the best fitting model of Scenario 2, further parameters were estimated. Table 7 shows the effective population sizes (N1–N4), divergence times (t1–t3) and mutation-related mean values (μmic, pmic, snimic). Among them, divergence times which are critical in the demographic scenario were further examined. The median values of the three divergence times of t1, t2 and t3 were 341 (95% CI: 89–894), 519 (95% CI: 174–1,820) and 1,480 (95% CI: 297–8,130) generations ago, respectively. It is quite difficult to determine the generation time in woody plant species (Tsuda et al., 2015). In peach, Yu et al. (2018) assumed the generation time of seven years (full reproductive age) for the demographic study. As *P. mume* shows similar growth patterns to peach, the same generation time for *P. mume* cultivars was applied for time conversion. Consequently, the median values of t1, t2 and t3 were estimated to be 2,387 (95% CI: 623–6,258), 3,633 (95% CI: 1,218–12,740) and 10,360 (95% CI: 2,079–56,910) years ago, respectively.

These results roughly suggest that Japanese and Taiwanese populations were separated through the geographic isolation with different climate conditions, and ornamental, fruit and small-fruit cultivars were recently differentiated based on human preference in Japan. Although these median values have broad 95% confidence intervals, the differentiation period between Japanese and Taiwanese populations seems to be overlapped with the bottleneck periods of peach species (Yu et al., 2018). This suggests that the global warming climate after the last glacial period may also promote the population expansion and differentiation in *P. mume*. In addition, divergence times among ornamental, fruit and small-fruit cultivars were closely estimated to the beginning period of *P. mume* utilization in Japan (Mega et al., 1988). Given that similar genetic structures tend to be shared within each cultivar group in Chapter 2, Japanese people may have selected *P. mume* trees based on the preferable traits such as flower shape and color, tree architecture, fruit size and so on.

In this chapter, the demographic history of Japanese and Taiwanese populations was roughly estimated in *P. mume*. For further study, more plant materials (including Chinese cultivars and wild relatives) and more marker data are necessary to clarify the domestication and differentiation process in *P. mume*.

Table 7. Demographic and mutation parameters estimated for Scenario 2.

Parameter[z]	Median	CI (0.025)[y]	CI (0.975)[y]
N1	946	314	3840
N2	4740	1330	9410
N3	3240	1050	8360
N4	5150	2020	9270
t1	341	89.1	894
t2	519	174	1820
t3	1480	297	8130
μmic	3.1×10^{-4}	1.7×10^{-4}	7.5×10^{-4}
pmic	0.29	0.19	0.3
snimic	6.6×10^{-8}	1.0×10^{-8}	2.2×10^{-6}

[z] N1–N4: Effective population sizes. t1–t3: Divergence times (generations). μmic: Mean mutation rate of SSR. pmic: Mean increase or decrease of the length of the locus during mutation events. snimic: Mean mutation rate of single nucleotide insertion or deletion.

[y] Boundaries of the 95% confidence intervals (CI).

3.4 Summary

More than 300 *P. mume* cultivars are distributed mainly in East Asia, including Japanese fruit, small-fruit and ornamental cultivars and Taiwanese cultivars. In order to estimate demographic history of three Japanese and one Taiwanese subpopulations, approximate Bayesian computation analysis was carried out using 20 SSR genotype datasets of 53 cultivars (20 fruit, 8 small-fruit, 20 ornamental and 5 Taiwanese cultivars). At first, the best fitting model (posterior probability: 0.501) was estimated among six probable scenarios, and median values of demographic parameters were computed. The generation time for *P. mume* cultivars was assumed to be seven years (full reproductive age) for time conversion. In the best scenario, Japanese and Taiwanese populations had first diverged at 10,360 (95% confidence interval (CI): 2,079-56,910) years ago, followed by the separation of ornamental cultivars among Japanese populations at 3,633 (95% CI: 1,218-12,740) years ago, and final differentiation between fruit and small-fruit cultivars at 2,387 (95% CI: 623-6,258) years ago. Although the divergence times were roughly estimated, the results suggest that Japanese and Taiwanese populations were separated through the geographic isolation with different climate conditions, and ornamental, fruit and small-fruit cultivars were recently differentiated based on human preference in Japan.

Chapter 4. Population structure analyses for the East Asian cultivars of Japanese apricot based on exon capture resequencing

4.1 Introduction

Japanese apricot (*P. mume*) is believed to have been domesticated firstly in China several thousand years ago, and then moved into Japan ca. 2,000 years ago, originally for ornamental purposes (Mega et al., 1988; Horiuchi et al., 1996; Faust et al., 2011). Currently, cultivars are widely diversified mainly based on their usages, such as for pickles ("umeboshi"), syrups/liquors, and ornamental flowers. However, in contrast to historical implications and conventional categorization, the genetic background of this species remains little known.

Population structure of *P. mume* have been investigated using PCR-based DNA markers in Chapter 2 and several other studies (Shimada et al., 1994; Hayashi, 2009; Yuying et al., 2011). However, due to the paucity of marker numbers and polymorphism detection abilities, information on the differentiation was limited, and conclusions of these studies were slightly different from each other. For example, Hayashi et al. (2008) concluded that *P. mume* could not be divided into subgroups except Taiwanese cultivars. On the other hand, Yuying et al. (2011) and Chapter 2 in this dissertation pointed out that fruit and ornamental cultivars tended to be clustered in each group. These differences may be caused by the limited number of PCR-based DNA markers, and more comprehensive approach is necessary to reveal the current population structure in *P. mume*. Therefore, in this chapter, the author conducted the exon-targeted resequencing of 129 genomes in the subgenus *Prunus*, Japanese apricot (*P. mume*), apricot (*P. armeniaca*), Japanese plum (*P. salicina*) and peach (*P. persica*). The data were merged with published resequencing data of 79 Chinese *P. mume* cultivars (Zhang et al., 2018) to infer the current population structure among the three East Asian geographic groups, Chinese, Japanese and Taiwanese cultivar groups of *P. mume*.

4.2 Materials and methods

4.2.1 Plant materials (Japanese and Taiwanese cultivars and Prunus relatives)

One hundred twelve Japanese and 5 Taiwanese cultivars of Japanese apricot (*P. mume*), and 7 apricot (*P. armeniaca*), 4 Japanese plum (*P. salicina*), and 1 peach (*P.*

persica) cultivars were used (Table 8). For Japanese cultivars of *P. mume*, the author used 55 fruit and 45 ornamental cultivars, 8 hybrids between *P. mume* and *P. armeniaca*, and 4 hybrids between *P. mume* and *P. salicina*. Cultivar categorization was based on Chapter 2 and the previous report (Hayashi et al., 2008). All plant materials were maintained at the Japanese Apricot Laboratory, Wakayama Fruit Experiment Station (Minabe, Wakayama, Japan).

Table 8. Japanese and Taiwanese cultivars and other *Prunus* species used in this study.

Species	Name[1]		ID[2]	Group[3]	Country	Location	JP acc. no.[4]	SRA acc. no.[5]	
P. mume	Ankoyabai	*	Jap_F1	F	Japan	-	-	DRR212437	
	Aojiku	*	Jap_F2	F	Japan	Nara	-	DRR212438	
	Benisashi	*	Jap_F3	F	Japan	Fukui	113065	DRR212439	
	Fudono	*	Jap_F4	F	Japan	-	170637	DRR212440	
	Fukuju	*	Jap_F5	F	Japan	-	-	DRR212441	
	Garyubai	*	Jap_F6	F	Japan	-	-	DRR212442	
	Gojiro	*	Jap_F7	F	Japan	Wakayama	172766	DRR212443	DRR212566
	Gyokuei	*	Jap_F8	F	Japan	Tokyo	170659	DRR212444	DRR212567
	Hachiro	*	Jap_F9	F	Japan	Ibaraki	-	DRR212445	
	Hanakami	*	Jap_F10	F	Japan	-	170639	DRR212446	
	Jizoume	*	Jap_F11	F	Japan	Wakayama	172768	DRR212447	
	Juro	*	Jap_F12	F	Japan	Kanagawa	172769	DRR212448	
	Kagajizo	*	Jap_F13	F	Japan	Ibaraki	-	DRR212449	
	Kahoku	*	Jap_F14	F	Japan	-	-	DRR212450	
	Kaidarewase	*	Jap_F15	F	Japan	Wakayama	-	DRR212451	
	Kairyouchida	*	Jap_F16	F	Japan	Wakayama	170661	DRR212452	
	Kensaki	*	Jap_F17	F	Japan	Fukui	170644	DRR212453	
	Kinotakara	*	Jap_F18	F	Japan	Mie	-	DRR212454	DRR212568
	Kinyuji	*	Jap_F19	F	Japan	Osaka	-	DRR212455	
	Kodama	*	Jap_F20	F	Japan	-	-	DRR212456	
	Kotsubunanko	*	Jap_F21	F	Japan	Wakayama	-	DRR212457	
	Kushino	*	Jap_F22	F	Japan	-	-	DRR212458	
	Misato1	*	Jap_F23	F	Japan	Wakayama	-	DRR212459	
	Naniwa	*	Jap_F24	F	Japan	-	172772	DRR212460	
	Nanko	*	Jap_F25	F	Japan	Wakayama	172773	DRR212461	
	NK14	*	Jap_F26	F	Japan	Wakayama	-	DRR212462	
	Okunoume	*	Jap_F27	F	Japan	-	-	DRR212463	
	Oshuku	*	Jap_F28	F	Japan	Tokushima	172777	DRR212464	
	Ozaki	*	Jap_F29	F	Japan	-	-	DRR212465	
	Rinshu-Fukui	*	Jap_F30	F	Japan	Fukui	-	DRR212466	
	Rinshu-Nara	*	Jap_F31	F	Japan	Nara	170647	DRR212467	
	Sadayuume	*	Jap_F32	F	Japan	-	-	DRR212468	
	Sakamoto	*	Jap_F33	F	Japan	-	-	DRR212469	
	Seiko	*	Jap_F34	F	Japan	Wakayama	-	DRR212470	
	Shiratama	*	Jap_F35	F	Japan	Wakayama	113054	DRR212471	
	Shirokaga	*	Jap_F36	F	Japan	-	172785	DRR212472	
	Shisen	*	Jap_F37	F	Japan	-	-	DRR212473	
	Suiko	*	Jap_F38	F	Japan	Ibaraki	-	DRR212474	
	Tenjin	*	Jap_F39	F	Japan	-	-	DRR212475	
	Tojikobai	*	Jap_F40	F	Japan	-	-	DRR212476	
	Toko	*	Jap_F41	F	Japan	Wakayama	-	DRR212477	
	Yakushi	*	Jap_F42	F	Japan	Wakayama	174252	DRR212478	
	Yogo1	*	Jap_F43	F	Japan	Wakayama	-	DRR212479	
	Yosei	*	Jap_F44	F	Japan	Wakayama	174255	DRR212480	
	Zaronbai	*	Jap_F45	F	Japan	-	-	DRR212481	

Table 8. (Continued).

Species	Name[1]		ID[2]	Group[3]	Country	Location	JP acc. no.[4]	SRA acc. no.[5]	
	Benio	*	Jap_FS1	FS	Japan	Wakayama	-	DRR212482	
	Hakuo	*	Jap_FS2	FS	Japan	Wakayama	-	DRR212483	
	Kinugasa	*	Jap_FS3	FS	Japan	Wakayama	-	DRR212484	
	Koshusaisho	*	Jap_FS4	FS	Japan	Nara	113057	DRR212485	
	Koyokoume	*	Jap_FS5	FS	Japan	Nara	-	DRR212486	
	Maezawakoume	*	Jap_FS6	FS	Japan	Nagano	-	DRR212487	
	Orihime	*	Jap_FS7	FS	Japan	Saitama	172776	DRR212488	
	Purplequeen	*	Jap_FS8	FS	Japan	Wakayama	-	DRR212489	
	Ryukyokoume	*	Jap_FS9	FS	Japan	Nagano	172779	DRR212490	
	Shinanokoume	*	Jap_FS10	FS	Japan	Nagano	-	DRR212491	
	Akebono	*	Jap_O1	O	Japan	-	172764	DRR212492	
	Asahitaki	*	Jap_O2	O	Japan	-	-	DRR212493	
	Benichidori	*	Jap_O3	O	Japan	-	-	DRR212494	
	Chasenbai	*	Jap_O4	O	Japan	-	-	DRR212495	
	Chinamume	*	Jap_O5	O	Japan	-	-	DRR212496	
	Eikan	*	Jap_O6	O	Japan	-	-	DRR212497	
	Gekkyuden	*	Jap_O7	O	Japan	-	-	DRR212498	
	Gofukushidare	*	Jap_O8	O	Japan	-	-	DRR212499	DRR212569
	Goshoko	*	Jap_O9	O	Japan	-	-	DRR212500	
	Hasegawashibori	*	Jap_O10	O	Japan	-	-	DRR212501	
	Hitoeryokugaku	*	Jap_O11	O	Japan	-	-	DRR212502	
	Ikuyonezame	*	Jap_O12	O	Japan	-	-	DRR212503	
	Jakobai	*	Jap_O13	O	Japan	-	-	DRR212504	
	Kagoshimako	*	Jap_O14	O	Japan	-	-	DRR212505	
	Kanbaishidare	*	Jap_O15	O	Japan	-	-	DRR212506	
	Kanseishidare	*	Jap_O16	O	Japan	-	-	DRR212507	
	Kinko	*	Jap_O17	O	Japan	-	-	DRR212508	
	Kurohikari	*	Jap_O18	O	Japan	-	-	DRR212509	
	Kurokumo	*	Jap_O19	O	Japan	-	-	DRR212510	
	Mangetsushidare	*	Jap_O20	O	Japan	-	170671	DRR212511	
	Meotoshidare	*	Jap_O21	O	Japan	-	-	DRR212512	
	Mera	*	Jap_O22	O	Japan	-	-	DRR212513	
	Michishirube	*	Jap_O23	O	Japan	-	170672	DRR212514	
	Morinoseki	*	Jap_O24	O	Japan	-	-	DRR212515	
	Morinoura	*	Jap_O25	O	Japan	-	-	DRR212516	
	Okinaume	*	Jap_O26	O	Japan	-	-	DRR212517	
	Omoinomama	*	Jap_O27	O	Japan	-	-	DRR212518	
	Osakazuki	*	Jap_O28	O	Japan	-	-	DRR212519	
	Sabashiko	*	Jap_O29	O	Japan	-	-	DRR212520	
	Seiryushidare	*	Jap_O30	O	Japan	-	-	DRR212521	
	Shinheike	*	Jap_O31	O	Japan	-	-	DRR212522	
	Shirobotan	*	Jap_O32	O	Japan	-	174237	DRR212523	
	Suishinbai	*	Jap_O33	O	Japan	-	-	DRR212524	
	Tagonotsuki	*	Jap_O34	O	Japan	-	-	DRR212525	
	Tamabotan	*	Jap_O35	O	Japan	-	174244	DRR212526	
	Tanfun	*	Jap_O36	O	Japan	-	-	DRR212527	
	Toji	*	Jap_O37	O	Japan	-	174249	DRR212528	
	Toyadenotaka	*	Jap_O38	O	Japan	-	-	DRR212529	
	Tsukushiko	*	Jap_O39	O	Japan	-	-	DRR212530	
	Unryu	*	Jap_O40	O	Japan	-	-	DRR212531	
	Unryubai	*	Jap_O41	O	Japan	-	-	DRR212532	
	Utsushiroyama	*	Jap_O42	O	Japan	-	-	DRR212533	
	Yaetoji	*	Jap_O43	O	Japan	-	-	DRR212534	
	Yanagawashibori	*	Jap_O44	O	Japan	-	-	DRR212535	
	Yokihi	*	Jap_O45	O	Japan	-	-	DRR212536	

Table 8. (Continued).

Species	Name[1]		ID[2]	Group[3]	Country	Location	JP acc. no.[4]	SRA acc. no.[5]	
	Bungo	*	Jap_AM1	AM	Japan	Oita	-	DRR212537	
	Fushida	*	Jap_AM2	AM	Japan	-	-	DRR212538	
	Jumbotakada	*	Jap_AM3	AM	Japan	Fukushima	-	DRR212539	
	Kanshikobai	*	Jap_AM4	AM	Japan	-	-	DRR212540	
	Kurodaume	*	Jap_AM5	AM	Japan	-	-	DRR212541	
	Musashino	*	Jap_AM6	AM	Japan	-	-	DRR212542	
	Seiyobai	*	Jap_AM7	AM	Japan	Hokkaido	172782	DRR212543	
	Taihei	*	Jap_AM8	AM	Japan	-	174241	DRR212544	
	Beninomai		Jap_SM1	SM	Japan	Gumma	-	DRR212545	
	PM1-1	*	Jap_SM2	SM	Japan	Ibaraki	-	DRR212546	
	Sumomoume	*	Jap_SM3	SM	Japan	Wakayama	174239	DRR212547	
	Tsuyuakane	*	Jap_SM4	SM	Japan	Ibaraki	-	DRR212548	
	85486	*	Tai_1	-	Taiwan	-	229937	DRR212549	
	Ellching	*	Tai_2	-	Taiwan	-	-	DRR212550	
	Hakufunbai	*	Tai_3	-	Taiwan	-	-	DRR212551	
	ST	*	Tai_4	-	Taiwan	-	-	DRR212552	
	Taiwan	*	Tai_5	-	Taiwan	-	174242	DRR212553	DRR212570
P. armeniaca	Harcot		Pa_1	-	Canada	-	174944	DRR212554	
	Heiwa	*	Pa_2	-	Japan	Nagano	174943	DRR212555	
	Niconicot		Pa_3	-	Japan	Ibaraki	-	DRR212556	
	Niigataomi		Pa_4	-	Japan	Niigata	174918	DRR212557	
	Ohisamacot		Pa_5	-	Japan	Ibaraki	-	DRR212558	
	Shingetsu		Pa_6	-	Japan	Nagano	-	DRR212559	
	Shinshuomi		Pa_7	-	Japan	Nagano	-	DRR212560	
P. salicina	Honeyrosa		Ps_1	-	Japan	Ibaraki	-	DRR212561	
	Oishiwase	*	Ps_2	-	Japan	Fukushima	112962	DRR212562	
	Soldum		Ps_3	-	USA	-	112977	DRR212563	
	Taiyo		Ps_4	-	Japan	Yamanashi	112982	DRR212564	
P. persica	Hakuho	*	Pp_1	-	Japan	Kanagawa	112532	DRR212565	

[1] Asterisks indicate the same materials used in chapter 2.

[2] Jap and Tai: Japanese and Taiwanese *P. mume* , respectively. Pa: *P. armeniaca* . Ps: *P. salicina* . Pp: *P. persica* .

[3] F, FS and O: Fruit, small-fruit and ornamental cultivars, respectively. AM and SM: putative hybrid cultivars between *P. mume* and *P.*

[4] JP numbers from Genbank of the National Institute of Agrofiological Sciences.

[5] Sequence Read Archive (SRA) accession numbers under DRA009691.

4.2.2 Target capture sequencing

Genomic DNA was extracted from young leaves using Nucleon PhytoPure (GE Healthcare, Chicago, IL, USA) and subjected to phenol/chloroform purification. The author employed a KAPA HyperPlus kit (Kapa Biosystems, Wilmington, MA, USA) to construct gDNA-seq libraries for an Illumina platform. Libraries were barcoded for each sample using single 8-bp NEXTflex adaptors (Bioo Scientific, Austin, TX, USA) and enriched by PCR using PrimeSTAR Max (Takara Bio, Shiga, Japan) with the following protocol: 3 min at 95℃, followed by eight cycles of 10 s at 95℃, 30 s at 65℃, 30 s at 72℃, and final extension for 5 min at 72℃.

To selectively retrieve libraries with exons, a myBaits Custom design kit was used to design 1–20-K probes (Arbor Biosciences, Ann Arbor, MI, USA), which uses biotinylated RNA probes to concentrate fragments carrying sequences of interest

(Gnirke et al., 2009) (Fig. 8), based on the published genomic and coding sequences of *P. mume* (Zhang et al., 2012). The author selected 29,621 non-redundant coding loci showing single hits with BLAST+ (MEGABLAST with -p 70 option) against the *P. mume* genome, for the subsequent bait designing. A 120-mer bait with 25–55 GC% per locus was randomly designed for each locus, and finally a bait set targeting 15,171 coding loci was obtained. An equal amount of constructed Illumina libraries (eight samples per tube) was pooled. Pooled libraries were purified by AMPure XP (Beckman Coulter, Indianapolis, IN, USA) and then electrophoresed on 1% agarose gel. A 300–700-bp area of DNA bands to re-extract libraries was cut out using a FastGene Gel/PCR extraction kit (NIPPON Genetics, Tokyo, Japan). Libraries were then subjected to target capture hybridization using myBaits Custom designed probes (Arbor Biosciences). Captured libraries were sequenced using the HiSeq 4000 platform (Illumina, San Diego, CA, USA) (paired-end 100 bp).

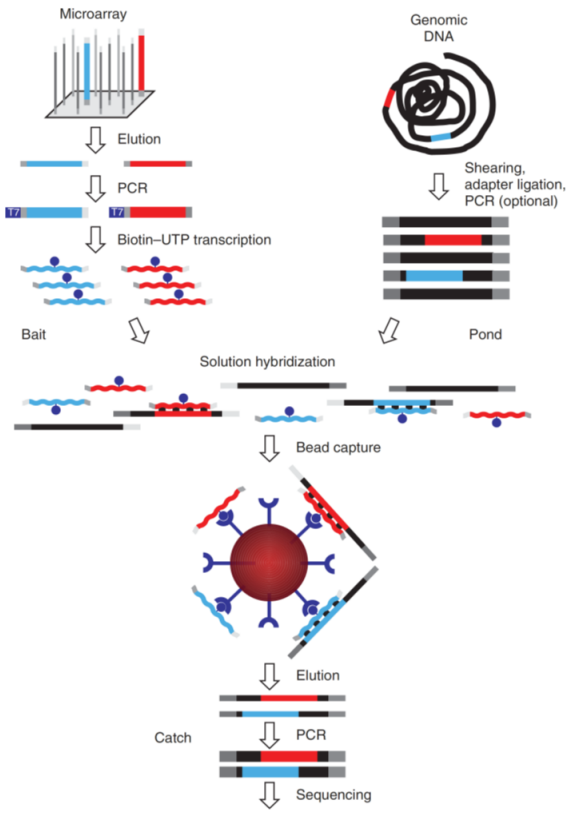

Fig. 8. Overview of target capture method by Gnirke et al. (2009).
Steps for the preparation of biotinylated RNA probes (wavy lines) (left), fragmented whole genome library (right) and selected and enriched target library (bottom) are illustrated. Bead capture allowed the selective retrieving of DNA fragments with targeted regions. Two independent targets for sequencing are indicated in red and blue. Adaptor sequences are shown in grey.

4.2.3 SNP calling

In addition to original sequencing data, published sequencing data were also used. Of the 348 *P. mume* cultivars in the whole-genome sequencing data reported by Zhang et al. (2018), 79 derived from China were selected and downloaded from Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) (Table 9). The data were selected to evenly contain all the P1 to P16 phylogenetic clusters reported by Zhang et al. (2018). Importantly, the clade P1 contains interspecific hybrids such as *P. mume* × *P. armeniaca* and *P. mume* × *P. salicina*. Raw reads were trimmed with no-demultiplex-allprep-8 (https://github.com/Comai-Lab/allprep) to select the reads with high quality (Phred score > 20 over a 5-bp window, length > 35-bp) and containing no 'N' and adapter sequences. The selected reads were mapped against LG1–8 of the peach (*P. persica*) v2.0 reference genome (Verde et al., 2017) using BWA-MEM with default parameters (Li, 2013). PCR duplicates were removed with OverAmp-3 (http://comailab.genomecenter.ucdavis.edu/index.php/Bwa-doall). SNPs were called using Samtools mpileup (Li et al., 2009) and VarScan2 mpileup2snp (Koboldt et al., 2009, 2012) with default settings (here, reffered to as "Primary_set"). From Primary_set, the author removed loci with >20% missing genotyping rate with PLINK (Purcell et al., 2007). Then, missing genotypes were imputed using Beagle 5.0 (Browning et al., 2018) with default settings ("Imputed_set"). The author also prepared "Cap_set" by removing the Chinese cultivars from Primary_set. Cap_set was subjected to sequencing depth estimation with VCFtools (Danecek et al., 2011). Imputed_set and Cap_set were used to estimate annotated genomic locations using SnpEff (Cingolani et al., 2012).

Table 9. Chinese cultivars (after Zhang *et al* ., 2018) used in this study.

Species	Name	ID in this study	Group[1]	Country	Location	SRA acc. no.
P. mume	Songchun	Chi_30	P1	China	Wuhan	SRX2369048
	Guifei	Chi_202	P1	China	Nanjing	SRX2369274
	Danhong Xingmei	Chi_270	P1	China	Taizhou	SRX2369295
	Qingming Wanfen	Chi_283	P1	China	Wuchang	SRX2369132
	Xiang Ruibai	Chi_396	P1	China	Beijing	SRX2369196
	Dalun Feimei	Chi_99	P2	China	Wuxi	SRX2369225
	Jinjin	Chi_142	P2	China	Wuhan	SRX2369078
	Wan Tiaozhi	Chi_178	P2	China	Nanjing	SRX2369264
	Huqiu Wanfen	Chi_200	P2	China	Suzhou	SRX2369278
	Yuanxia Xiaomei	Chi_250	P2	China	Wuxi	SRX2369116
	Danfen Chuizhi	Chi_14	P3	China	Nanjing	SRX2369345
	Xiantao Chuizhi	Chi_24	P3	China	Wuhan	SRX2374101
	Hubei Hanhong	Chi_49	P3	China	Wuhan	SRX2369355
	Bachong Hanhong	Chi_104	P3	China	Wuxi	SRX2369062
	Liu Banhong	Chi_181	P3	China	Wuhan	SRX2369095

Table 9. (Continued).

Species	Name	ID in this study	Group[1]	Country	Location	SRA acc. no.
	Laoshan Gongfen	Chi_55	P4	China	Shandong	SRX2369322
	Jinguang	Chi_100	P4	China	Wuxi	SRX2369309
	Micong Wanfen	Chi_218	P4	China	Nanjing	SRX2369283
	Nanjing Zaohong	Chi_219	P4	China	Nanjing	SRX2369284
	Yuheng Chuizhi	Chi_303	P4	China	Nanjing	SRX2369141
	Jinhong Chuizhi	Chi_22	P5	China	Wuhan	SRX2369348
	Danban Zhusha	Chi_96	P5	China	Nanjing	SRX2369307
	Xizhi Zhusha	Chi_192	P5	China	Wuhan	SRX2369099
	Luotian Xiaomozhusha	Chi_281	P5	China	Loutian	SRX2369130
	Yichong Tangmei	Chi_300	P5	China	Wuxi	SRX2369302
	Danzhuang Gongfen	Chi_113	P6	China	Nanjing	SRX2369232
	Jiang Nan	Chi_132	P6	China	Wuhan	SRX2369245
	Zao Ningxin	Chi_133	P6	China	Wuhan	SRX2369075
	Honghua Wantiao	Chi_269	P6	China	Nanjing	SRX2369291
	Long Youmei	Chi_31	P7	China	Nanjing	SRX2369049
	Jian Jingmei	Chi_65	P7	China	Nanjing	SRX2369056
	Fenbai Gongfen	Chi_129	P7	China	Nanjing	SRX2369244
	Danban Zaolve	Chi_164	P7	China	Wuhan	SRX2369088
	Mi Danlv	Chi_166	P7	China	Suzhou	SRX2369089
	Danhong Chuizhi	Chi_5	P8	China	Wuhan	SRX2369336
	Danfen Chuizhi	Chi_8	P8	China	Wuhan	SRX2369334
	Hanfen Chuizhi	Chi_17	P8	China	Wuhan	SRX2369343
	Moshan Gongfen	Chi_153	P8	China	Wuhan	SRX2369082
	Shui Zhusha	Chi_359	P8	China	Wuhan	SRX2369182
	Dayun Zhaoshui	Chi_48	P9	China	Wuhan	SRX2369356
	Fenxia	Chi_128	P9	China	Wuhan	SRX2369072
	Danlun Zhusha	Chi_265	P9	China	Wuhan	SRX2369292
	Moshan Shuizhusha	Chi_311	P9	China	Wuhan	SRX2369145
	Jingdezheng Yemei	Chi_418	P9	China	Jingdezhen	SRX2369209
	Wanhua Gongfen	Chi_103	P10	China	Wuhan	SRX2369061
	Caishan Gongfen	Chi_147	P10	China	Huangmei	SRX2369080
	Xuemei	Chi_187	P10	China	Wuhan	SRX2369097
	Zaohua	Chi_245	P10	China	Nanjing	SRX2369115
	Kaidi	Chi_246	P10	China	Nanjing	SRX2369288
	Danyun	Chi_68	P11	China	Wuhan	SRX2369366
	Qianye Gongfen	Chi_109	P11	China	Wuhan	SRX2369230
	Xiao Lve	Chi_165	P11	China	Wuhan	SRX2369255
	Midan Tiaozhi	Chi_176	P11	China	Wuhan	SRX2369313
	Xiao Yudie	Chi_360	P11	China	Wuhan	SRX2369183
	Lianhu Gongfen	Chi_36	P12	China	Wuhan	SRX2369388
	Xuehai Gongfen	Chi_57	P12	China	Wuhan	SRX2369376
	Xueyu	Chi_62	P12	China	Wuhan	SRX2369385
	Fenkou	Chi_102	P12	China	Sichuan	SRX2369226
	Quanzhou Xiaofen	Chi_140	P12	China	Quanzhou	SRX2369247
	Chaotang Gongfen	Chi_41	P13	China	Wuhan	SRX2369387
	Jiangsha Gongfen	Chi_105	P13	China	Wuhan	SRX2369063
	Shaoan Guomei2	Chi_402	P13	China	Shaoan	SRX2369198
	Mianning Guomei	Chi_413	P13	China	Mianning	SRX2369206
	Huangshan Yemei2	Chi_415	P13	China	Huangshan	SRX2369208

Table 9. (Continued).

Species	Name	ID in this study	Group[1]	Country	Location	SRA acc. no.
	Honhfen Taige	Chi_72	P14	China	Chongqing	SRX2369363
	Laoren Meidahong	Chi_116	P14	China	Sichuan	SRX2369066
	Gongchun	Chi_117	P14	China	Chengdu	SRX2369239
	Fenghong	Chi_135	P14	China	Nanjing	SRX2369076
	Xiaohong Changxu	Chi_158	P14	China	Wuhan	SRX2369256
	Dayu Zhaoshui	Chi_44	P15	China	Wuhan	SRX2369350
	Danyun Zhusha	Chi_95	P15	China	Wuhan	SRX2369223
	Bianban Dahong	Chi_130	P15	China	Wuhan	SRX2369073
	Shuihong Changsi	Chi_206	P15	China	Nanjing	SRX2369317
	Qingxin	Chi_313	P15	China	Kunming	SRX2369146
	Wan Lve	Chi_229	P16	China	Kunming	SRX2369287
	Nandaping Gumei2	Chi_332	P16	China	Eryuan	SRX2369160
	Duan Ruifen	Chi_353	P16	China	Kunming	SRX2369178
	Xiaoxi Meifen	Chi_358	P16	China	Kunming	SRX2369181
	Xiaoxi Baimei	Chi_373	P16	China	Lijiang	SRX2369187

[1] Classified by Zhang *et al* . (2018).

4.2.4 Population structure analysis

SNP sets for population structure analyses were prepared based on Imputed_set. The author first extracted cultivars of interest and removed loci with a minor allele frequency (MAF) <0.03 and that violated the Hardy–Weinberg equilibrium (P <0.0001) using PLINK. The SNPs with high (r^2 >0.5) linkage disequilibrium (LD) within a 50-SNP window with 3 SNPs shifting were further pruned using PLINK (--indep 50 3 2).

Population structure was estimated using three methods: principal component analysis (PCA), Bayesian clustering, and maximum likelihood phylogenetic analysis. PCA was performed using smartpca of EIGENSOFT (Patterson et al., 2006). ADMIXTURE (Alexander et al., 2009) was used for Bayesian clustering. The author assumed K = 2–10, and 10 simulations were carried out for each K value. The author then compiled the results of 10 simulations for each K using CLUMPP (Jakobsson and Rosenberg, 2007). The most optimal K was estimated based on cross-validation error (CVE) values calculated according to the ADMIXTURE manual. In the present study, the most optimal K was estimated to be four (CVE = 0.328). A maximum likelihood (ML) phylogenetic tree was constructed using SNPhylo (Lee et al., 2014) with 1,000 bootstrap replications.

For detection of linkage disequilibrium, based on Imputed_set, the author removed samples Chi_30, 202, 270, 283, 396, Jap_AM1–8, and SM1–4, which were previously considered to be interspecific hybrids in Chapter 2 and previous studies (Hayashi et al., 2008; Zhang et al., 2018), and additionally Chi_250 and Jap_O2, which

were newly classified as “Admixed” in this chapter. Pairwise LD was computed using PopLDdecay (Zhang et al., 2019) with -MaxDist 10000, -MAF 0.03, and -Het 0.75 options. The Plot_Multipop function was then used to calculate moving averages of LD for each 10-kb bin.

For detection of genetic differentiation and identity by descent (IBD), the same SNP set as that used for the above population structure analyses was used. Pairwise Weir and Cockerham weighted F_{ST} was calculated using VCFtools. IBD was estimated using pairwise pi-hat values from PLINK.

4.3 Results and discussion

4.3.1 Efficacy of targeted resequencing in Prunus cultivars

From the targeted resequencing of 129 *Prunus* cultivars (117 *P. mume*, 7 *P. armeniaca*, 4 *P. salicina*, and 1 *P. persica*), 1,096,007,397 of a total 1,177,780,940 reads (93.1%) (deposited at DRA009691; Table 8) were mapped onto LG1–8 of the peach v2.0 genome, including 402,859,421 uniquely mapped reads (34.2%). In Cap_set, the author could identify a total of 489,420 SNPs with an average depth of 29.8×, of which each cultivar ranged from 15.2× to 60.2×. SnpEff analysis revealed that total 94.4% of SNPs were located on a genic region (35.3%) or its upstream and downstream regions (26.8% and 32.3%, respectively) (Table 10). In Imputed_set, the author obtained a total of 148,953 SNPs, of which the SnpEff result was almost consistent with that of Cap_set (Table 10). Thus, the author could successfully and cost-effectively obtain SNPs based on exon capture in subgen. *Prunus* cultivars.

Table 10. Annotated genomic locations of SNPs derived from targeted resequencing.

	SNPs (%)	
Annotated region	Cap_set[1]	Imputed_set[1]
Upstream (~5 kb)	26.8	26.1
5'-UTR[2]	1.5	1.1
Exon[2]	16.2	20.3
Intron[2]	14.4	13.1
Splice site[2]	0.8	0.9
3'-UTR[2]	2.4	2.2
Downstream (~5 kb)	32.3	31.9
Intergenic	5.6	4.5

[1] Cap_set: not filtered set without Chinese cultivars.
Imputed_set: filtered with missing loci < 0.2 and imputed with Beagle 5.0.
[2] A part of genic region.

4.3.2 Definition of population structure

A total of 14,310 selected SNPs were used for PCA, ADMIXTURE and ML phylogenetic analyses to reveal the population structure among 208 *Prunus* cultivars (79 Chinese cultivars were added to the 129 cultivars). In all three analyses, Japanese apricot (*P. mume*), apricot (*P. armeniaca*), and Japanese plum (*P. salicina*) were clearly found to form species specific clusters (Figs. 9, 10). This was also supported by the F_{ST} values among the species (Table 11). Hypothetical interspecific hybrids of *P. mume* (Jap_AM1–8, SM1–4, Chi_30, 202, 270, and 396) in Chapter 2 and other studies (Hayashi et al., 2008; Zhang et al., 2018) were positioned between *P. mume* and the other *Prunus* species, supporting that they are "Admixed" individuals (Figs. 9, 10). Importantly, Chinese and Japanese cultivars of *P. mume* were clustered into separate groups, whereas Taiwanese cultivars were clustered with Japanese cultivars (Fig. 9a, b). The ML tree (Fig. 10) also supported that the *P. mume* cluster was largely divided into Chinese and Japanese (with Taiwanese) clades, with statistical support (bootstrap > 60 in ML), including minor exceptions (Jap_O6, 9, 14, 18, 27, 42, 44, and 45; green stars in Fig. 10). In the IBD analysis, some pairs of Japanese and Taiwanese cultivars were inferred to be in first- or second-degree relationships (pi-hat = 0.25–0.5; Fig. S2). Conversely, some pairs of Chinese and Japanese cultivars showed first-degree relationships (pi-hat = 0.5), but most combinations were genetically distinct (Fig. 11). The highest F_{ST} value was observed between Chinese and Taiwanese cultivar groups, in contrast to the geographical proximity (Table 11). These results differ from conventional (or empirical) observations, which have indicated that Japanese cultivars of *P. mume* were originally introduced from China to Japan relatively recently (ca. 2,000 years ago) via human activities (Horiuchi et al., 1996). Based on the ML tree, Japanese cultivars showed weak differentiation depending on their characteristics or applications by human (Fig. 10). Although subpopulations of fruit, small-fruit and ornamental cultivars was not clearly divided in PCA, ADMIXTURE and F_{ST} (Figs. 12, 9b, Table 12), in the ML tree, the majority of fruit (36 of 45 cultivars), small-fruit (9 of 10 cultivars) and ornamental (25 of 45 cultivars) cultivars belonged to the same cluster (Fig. 10). This supports the possibility that human preference triggered a recent differentiation of Japanese population from the same genetic resources as described in Chapter 2 and previous reports (Horiuchi et al., 1996; Hayashi et al., 2008).

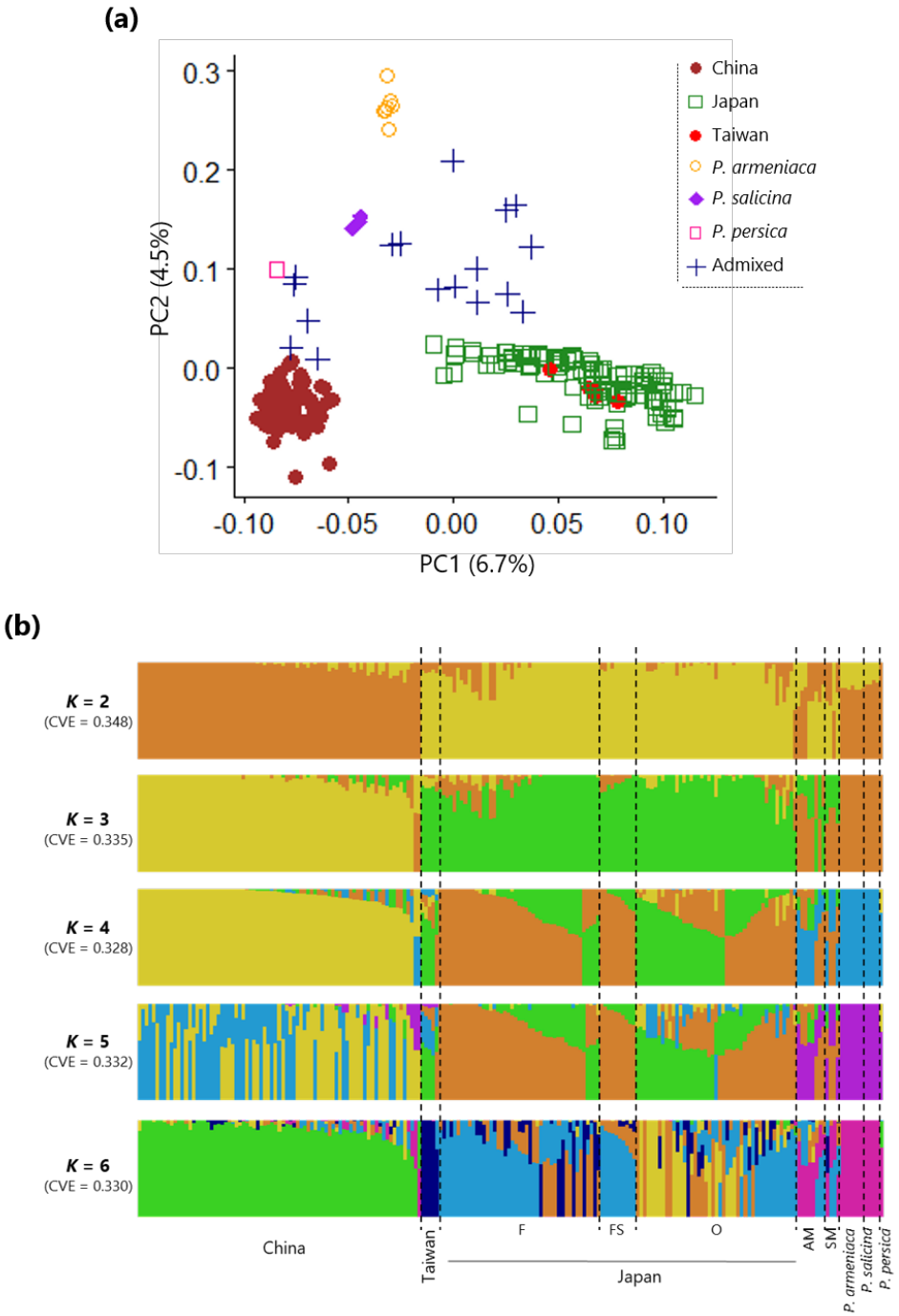

Fig. 9. Population structure analysis in *Prunus mume*. (a) Principal component analysis (PCA) of all the 208 *Prunus* cultivars. (b) Proportion of ancestry for all the 208 *Prunus* cultivars from $K = 2$–6 inferred with ADMIXTURE. Proportions of the membership to each cluster are shown with the lengths of the colored bar (y-axis). CVE: cross validation error. F: fruit cultivars, FS: small-fruit cultivars, O: ornamental cultivars, AM: putative hybrids with *P. armeniaca*, SM: putative hybrids with *P. salicina*.

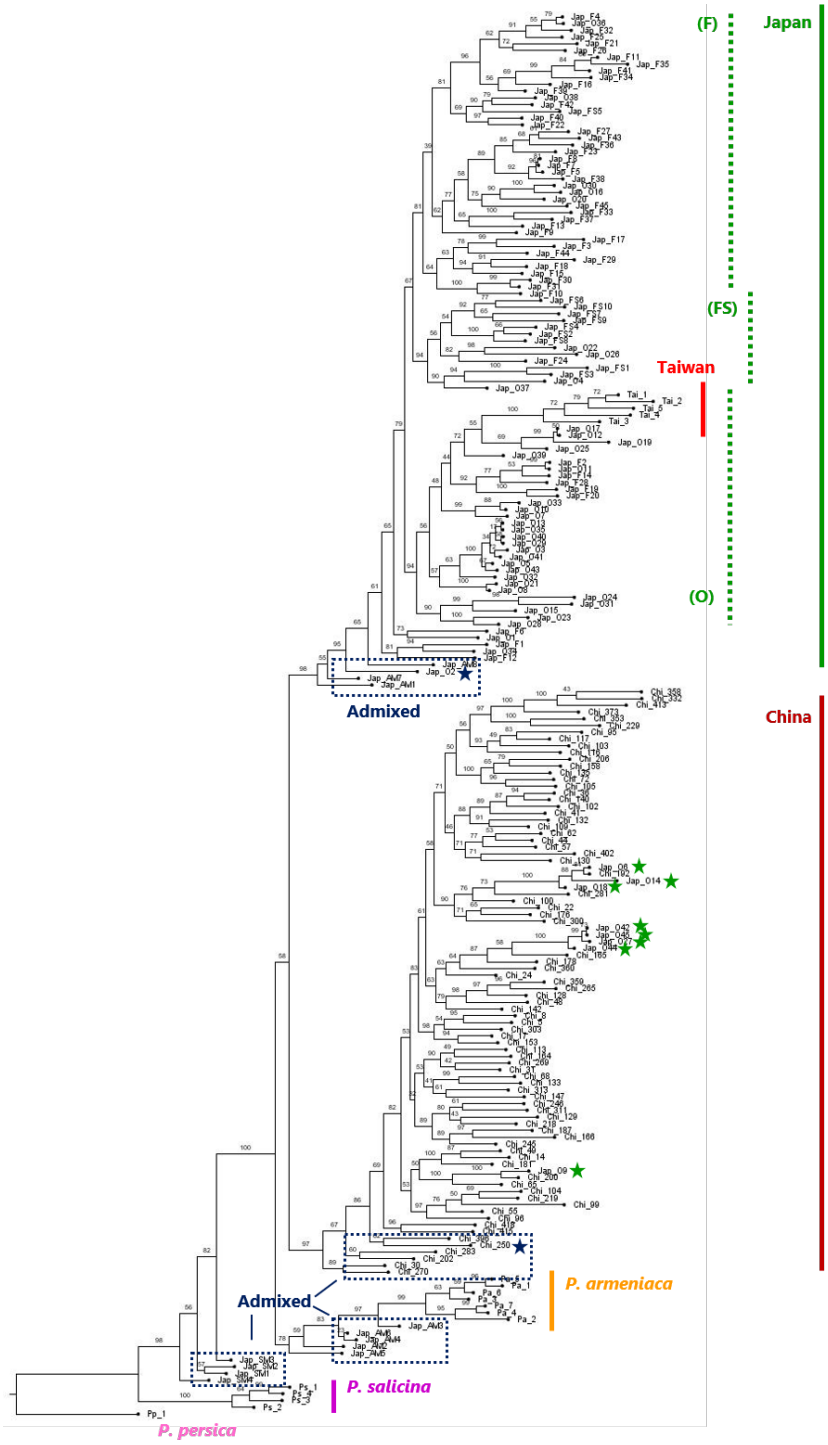

Fig. 10. Maximum likelihood phylogenetic tree inferred with all 208 *Prunus* cultivars. Values beside the nodes indicate bootstrap values generated with 1000 replications. *P. mume* cultivars clustered with admixed cultivars are shown with navy stars. Dotted lines within Japanese clusters indicate the characteristic clusters for fruit (F), small-fruit (FS) and ornamental (O) cultivars. Green stars indicate Japanese cultivars in Chinese clusters. Chi: Chinese cultivars, Jap: Japanese cultivars, Tai: Taiwanese cultivars, F: fruit cultivars, FS: small-fruit cultivars, O: ornamental cultivars, AM: putative hybrids with *P. armeniaca*, SM: putative hybrids with *P. salicina*, Pa: *P. armeniaca*, Ps: *P. salicina* and Pp: *P. persica*.

Table 11. Pairwise F_{ST} among *Prunus* cultivars.

	China[1]	Japan[1]	Taiwan[1]	*P. armeniaca*
China				
Japan	0.075			
Taiwan	0.151	0.076		
P. armeniaca	0.203	0.169	0.270	
P. salicina	0.207	0.181	0.301	0.324

[1] China, Japan and Taiwan: Chinese, Japanese and Taiwanese *P. mume* cultivars, respectively.

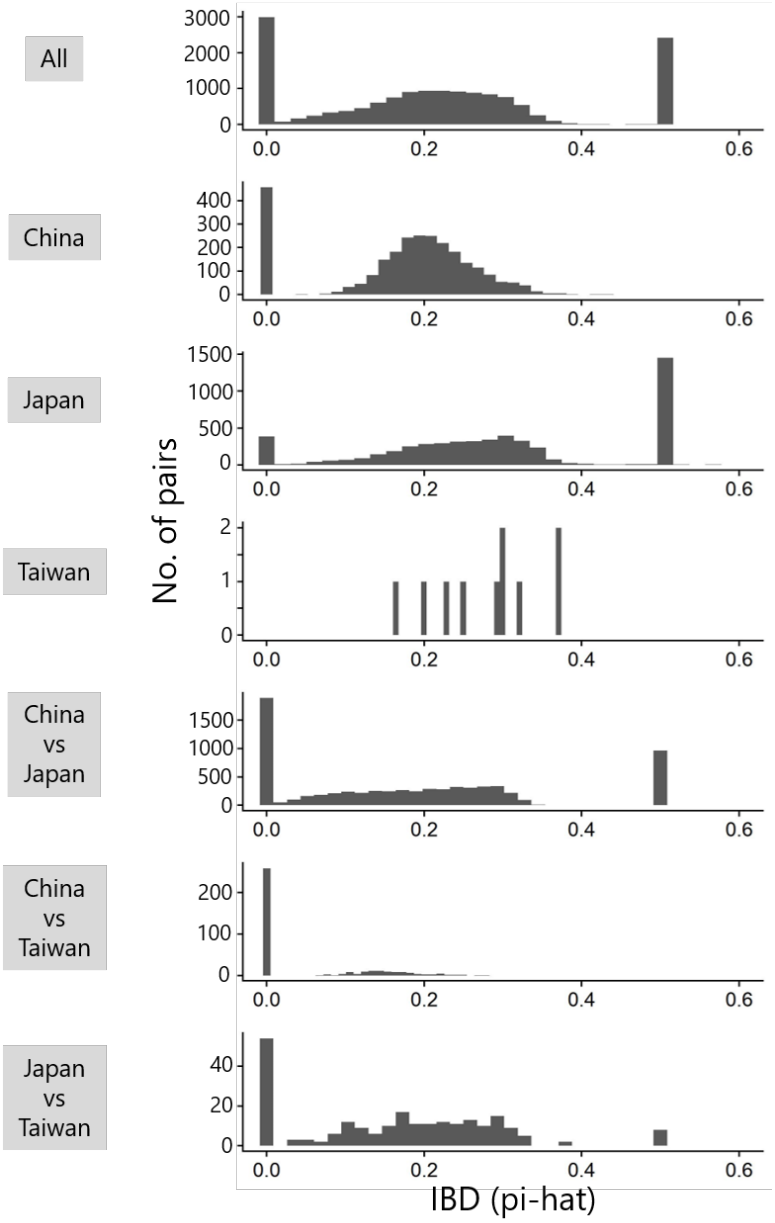

Fig. 11. Pairwise identity by descent (IBD) proportions in *Prunus mume* cultivars.

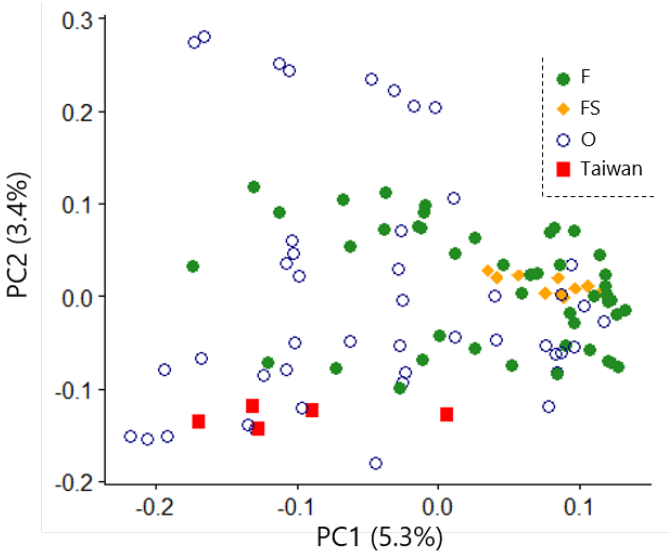

Fig. 12. Principal component analysis (PCA) of Japanese and Taiwanese cultivars of *Prunus mume*. Percentage of the total variation explained by the PC is shown in parentheses. F: fruit cultivars, FS: small-fruit cultivars, O: ornamental cultivars.

Table 12. Pairwise F_{ST} among Japanese *Prunus mume* cultivars.

	F[1]	FS[1]	O[1]	*P. armeniaca*
F				
FS	0.026			
O	0.020	0.034		
P. armeniaca	0.185	0.240	0.169	
P. salicina	0.197	0.254	0.184	0.331

[1] F, FS and O: Fruit, small-fruit and ornamental groups in Japanese *P. mume* cultivars, respectively.

LD mostly decayed within ca. 100 kb in all the *P. mume* groups surveyed in this chapter (Fig. 13), which is much longer than in *P. armeniaca* but shorter than in *P. persica* (Akagi et al., 2016; Mariette et al., 2016; Yu et al., 2018). The extent of LD was slightly different among cultivar groups. For example, LD decayed slower in Japanese cultivars than in the others (Fig. 13a). Within the Japanese cultivars, ornamental cultivars exhibited further slower LD decay (Fig. 13b), presumably due to their narrow genetic resources and frequent utilization of bud-sport for development of new cultivars especially after the Edo Period (Mega et al., 1988; Horiuchi et al., 1996).

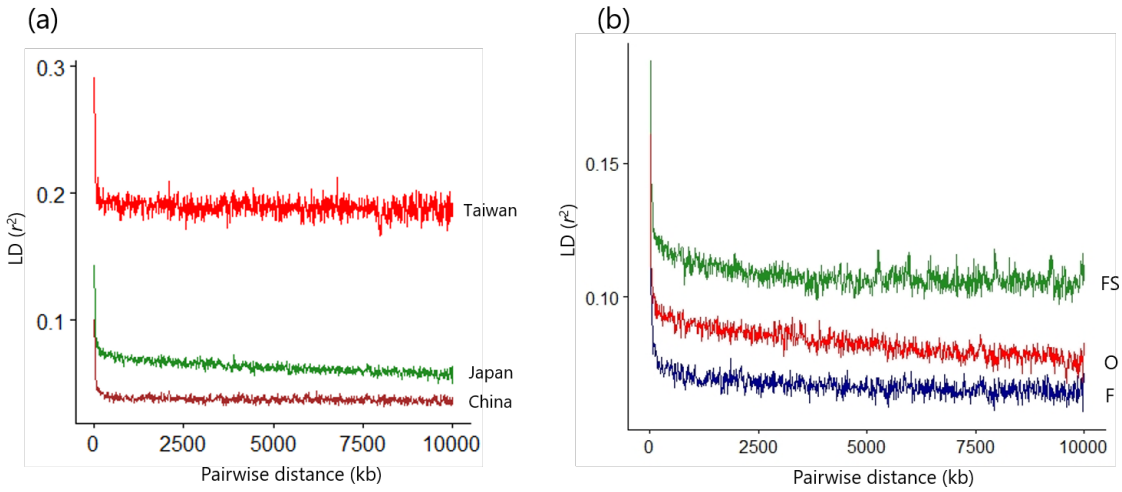

Fig. 13. Patterns of linkage disequilibrium decay in (a) Chinese, Japanese and Taiwanese cultivars and (b) Japanese fruit (F), small-fruit (FS) and ornamental (O) cultivars of *Prunus mume*.

4.3.3 Completed and ongoing population differentiation in P. mume

In this chapter, the author revealed that Chinese and Japanese cultivars of *P. mume* showed distinct population differentiation (Figs. 9, 10). Conversely, Taiwanese cultivars belonged to the Japanese clade but clustered independently from the Japanese cultivars, consistent with the results of Chapter 2 and the previous study (Hayashi et al., 2008). These results suggest that the differentiation of Chinese and Japanese populations predated that of the Taiwanese population, which is inconsistent with the conventional belief that *P. mume* cultivars were derived from Chinese ones, and recently (ca. 2,000 years ago), people have been introduced to other regions (Mega et al., 1988; Faust et al., 2011). There is a record describing wild *P. mume* accessions collected from western Japan that have been clonally maintained (https://agriknowledge.affrc.go.jp/RN/3030041889; in Japanese). Those samples may have important genetic information related to the origin of Japanese cultivars. The population structure of domesticated crops may be directly investigated using ancient DNA (aDNA) extracted from ancient plant remains (Frantz et al., 2016; Mascher et al., 2016; Kistler et al., 2018; Allaby et al., 2019; Narasimhan et al., 2019; Smith et al., 2019). Thus, collaboration between geneticists and archaeologists will make rapid progress in studies on crop domestication. In Japanese population of *P. mume*, fruit, small-fruit, and ornamental cultivars tended to form their subgroups, suggesting that the Japanese cultivars are differentiating based on human preference. These results shed

light on our understanding of the current population structure in *P. mume*.

4.4 Summary

Japanese apricot (*Prunus mume*) is believed to have been domesticated firstly in China, and then have been moved into Japan, originally for ornamental purposes. Currently, the cultivars are widely diversified mainly based on their usages. However, in contrast to these historical implications and conventional categorizations, the genetic background of this species is little known, and the previous studies using limited number of PCR-based DNA markers are insufficient to reveal the current population structure in *P. mume*. Therefore, in this chapter, the author conducted the exon-targeted resequencing of 129 genomes in the subgenus *Prunus*, Japanese apricot (*P. mume*), apricot (*P. armeniaca*), Japanese plum (*P. salicina*) and peach (*P. persica*). The data were merged with published resequencing data of 79 Chinese *P. mume* cultivars. Principal component analysis, ADMIXTURE and maximum likelihood phylogenetic analysis indicated that Japanese apricot (*P. mume*), apricot (*P. armeniaca*), and Japanese plum (*P. salicina*) form distinct clusters. Importantly, Chinese and Japanese cultivars of *P. mume* were clustered into separate groups, whereas Taiwanese cultivars were clustered with Japanese cultivars. In addition, most of the fruit (36 of 45 cultivars), small-fruit (9 of 10 cultivars) and ornamental (25 of 45 cultivars) cultivars belonged to the same phylogenetic cluster. This would support the possibility that the differentiation of Chinese and Japanese populations predated that of the Taiwanese population, and subsequently human preference triggered a recent differentiation of Japanese population from the same genetic resources. These results shed light on our understanding of the current population structure in *P. mume*.

Chapter 5. Genomic signatures for natural selection and interspecific introgression in East Asian cultivars of Japanese apricot

5.1 Introduction

Domestication and population differentiation often involve considerable phenotypic changes in crops (Diamond, 2002; Purugganan and Fuller, 2009; Zeder, 2015). Mainly natural mutations are thought to drive these changes, while occasional interspecific introgression can also potentially contribute (Baack and Rieseberg, 2007; Harrison and Larson, 2014; Gaut et al., 2015; Suarez-Gonzalez et al., 2018). For instance, domesticated apple (*Malus domestica*), which originated from a wild apple species distributed in Central Asia (*M. sieversiii*), has recently experienced additional genomic introgression from another wild species (*M. sylvestris*) (Cornille et al., 2012). Citrus and olive are also suggested to have complicated evolutionary pathways involving interspecific introgression from related or ancestral wild species (Wu et al., 2014, 2018; Diez et al., 2015). Hexaploid bread wheat (*Triticum aestivum*) is a notable example of herbaceous crops with a drastic domestication process. This species was developed from a dynamic hybridization between tetraploid emmer wheat (*T. turgidum*) and diploid Tausch's goatgrass (*Aegilops tauschii*) and subsequent introgression from other species promoted cultivar differentiation (Molnár-Láng et al., 2015; He et al., 2019). Owing to the evolutional importance of interspecific introgression, much previous research on a variety of species has inferred the presence of interspecific introgression in current populations. Notwithstanding, few studies have further estimated the genome-wide distribution of introgressed fragments and their importance for domestication/population differentiation events.

Introduced mutations favorable to environmental adaptation or human preference might be subjected to natural or artificial selection. When a particular locus experiences a strong selection pressure, the genetic diversity of adjacent genomic regions is reduced as well as the targeted locus itself, which is known as a "selective sweep" (Stephan, 2019). Therefore, the genetic factors playing important roles in the formation of current populations can be estimated by characterizing genome-wide selective sweep profiles (Clark et al., 2004; Sabeti et al., 2007; Kosova et al., 2010; Ishii et al., 2013; Akagi et al., 2016; Lee et al., 2016; Pankin et al., 2018; Nadachowska-Brzyska et al., 2019). Selective sweep profiles have been well studied especially in annual crops such as rice (*Oryza sativa*). In annual crops, where selected alleles are thought to be fixed in homozygous state, patterns of selective sweeps have been identified mainly using site

frequency spectrum (SFS)-based methods using the reduction in genetic diversity as the index. Conversely, perennial crops (or tree crops) have more complicated genomic/genetic conditions, mainly due to vegetative propagation, frequent outcrossing, and long generation time. Therefore, a selected allele in perennial crops is expected to be maintained in a heterozygous manner, which would be quite similar to animal (including human) genomes, requiring haplotype-based detection of selective sweeps (Voight et al., 2006; Sabeti et al., 2007).

The genus *Prunus* includes a wide variety of major tree crops consumed worldwide, such as peach (*P. persica*), sweet cherry (*P. avium*), plum (*P. salicina*), apricot (*P. armeniaca*), and almond (*P. dulcis*). Japanese apricot is morphologically similar to apricot and plum, and they are all nested in the subgenus *Prunus* (Bortiri et al., 2001). Species of the subgen. *Prunus* are partially compatible for interspecific crossing (Yamaguchi et al., 2018; Morimoto et al., 2019). Therefore, recent breeding programs often utilize interspecific crossing of the subgen. *Prunus*, such as "Sumomo-ume" (*P. salicina* × *P. mume*), or "Pluot" (*P. salicina* × *P. armeniaca*) (Kyotani et al., 1988; Brantley, 2004; Yaegaki et al., 2012).

Given the above information, it would appear that Japanese apricot and related species in the subgen. *Prunus*, have undergone complicated evolutionary processes, involving natural or artificial selections and potential introgressions among them. Here, to clarify the evolutionary paths to establish the current subgen. *Prunus*, mainly for *P. mume*, the author analyzed genome-wide single nucleotide polymorphisms (SNPs) based on targeted resequencing of ca. 15,000 exons in East Asian *P. mume* cultivars obtained in Chapter 4. In this chapter, an integrative analysis of selective sweeps based not only on SFS but on extended haplotype homozygosity (EHH) were performed to infer the natural or artificial selections involved in the establishment or population differentiation of *P. mume* (section 5.2). In addition, the transition of fragmental genetic structures was assessed to detect regions of interspecific introgressions (section 5.3). The results successfully inferred the importance of interspecific introgressions and lineage-specific selections during the evolution of Japanese apricot.

5.2 Detection of selective sweeps potentially involved in the establishment or population differentiation of Japanese apricot

5.2.1 Materials and methods

5.2.1.1 Plant materials and datasets

A total of 196 *P. mume* cultivars (79 Chinese, 112 Japanese and 5 Taiwanese cultivars) were used to detect selective sweeps. Chi_30, 202, 270, 283, 396, Jap_AM1–8, and SM1–4, which were previously considered to be interspecific hybrids (Hayashi et al., 2008; Zhang et al., 2018), and additionally Chi_250 and Jap_O2, which were newly classified as "Admixed" in Chapter 4 were excluded from the subsequent analyses. SNP sets were prepared based on Imputed_set developed in Chapter 4. The author first extracted cultivars of interest and removed loci with a minor allele frequency (MAF) <0.03 using PLINK (Purcell et al., 2007).

5.2.1.2 Identification of selective sweeps

To estimate genomic regions that experienced natural or artificial selection, the author used the following methods: composite likelihood ratio (CLR) (Nielsen et al., 2005), nSL (Ferrer-Admetlla et al., 2014), and XP-EHH (Sabeti et al., 2007) tests. Of these, the CLR test is a method based on SFS, which detects deviations in allele frequency from neutrality at each site. This assumes that a selected allele is fixed in a population (as in annual crops). Conversely, nSL and XP-EHH analyses are based on extended haplotype homozygosity (EHH), which detects elongated linkage disequilibrium blocks around a selected core allele. EHH-based methods work well if a selected allele is not completely fixed in a population but is maintained in a heterozygous state (as in trees, humans, and other animals) (Sabeti et al., 2007; Kosova et al., 2010; Akagi et al., 2016; Lee et al., 2016; Nadachowska-Brzyska et al., 2019). XP-EHH compares EHH values between paired populations (e.g., ancestral and derived populations) and can detect selective sweeps related to population differentiation.

SweeD (Pavlidis et al., 2013) (with -grid 500 flag) was used for calculation of CLR values. Neutral thresholds were determined according to Nielsen et al. (2005). A thousand simulated neutral genotype datasets were first generated using ms (Hudson, 2002), based on the observed number of polymorphic sites (S) and sample size (n) of each subpopulation. Next, we ran SweeD with -grid 500 flag using simulated genotype sets to obtain neutral CLR values. Neutral thresholds were determined as 99% percentile values for each subpopulation. Selscan v 1.2.0a (Szpiech and Hernandez, 2014) was used to perform nSL and XP-EHH analyses. The author assumed that genetic position was equal to physical position in the XP-EHH analysis. An SNP dataset was generated based on Imputed_set. Genotype sets for each subpopulation (China, Japan, Taiwan, ornamental, fruit, and small-fruit) were first extracted. Loci with MAF <0.03 were filtered for each

dataset, and subsequently, haplotypes were phased using Beagle 5.0. Unstandardized nSL and XP-EHH values were Z-scored using the norm function of selscan and transformed to *P* values.

5.2.2 Results and discussion

5.2.2.1 Identification of selective sweeps related to population differentiation in P. mume

Alleles that have undergone positive selections showed i) reduction in genetic diversity, and ii) extension of haploblock, in adjacent genetic regions and a targeted locus itself (Stephan, 2019). An SFS-based method, SweeD was used to detect i), whereas extended haplotype homozygosity (EHH)-based method, nSL and XP-EHH was used to identify ii).

For Chinese and Japanese populations, the results for SweeD and single-population EHH-based nSL analyses were different (Figs. 14a, 15, Table S5). Especially in the Chinese population, no SweeD peak exceeded the neutral threshold (Fig. 15). Tree crops have specific characters, such as long generation time and frequent vegetative propagation, suggesting that selected alleles may have not been completely fixed, like in humans (Voight et al., 2006; Akagi et al., 2016). Sites with only SFS-based peaks (with no EHH-based peaks) are thus suspected to be derived from the occasional reduction of nucleotide diversity, namely, the substitution ratio in the *P. mume* genome, distortion of the availability in SNPs, or simple drift (Akagi et al., 2016). Therefore, in *P. mume*, the author mainly focused on the results of the EHH-based analyses. Only in the Taiwanese cultivar group was the pattern of significant peaks of SweeD similar to that of nSL (Figs. 14a, 15, Table S5), suggesting that Taiwanese cultivars may have experienced stronger selection pressure than the other cultivar groups.

In the nSL analysis, a strong peak ($P < 1e^{-4}$) on chromosome 8 (ca. 6.7–6.8 Mb), which was common in all Chinese, Japanese, and Taiwanese cultivar groups of *P. mume* was identified (Fig. 14a). A peak on chromosome 6 was also common in three geographic groups, although the Japanese peak was not significant. Strong peaks common in Chinese and Japanese cultivars were observed on chromosome 2. Geographically specific peaks were found on chromosomes 2 and 8 in Chinese, chromosomes 3 and 4 in Japanese, and chromosomes 1, 2, 4, 5, 7 and 8 in Taiwanese populations. These results suggest that the common ancestor of *P. mume* underwent certain positive selection, such as on chromosomes 6 and 8, and thereafter established the three populations based on geographical separation and independent selection.

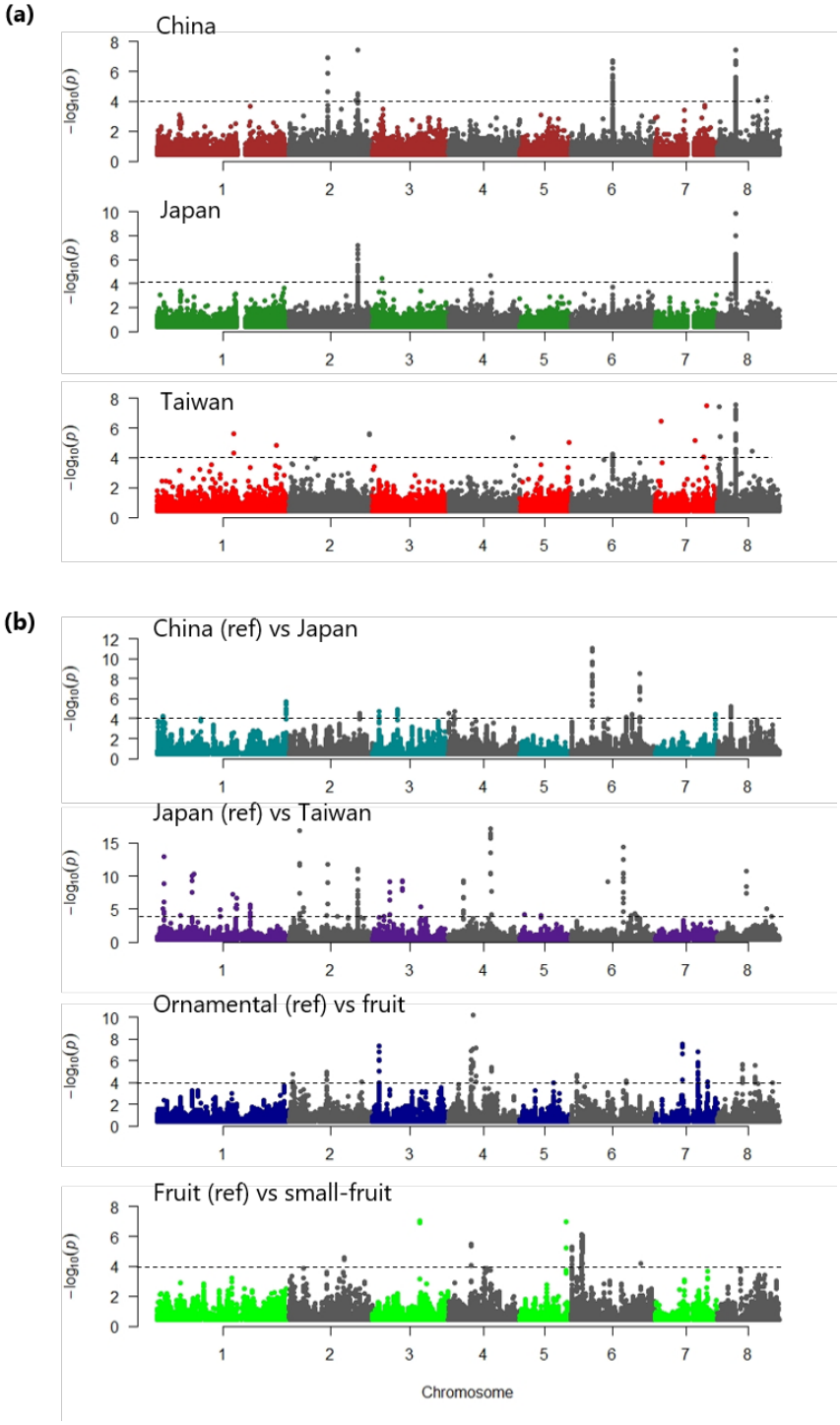

Fig. 14. Identification of selective sweeps in Chinese, Japanese, and Taiwanese cultivars of *Prunus mume* based on, (a) single-population nSL and (b) dual-population XP-EHH analyses. *P* values calculated with normalized nSL values are shown in $-\log_{10}$ scale. Detailed information of SNPs with $-\log_{10}P > 4$ (dotted line) is summarized in Tables S5 (nSL) and S6 (XP-EHH).

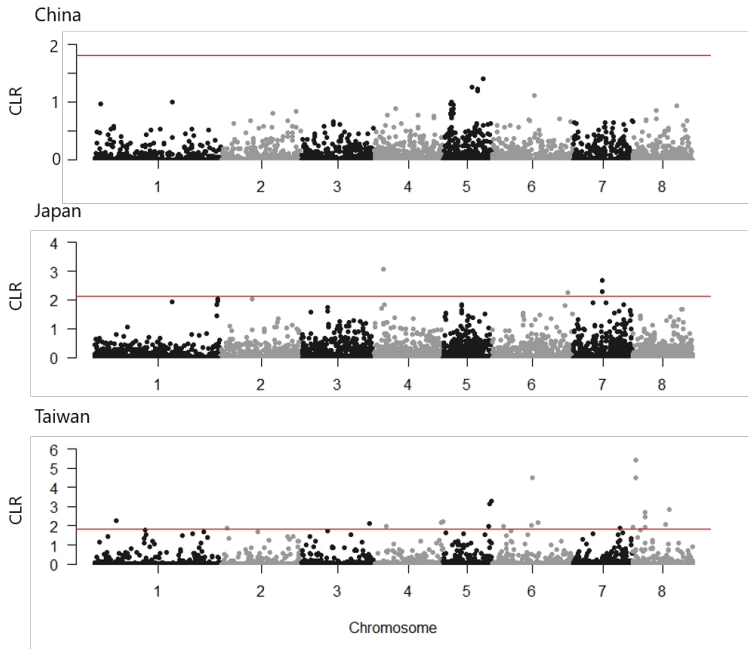

Fig. 15. Identification of selective sweeps in Chinese, Japanese and Taiwanese cultivars of *Prunus mume* based on site frequency spectrum (SFS)-based SweeD (CLR) analysis. Red bar indicates neutral threshold.

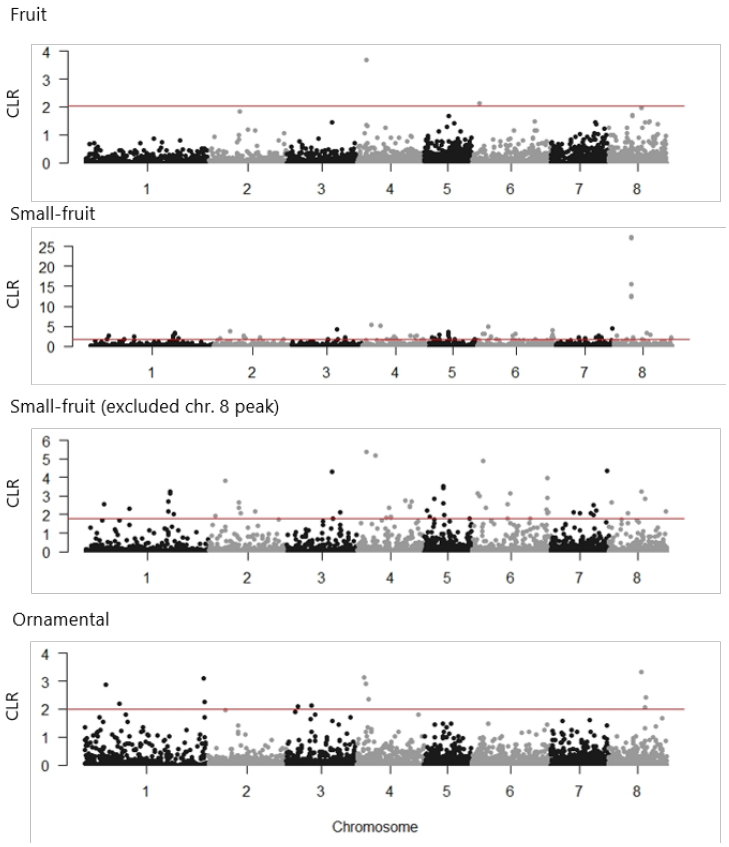

Fig. 16. See next page for caption.

Fig. 16. Identification of selective sweeps in fruit, small-fruit and ornamental cultivars of *Prunus mume* based on site frequency spectrum (SFS)-based SweeD (CLR) analysis. In small-fruit cultivars, Manhattan plot was redrawn with extremely high peaks in chromosome 8 removed. Red bar indicates neutral threshold.

Furthermore, the author conducted two population-based XP-EHH analyses in (i) the Chinese cultivars (reference) vs. the Japanese cultivars (derived), (ii) the Japanese cultivars (reference) vs. the Taiwanese cultivars (derived), (iii) the ornamental cultivars (reference) vs. the fruit cultivars (derived), and (iv) the fruit cultivars (reference) vs. the small-fruit cultivars (derived) (Fig. 14b). The XP-EHH can focus on the selected alleles that are highly differentiated between populations (Sabeti et al., 2007). Here, negative and positive normalized XP-EHH values indicated that an extended haploblock was observed in the reference and derived populations, respectively (Szpiech and Hernandez, 2014). Most of the strong peaks ($P < 1e^{-4}$; Table S6) were not overlapped with those detected in the nSL analysis (Fig. 14b). In the Chinese vs Japanese analysis, significant peaks were observed on all chromosomes except for chromosome 5 (Fig. 14b). The genomic positions of these peaks were different from those in SweeD (Fig. 15, Table S6), indicating that they have not yet perfectly fixed in the population. Most strong XP-EHH peaks in Japanese vs. Taiwanese groups showed negative values, except for ca. 33.9 Mb of chromosome 1 (Table S6), indicating that Japanese cultivars underwent much more extensive selection in the differentiation from Taiwanese cultivars. In the analyses of Japanese ornamental vs. fruit cultivars, and fruit vs. small-fruit cultivars, many significant peaks were found, presumably involving ongoing selection in favor of human preference in Japan. Especially, peaks of ca. 3.9 Mb in chromosome 6 overlapped with the SweeD peak (Figs 14b, 16, Table S6), suggesting the strong selection pressure for the small-fruit trait in Japanese cultivars. Accordingly, tests for selection in *P. mume* populations identified its tree-crop-specific patterns for natural or artificial selection.

5.2.2.2 Contribution of positive selections in the establishment of P. mume

A higher number of strong (and successive) peaks were observed in EHH-based scans than in an SFS-based SweeD analysis. This suggested that, in tree (or perennial) crops, most selected alleles are maintained in a heterozygous state, as suggested previously by Akagi et al. (2016). Since most SNPs called in the exon capture approach were located around protein coding regions (Table 10), it is expected that we may be able to detect functional nucleotide polymorphisms (FNPs) even with low sequencing

coverages for each individual (Bamshad et al., 2011; Kaur and Gaikwad, 2017). We could detect candidate genes potentially involved in environmental adaptation (Fig. 14, Tables S5, S6). For instance, in the regions with strong selective sweep, leucine rich repeat containing proteins (e.g., Prupe1G161800, Prupe4G157900, Prupe8G046600, Prupe.8G012000 and Prupe.8G012800), and receptor-like kinases (e.g. Prupe.6G183600 and Prupe.6G261400) would commonly contribute to pathogen recognition pathways (Ellis et al., 2000). The BTB/POZ-MATH-TRAF-like protein (Prupe.1G107200) is potentially associated with virus resistance in *P. armeniaca* (Mariette et al., 2016). Genes potentially involved in stress responses (e.g. Prupe.2G089100, Prupe.2G145200 and Prupe.3G110300) (Vij and Tyagi, 2008; Cheng et al., 2011) were also identified (Tables S5, S6). These results suggest that selection on biotic or abiotic stress responsive genes may have contributed to the geographic separation of Chinese, Japanese and Taiwanese cultivars.

5.3 Estimation of interspecific fragment introgressions between Japanese apricot and related *Prunus* species

5.3.1 Materials and methods

5.3.1.1 Plant materials and datasets

A total of 196 *P. mume* cultivars (79 Chinese, 112 Japanese and 5 Taiwanese cultivars) were used to fragmental introgression analyses. Chi_30, 202, 270, 283, 396, Jap_AM1–8, and SM1–4, which were previously considered to be interspecific hybrids (Hayashi et al., 2008; Zhang et al., 2018), and additionally Chi_250 and Jap_O2, which were newly classified as “Admixed” in Chapter 4 were excluded from the subsequent analyses. For SNP data preparation, the author first removed admixed cultivars from Imputed_set developed in Chapter 4 and subsequently paired Chinese, Japanese, and Taiwanese cultivars with apricots (*P. armeniaca*) or Japanese plums (*P. salicina*). Then, loci with MAF <0.03 were filtered with PLINK (Purcell et al., 2007).

5.3.1.2 Genetic differentiation in genome fragments among subgen. Prunus species

Here, estimation of introgressed genomic positions based on sliding window characterization for indices of genetic differentiation were conducted. To do this, the transition of three indices for population differentiation were assessed: (i) value of the

first principal component (PC1) in PCA, (ii) Q value of the ADMIXTURE analysis with $K = 2$, and (iii) Jost's D value (Jost, 2008), with the sliding window approach. The author conducted "Bin-PCA" and "Bin-Admixture" analyses, which refer PCA from scikit-learn (Pedregosa et al., 2011) and ADMIXTURE (Alexander et al., 2009), to consecutively calculate PC1 and Q values, respectively. The author used 1-Mb bin and 500-kb walking size in Bin-PCA and Bin-Admixture analyses. PC1 values were Z-transformed based on the equation: $zPC1 = (PC1-\mu PC1)/\sigma PC1$. Here, μPC1 and σPC1 indicate the average and standard deviation of PC1, respectively. Q values of Bin-Admixture were transformed into absolute values of the difference between each Q value of *P. mume* individuals (Pm_indv.) and average Q values for *P. armeniaca* or *P. salicina* (related_ave.) as follows: $dQ = |Q_{Pm\_indv.}-Q_{related\_ave.}|$. For calculating Jost's D, vcfWindowedFstats in pypgen 0.2.1 (https://pypi.org/project/pypgen/) was used with a window size of 1 Mb.

5.3.1.3 Detailed analysis on the loci with selective sweep and interspecific introgression

To further examine the inferred region with characteristic selection or introgression, the author focused on some specific regions. Especially, the author narrowed the possibly introgressed regions on chromosome 8 with Bin-Admixture setting the 50-kb bin and 25-kb walking size. Bins with <10 SNPs were removed. Neighbor-joining phylogenetic analysis was performed to visualize the allelic evolution in the specific regions with TASSEL 5.0 (Bradbury et al., 2007), using extracted SNPs of interested regions from Imputed_set. Roots for phylogenetic trees were determined at midpoints using FigTree v1.4.4 (https://github.com/rambaut/figtree/releases).

5.3.2 Results and discussion

5.3.2.1 Fragmental interspecific introgression in subgen. Prunus

In 1-Mb bins, the zPC1 (Bin-PCA) and proportion of the Q values in Admixture with $K = 2$ (Bin-Admixture), were conducted to scan for genome-wide transition of genetic differentiation or potential introgressions, for *P. mume* vs *P. armeniaca* or *P. salicina* (Figs. S1, S2). The Bin-PCA and Bin-Admixture showed mostly consistent results (Figs. S1, S2). In most of the chromosomes, *P. mume* genomes showed signs of fragmental interspecific introgressions (or no clear differentiation between species) from *P. armeniaca* (Fig. S1) or from *P. salicina* (Fig. S2). When the three geographic groups were compared, Japanese cultivars showed the most frequent signals of interspecific

introgressions from *P. armeniaca* or *P. salicina* (Figs. S1b, S2b), while Taiwanese cultivars rarely showed them (Figs. S1c, S2c). The author also calculated a distance-matrix-based Jost's *D* statistic in 1-Mb bins. Jost's *D* values tended to be low in the genomic regions with interspecific introgression signals in Bin-PCA and Bin-Admixture (Figs. S1, S2). However, unlike Bin-PCA and Bin-Admixture, the transition of Jost's *D* values substantially fluctuated in most chromosomes (Figs S1, S2). The overlapped regions of Bin-PCA, Bin-Admixture and Jost's *D* signals may be a strong signature of interspecific introgressions, indicating the especially low allelic divergence between *P. mume* and relatives.

Importantly, some signals of interspecific introgressions were overlapped with the selective sweep (nSL peaks) (Fig. 17; hereafter, "introgression-sweep"), indicating that introgressed regions may have been positively selected in the evolution of *P. mume* populations. They were located on chromosomes 6 and 8 in Chinese cultivars, 2, 3, 4, 6, and 8 in Japanese cultivars, and 6 and 8 in Taiwanese cultivars. In chromosomes 6 and 8, nSL signals harbor fragment introgressions from both *P. armeniaca* and *P. salicina* (Fig. 17a-c). In chromosome 2, 3 and 4 of Japanese cultivars, introgression signals were accompanied by nSL peaks, independently of other groups (Fig. 17b). These results suggest that interspecific introgressions independently contributed to the establishment of not only *P. mume* but also of each geographical group following the divergence of these subpopulations.

It is worth noting that the strongest introgression-sweep signal was detected around 6 Mb on chromosome 8 (Fig. 18a,b). Fine assessment of the Q values with shorter bins (50 kb) increased the resolution for the genomic region with overlap of Bin-Admixture and nSL signals around 6.7–7.1 Mb (Fig. 18c). Next, the author compared evolutionary topologies constructed from the SNPs in region 1 (6.0–6.7 Mb), region 2 (6.7–7.1 Mb), region 3 (7.1–8.5 Mb) and in the whole chromosome 8, according to an approach by Choi and Purugganan (2018) (Fig. 18d–g). The topology for the whole chromosome 8 (Fig. 18d) was mostly consistent with that for whole genome (Fig. 10 in Chapter 4) in accordance with the divergence of species and populations. For regions 1–3, only region 2 showed a distinct topology with alleles of *P. mume* and related species in the subgen. *Prunus* grouped together, while alleles putatively introgressed and under selection were nested in a single clade with *P. salicina* alleles (Fig. 18f), showing very small genetic differentiation (alleles with green band in Fig. 18f). Consequently, region 2 (6.7–7.1 Mb) is thought to have been exposed to positive selection pressure, which may be associated with the adaptive evolution of *P. mume* to import advantageous alleles from other species.

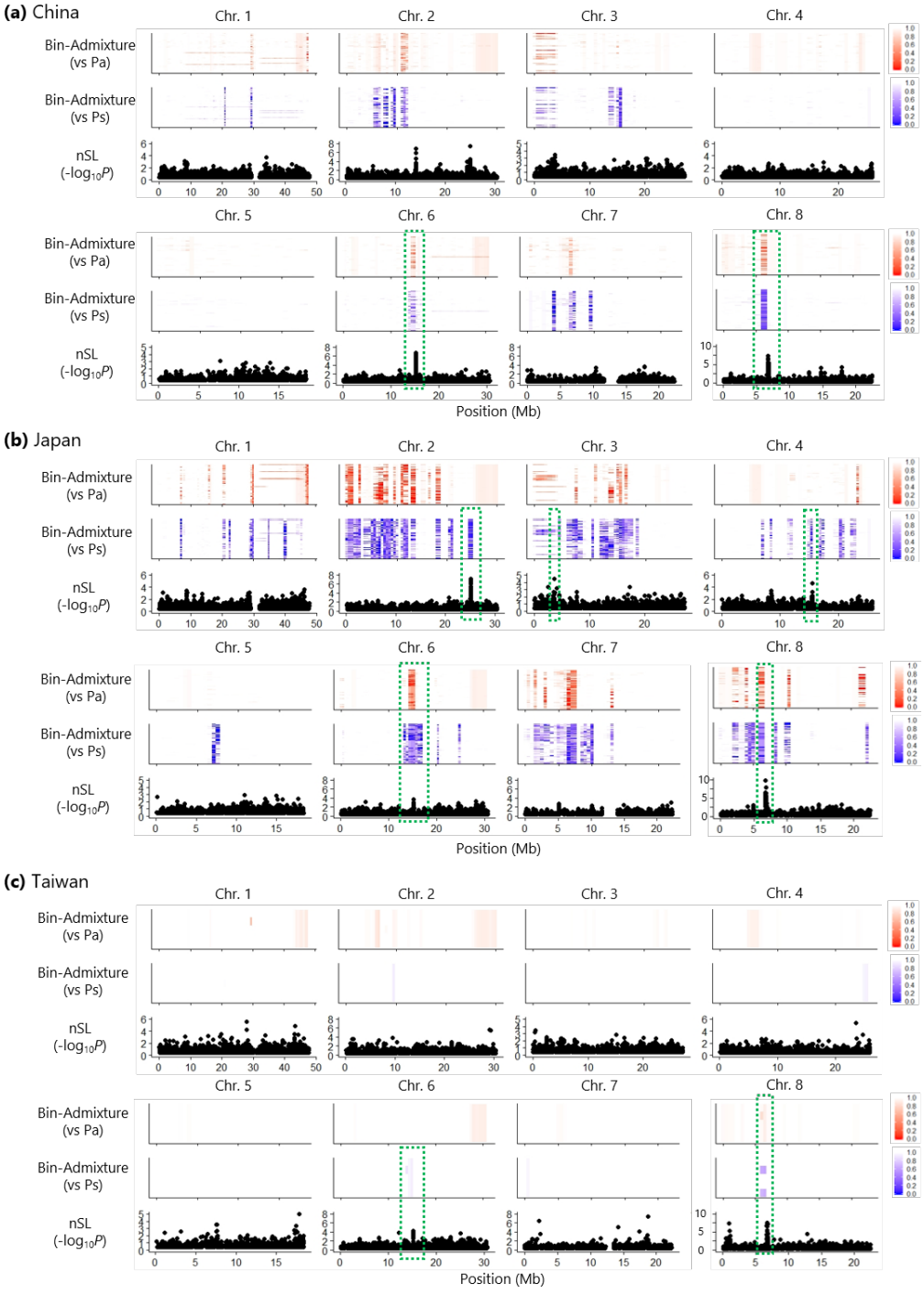

Fig. 17. Examination of conformity between putative interspecific introgression and selective sweep signals. Genomic locations of Bin-Admixture signals were compared with those of nSL peaks. Bin-Admixture (1-Mb-binned) analyses with *Prunus armeniaca* (Pa, red signals) and *P. salicina* (Ps, blue signals), and nSL scans were performed in, (a) Chinese, (b) Japanese, and (c) Taiwanese cultivars of *P. mume*. The degree of introgression is indicated by the color scales to the right (0: highly introgressed–1: not introgressed). Potential introgression-sweep regions are highlighted by green dotted rectangles.

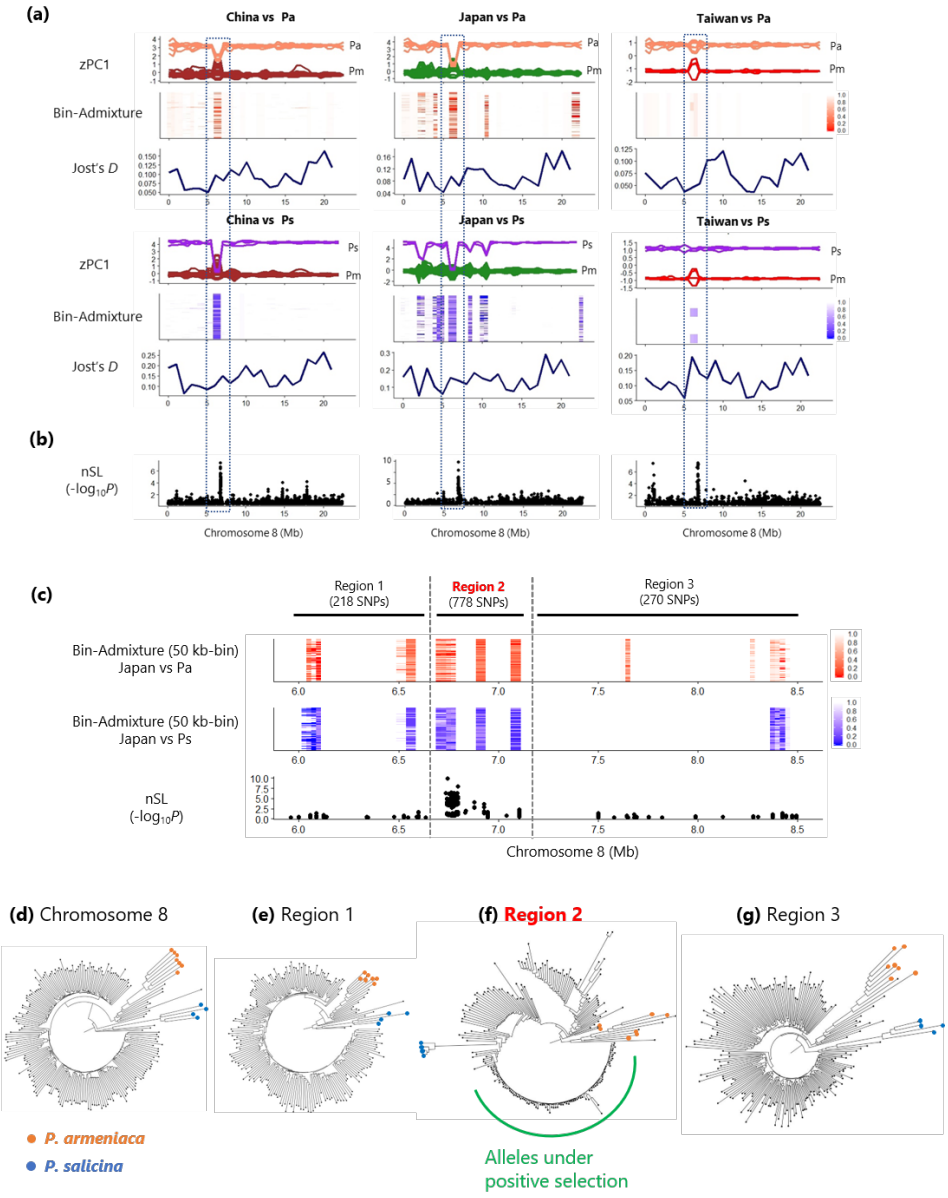

Fig. 18. Consistent region for interspecific introgression and selective sweep on chromosome 8. (a) Bin-PCA, Bin-Admixture, and Jost's *D* patterns in Chinese, Japanese, and Taiwanese cultivars, in comparison to *Prunus armeniaca* (Pa, upper panels) and *P. salicina* (Ps, lower panels). The degree of introgression in Bin-Admixture is indicated by the color scales to the right (0: highly introgressed–1: not introgressed). (b) Transitions of nSL for an index of selective sweep. *P* values calculated with normalized nSL values were shown in $-\log_{10}$ scale. (c) Close up of the genomic fragment (ca. 6.0–8.5Mb) showing overlapped nSL peaks and potential interspecific-introgression, in Japanese cultivars of *P. mume*. We further divided this region into three sub-fragments (Region 1–3) according to the pattern of nSL plots, to assess their phylogenetic relationships. (d) Neighbor-joining phylogenetic trees with the whole SNPs in chromosome 8 and with the SNPs in Regions 1–3. The tree for Region 2 showed a topology inconsistent with the whole chromosome 8 and the flanking regions (Region 1 and 3), and was also inconsistent with the estimated speciation pattern of the subgenus *Prunus*. Alleles that underwent potential selective sweeps were indicated with a green solid line.

5.3.2.2 Contribution of genomic fragments undergoing interspecific introgressions in the evolution of P. mume

According to the results of Bin-PCA, Bin-Admixture, and Jost's *D* analyses, it was suggested that substantial fractions of *P. mume* genomes have frequently exchanged genomic fractions with related species of the subgen. *Prunus*. Large fractions of introgression may indicate that *P. mume*, especially in Japanese cultivars, have experienced a limited number of generations since the interspecific hybridization. Two genomic regions on chromosomes 6 and 8 were detected to have interspecific introgressions in Chinese, Japanese, and Taiwanese cultivar groups (Figs 17, 18). Particularly, a 6.74–6.80-Mb region in the region 2 on chromosome 8 contained high nSL peaks commonly detected among three groups (Table S5). Although this region carries no genes in the reference peach (*P. persica*) genome, we can propose the two following possibilities: 1) only in the *P. mume* genome, the selected haploblock in the corresponded region harbors candidate genes, and 2) this region includes *cis*-regulatory elements affecting gene expression in the flanking regions. For hypothesis 2), often, *cis*-elements were located distantly upstream (>10 kb) of the genes (Clark et al., 2004; Konishi et al., 2006; Ishii et al., 2013; Ricci et al., 2019). Prupe.8G057100 (mitochondrial transcription termination factor) is located ca. 80 kb from the selected haploblock (Fig. 19). Mitochondrial transcription termination factors have been reported to be related to abiotic stress response by controlling the expression level of nuclear genes (Quesada, 2016). Another region on chromosome 6 (ca. 15.2–15.3 Mb) also showed common introgression-sweep (Figs. 17, 20, Table S5). Although this region also harbored no annotated genes in the peach reference genome as well as the described introgression-sweep region on chromosome 8, it may have been important in the evolution of *P. mume*.

Other than the introgression-sweep commonly underwent among the geographical cultivar groups, our exon capture sequencing would allow efficient identification of FNPs in introgression-sweep regions specific to each cultivar groups. Japanese cultivars have the largest fractions of interspecific introgressions, and they also show several geographically unique introgression-sweep regions in nSL (Fig. 17b, Table S5). Of them, interleukin-1 receptor-associated kinase 1 (IRAK1) (Prupe.2G223200), which harbors the introgression signal of *P. salicina* on chromosome 2 (ca. 25 Mb), may act for pathogen recognition pathways (Jebanathirajah et al., 2002; Dardick and Ronald, 2006). On chromosome 3 (ca. 3.6 Mb), 3-epi-6-deoxocathasterone 23-monooxygenase (Prupe.3G050900) was reported to be associated with brassinosteroid biosynthesis (Ohnishi et al., 2006) and is potentially involved in the fruit-enlargement process.

Premnaspirodiene oxygenase (Prupe.4G237900) on chromosome 4 (ca. 15.5 Mb) is a kind of cytochrome P450, which participates in terpene biosynthesis (Weitzel and Simonsen, 2015). Terpenes are the largest class of plant-derived compounds, that have numerous potential applications across food, beverage, pharmaceutical, cosmetic and agriculture industries (Boutanaev et al., 2015). These results suggest that Japanese cultivars might import genetic factors from other species to satisfy the preference of people in Japan.

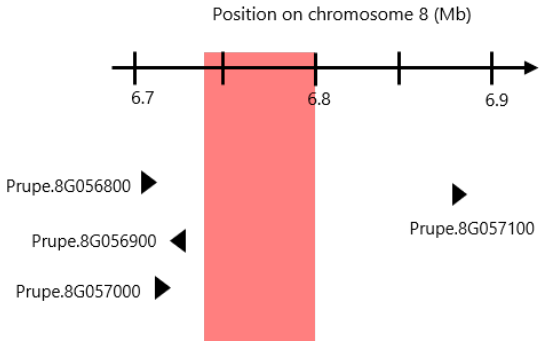

Fig. 19. Positions of annotated genes adjacent to the candidate region (red) on chromosome 8.

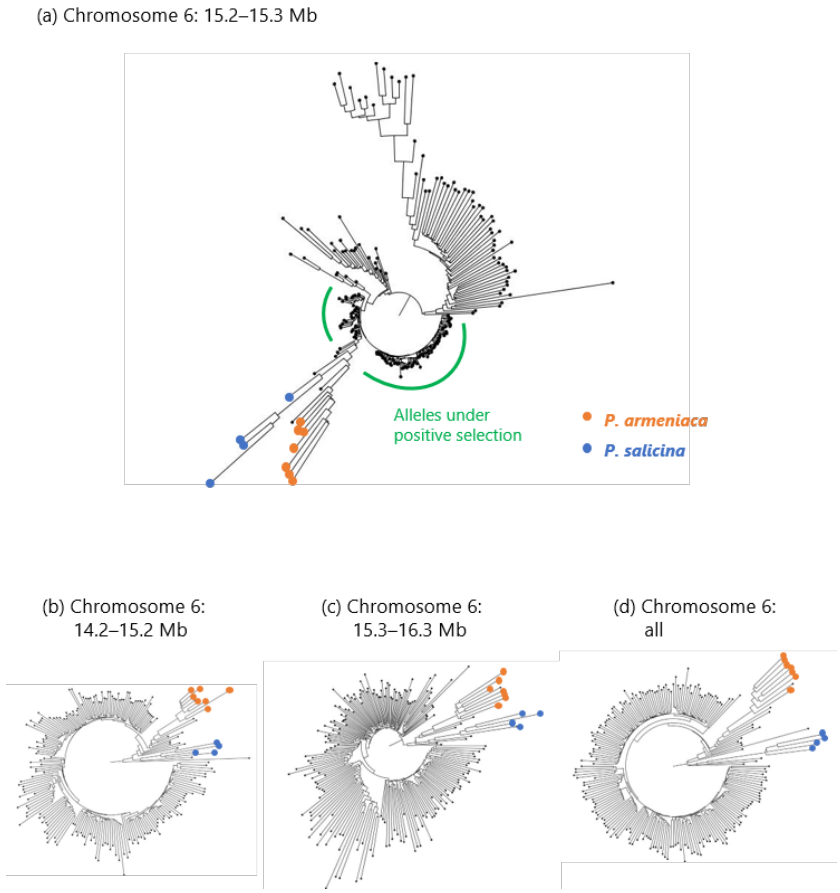

Fig. 20. See next page for caption.

Fig. 20. Neighbor-joining phylogenetic trees with the SNPs in the (a) 15.2–15.3, in its (b) upstream and (c) downstream 1-Mb regions and (d) with the whole SNPs in the chromosome 6. The tree for the 15.2–15.3-Mb region showed a topology inconsistent with the whole chromosome 6 and the flanking regions (14.2–15.2- and 15.3–16.3-Mb regions), and also inconsistent with the estimated speciation pattern of the subgenus *Prunus*. Alleles which undergone potential selective sweeps were indicated with a green solid line.

5.4 Summary

Domestication and population differentiation often involve considerable phenotypic changes in crops. Natural mutations are thought to mainly drive these changes. Besides, occasional interspecific introgressions have potential to contribute for them. Many studies have been tried to clarify the evolutional importance of interspecific introgressions among current crop populations. However, few studies have further estimated the genome-wide distribution of introgressed fragments and their importance. Japanese apricot and the related species in the subgenus *Prunus* are known to have undergone complicated evolutionary processes, involving natural or artificial selections and potential introgressions. Here, to clarify the evolutionary paths to establish the current subgen. *Prunus*, mainly for *P. mume*, the author analyzed genome-wide SNPs based on targeted resequencing of ca. 15,000 exons in East Asian *P. mume* cultivars.

Site frequency spectrum (SFS)-based and extended haplotype homozygosity (EHH)-based approaches were employed for detecting selective sweeps, and observed a higher number of strong (and successive) peaks in EHH-based scans than in an SFS-based analysis. This suggests that, in tree (or perennial) crops, most selected alleles are not completely fixed. We could detect candidate genes potentially involving environmental adaptation, suggesting that the selections on biotic or abiotic stress responsive genes may have contributed to geographic separation of Chinese, Japanese and Taiwanese cultivars.

Sliding window characterization of the indexes for genetic differentiation identified interspecific fragment introgressions between *P. mume* and related species (plum and apricot). These regions often exhibited strong selective sweeps formed in the paths of establishment of *P. mume*, suggesting that *P. mume* has frequently imported advantageous genes from other species in the subgenus *Prunus*, as adaptive evolution. These findings shed light on the complicated nature of evolution with interspecific introgression and natural or artificial selection in fruit tree crop.

Chapter 6. General discussion

In the present study, genetic diversity of Japanese apricot (*P. mume*) was investigated using a series of the latest analyses of molecular population genetics to reveal the evolutional process in *P. mume*. In Chapters 2 and 3, microsatellite markers were newly designed based on the reference genome of *P. mume*, and current population structure and demographic history of Japanese and Taiwanese cultivars of *P. mume* were roughly estimated. In Chapter 4, using exon-targeted resequencing combined with the published data for the Chinese cultivars, the population structure of *P. mume* cultivars was investigated. As a result, the current population structure and diversity of East Asian cultivars of *P. mume* were clearly estimated. They seemed to be associated with geographic separation (e.g., China, Japan and Taiwan) and human preference (e.g., fruit and flower ornamental purposes). Further, in Chapter 5, the genomic regions responsible for the formation of population structure of *P. mume* were estimated. Interestingly, some of these regions were associated with the interspecific introgressions, indicating some introgressed genomic factors may be related to natural or artificial selection in the evolution of *P. mume*.

In this chapter, the general discussion on the evolution of *P. mume* is shown based on the overall results in the past and the present studies.

6.1 The origin and diversity of East Asian cultivars of Japanese apricot

It has empirically been considered that the *P. mume* cultivars have been originated from China, and subsequently, people have distributed them into other East Asian countries such as Japan, Taiwan and so on (Mega et al., 1988; Horiuchi et al., 1996; Faust et al., 2011). However, the origin and diversification process of *P. mume* cultivars have remained to be clarified because of the paucity of molecular marker analyses. Therefore, in the previous and the present studies, PCR-based molecular markers (e.g., RAPD and SSR markers) (Shimada et al., 1994; Hayashi et al., 2008; Yuying et al., 2011; Chapter 2) and next generation sequencing-based analyses (Zhang et al., 2018; Shi et al., 2020; Chapter 4) have been employed to assess the current genetic diversity and the population structure of *P. mume*. Importantly, the Chinese cultivars of *P. mume* were revealed to be genetically distinct from Japanese and Taiwanese cultivars (Figs. 9 and 10) in Chapter 4. This gives a new hypothesis that many Japanese cultivars might be originated in wild *P. mume* populations distributed in Japan (or differentiated from Chinese populations early in the evolutional history). Taiwanese cultivars, which is

characterized by very early flowering and leafing traits, were reported to form a different cluster from Chinese and Japanese cultivars in the previous studies (Shimada et al., 1994; Hayashi et al., 2008; Shi et al., 2020), but in Chapter 4, the genetic differentiation between Japanese and Taiwanese cultivars were found to be relatively small (Figs. 9, 10 and 12, Table 11). These conclusions may be due to the different evolutional speeds between microsatellite and single nucleotide polymorphism (SNP) markers (Haasl and Payseur, 2011). The microsatellites may rapidly mutate and can detect recent signatures of molecular evolution compared to the SNPs.

Japanese *P. mume* cultivars have been historically divided into fruit and ornamental cultivars (Mega et al., 1988; Horiuchi et al., 1996). Fruit cultivars are further classified into large- to small-sized fruit cultivars (Horiuchi et al., 1996). According to the results by Yuying et al. (2011) and in Chapters 2 and 4 in the present study, there might be population differentiation between fruit and ornamental cultivars (Figs. 4, 5, 6 and 10). Small-fruit cultivars tended to belong to a single cluster (Figs. 5, 6 and 10), suggesting that most small-fruit cultivars may have derived from the single origin promoted by the usage of Japanese people.

Although the relative genetic diversity of *P. mume* is difficult to be measured, the extent of linkage disequilibrium (LD) in *P. mume* was indicated to be low (decayed within ca. 100 kb) (Fig. 13), compared to that in peach (decayed within ca. 1,000–2,500 kb) (Akagi et al., 2016). This suggest that *P. mume* populations have higher genetic diversity than peach cultivars. Compared to Chinese cultivars of *P. mume*, Japanese cultivars showed longer LD (Fig. 13) and higher number of cultivar pairs with first degree relationships (pi-hat = 0.5) (Fig. 11). These results suggest that genetic diversity of Japanese cultivars of *P. mume* may have been narrowed down possibly by the recent breeding using the limited genetic resources with preferable phenotypes. Therefore, other cultivar groups with different genetic background, such as Chinese and Taiwanese cultivars may be important sources to maximize the genetic diversity.

6.2 Roles of interspecific introgression and natural selection in Japanese apricot

Population differentiation of *P. mume* may have been induced by geographic separation (e.g., China, Japan and Taiwan) and human preference (e.g., ornamental or fruit purpose) as discussed above. The logs of these evolutional paths may be found in the genomic sequences of current populations (selective sweeps) (Stephan, 2019). Therefore, in Chapter 5, two different approaches: site frequency spectrum (SFS)-based and extended haplotype homozygosity (EHH)-based methods were applied for genome-

wide selective sweep scan. As a result, EHH-based method (Fig. 14) showed more potential to detect selective loci than SFS-based method (Figs. 15 and 16) in *P. mume*. These indicate that *P. mume* experienced a tree-crop specific selection like peach (Akagi et al., 2016), in which selected alleles are not completely fixed due to frequent outcrossing and vegetative propagation. Many candidate genes associated with geographic separation seemed to be related to disease resistance and stress response (Tables S5 and S6). Also, many selective sweeps were detected between ornamental and fruit cultivars and between fruit and small-fruit cultivars, suggesting the selection based on the people preference (Fig. 14b, Table S6). Although candidate genes located around these selective sweeps were not functionally characterized in *Prunus* species, these results are informative for future gene identification on phenotypic differences among cultivar groups. Thus, selective sweeps identified in the present study may reflect the actual evolutional history of *P. mume*.

The importance of interspecific introgression from *P. armeiaca* (apricot) and *P. salicina* (Japanese plum) have long been discussed (Yoshida, 1984; Mega et al., 1988; Yoshida and Yamanishi, 1988; Horiuchi et al., 1996). Recently, molecular analyses based on PCR-based DNA markers have detected some interspecific hybridization evidences between *P. mume* and *P. armeniaca* (Shimada et al., 1994; Hayashi et al., 2008). However, they could identify them in the typical interspecific hybrids such as 'Bungo', which may experience a few generations after the interspecific cross. Interspecific introgressions are thought to be important also in the evolutions of other annual or perennial crops, such as rice (Choi and Purugganan, 2018), wheat (He et al., 2019), maize (Hufford et al., 2013; Brandenburg et al., 2017), apple (Cornille et al., 2012), and olive (Diez et al., 2015; Gros-Balthazard et al., 2019). These reports have pointed out that crop-wild introgressions contributed for the transfer of wild beneficial alleles into domesticates under new climatic and agricultural conditions. On the other hand, in woody crops, the genomic landscapes of interspecific introgressions and their contributions have been poorly understood. Therefore, in Chapter 5, indices for population structure were consecutively calculated (Bin-PCA, Bin-Admixture and Jost's *D*) to estimate positions of interspecific introgressed genomic fragments. As a result, many candidate regions for interspecific introgression from *P. armeniaca* and *P. salicina* were detected in *P. mume* (Figs. S1 and S2), indicating the complicated evolutional process in *P. mume*. Proportion of introgressed genomic region seemed to be the highest in Japanese cultivars and the lowest in Taiwanese cultivars (Figs. S1 and S2). This suggests that introgressed genetic factors from *P. armeniaca* and *P. salicina* may be advantageous for adaptation in Japan (or use by Japanese people) in *P. mume*, as suggested by the previous reports (Yoshida, 1984; Mega

et al., 1988; Horiuchi et al., 1996).

The present study detected many naturally/artificially selected regions derived from interspecific introgressions ("introgression-sweep", Fig. 17), and they seemed to considerably contribute for the establishment of current *P. mume* populations. Of these, common introgression-sweeps on chromosomes 6 and 8 (Figs. 17, 18 and 20) may be the most important among *P. mume* cultivars. In addition, there were several specific introgression-sweep signatures for Japanese cultivars of *P. mume* (Fig. 17b), suggesting that introgressed genetic factors were naturally/artificially involved during the formation of Japanese cultivars.

6.3 Inferences for the evolution of Japanese apricot

The present results propose a new evolutionary model for the establishment of the current *P. mume* populations, where frequent interspecific introgressions with natural/artificial selections were involved (Fig. 21). Several hybridization events might occur in differentiating three species of the subgenus *Prunus*: *P. mume*, *P. armeniaca* and *P. salicina*. When the species of *P. mume* was generated, advantageous interspecific introgressions on chromosomes 6 and 8 (Fig. 17, 18 and 20) were first occurred. After the emergence of the prototype, three core cultivated populations were further differentiated in China, Japan and Taiwan. Probably, some introgressed regions from *P. armeniaca* and *P. salicina* may contribute for the positive selections on establishment of each population (e.g., chromosomes 2, 3 and 4 in Japanese cultivar, Fig. 17b). In addition, among Japanese *P. mume* cultivars, ornamental, fruit and small-fruit cultivars were further found to form sub-clusters, suggesting that human preference-associated selections worked for their differentiation (Fig. 14b, Table S6).

The evolution of Japanese apricot is just started to be revealed. The present findings shed light on the complicated nature of evolution with interspecific introgressions and natural or artificial selection in a fruit tree crop.

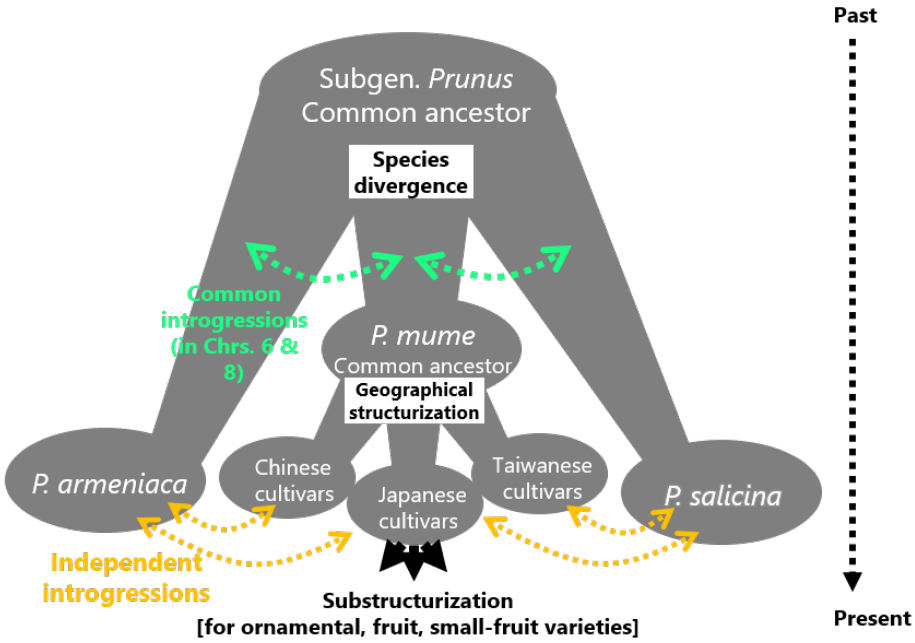

Fig. 21. Tentative process of evolution in *Prunus mume*. The three subgenus *Prunus* species, Japanese apricot (*P. mume*), apricot (*P. armeniaca*), and Japanese plum (*P. salicina*) may have diverged from their common ancestor, experiencing mutual hybridizations. Important introgressions commonly detected on chromosomes 6 and 8 may have been selected during the formation of *P. mume* common ancestor. *P. mume* may have further experienced independent introgressions and selections, resulting in differentiation based on geographical separation and human preference.

Conclusion

Domestication and cultivar differentiation of *P. mume* have little been elucidated despite the importance of this species in the East Asia. In the current study, genetic diversity of *P. mume* was investigated using the latest molecular and population genetic analyses to reveal the evolutional process in *P. mume* as a fruit tree.

In Chapter 2, microsatellite markers were newly designed based on the reference genome of *P. mume*. They were used successfully for fingerprinting most of the *Prunus* cultivars examined (124 *P. mume* cultivars and one cultivar each of *P. armeniaca*, *P. salicina*, *P. persica*, and *P. dulcis*), and the resulting genotype data were used to examine the genetic differentiation of six Japanese apricot cultivar groups, including those producing normal fruit, small-fruit, and ornamental flowers, as well as Taiwanese cultivars, putative hybrids of *P. armeniaca* and *P. mume*, and putative hybrids of *P. salicina* and *P. mume*. Phylogenetic cluster analysis showed three clades with high support values: one clade comprised the putative *P. armeniaca* × *P. mume* hybrids, and the two others included Taiwanese and ornamental cultivars. The rest of the accessions were clustered into two wide clusters, but not clearly divided into the respective cultivar groups. These results indicate that many factors such as human preference, geographical separation, introgression, and local breeding, may have been involved to form the present complex genetic structure in Japanese apricot.

In Chapter 3, an approximate Bayesian computation (ABC) analysis was carried out using microsatellite markers developed in Chapter 2 to clarify the differentiation history among Japanese (fruit, small-fruit and ornamental) and Taiwanese cultivars. In the best scenario, Japanese and Taiwanese populations were estimated to have first diverged, followed by the separation of ornamental cultivars among Japanese populations, and final differentiation between fruit and small-fruit cultivars. The results roughly suggest that Japanese and Taiwanese populations were separated through the geographic isolation with different climate conditions, and ornamental, fruit and small-fruit cultivars were recently differentiated based on human preference in Japan.

In Chapter 4, using target capture method by next generation sequencing (NGS), the population structure of current *P. mume* cultivars in the East Asia was re-estimated based on the SNPs in ca. 15,000 targeted exons, merged with published resequencing data of 79 Chinese *P. mume* cultivars. Principal component analysis, ADMIXTURE and maximum likelihood phylogenetic analysis indicated that Japanese apricot (*P. mume*), apricot (*P. armeniaca*), and Japanese plum (*P. salicina*) form distinct clusters. Importantly, Chinese and Japanese cultivars of *P. mume* were clustered into separate groups, whereas

Taiwanese cultivars were clustered with Japanese cultivars. In addition, most of the fruit (36 of 45 cultivars), small-fruit (9 of 10 cultivars) and ornamental (25 of 45 cultivars) cultivars belonged to the same phylogenetic cluster. This would support the possibility that the differentiation of Chinese and Japanese populations predated that of the Taiwanese population, and subsequently human preference triggered a recent differentiation of Japanese population from the same genetic resources.

In Chapter 5, the genomic regions associated with geographic isolation (e.g., China, Japan and Taiwan) and human usage (e.g., for fruit or ornamental purposes in Japan) were estimated. Furthermore, the interspecific introgressions from *P. armeniaca* and *P. salicina* in the *P. mume* were examined using above genome-wide SNPs in East Asian *P. mume* cultivars. Site frequency spectrum (SFS)-based and extended haplotype homozygosity (EHH)-based approaches were employed for detecting selective sweeps. A higher number of strong (and successive) peaks were observed in EHH-based scans than in an SFS-based analysis. This suggested that, in tree (or perennial) crops, most selected alleles are not completely fixed. Candidate genes potentially involving adaptation to local environment and human preference were detected. Sliding window characterization of the indices for genetic differentiation identified interspecific fragment introgressions between *P. mume* and related species (plum and apricot). Importantly, these regions often exhibited strong selective sweeps formed in the paths of establishment of *P. mume*, suggesting that *P. mume* has frequently imported advantageous genes from other species in the subgenus *Prunus*, as adaptive evolution.

In the present study, the current population structure of East Asian *P. mume* was revealed in detail, and genome-wide profiles of interspecific fragment introgression with positive selection were unveiled for the first time in the fruit tree crops. Interspecific introgressions are evolutionally important because they could contribute to rapid distribution and adaptation to new climatic and agricultural conditions through importing advantageous genes from the relative species. The evolution of Japanese apricot is just started to be revealed. The present findings would shed light not only on the evolutional studies of fruit trees but on the future breeding programs of *P. mume* with maximizing the genetic diversity.

References

Akagi, T., Hanada, T., Yaegaki, H., Gradziel, T. M., and Tao, R. (2016). Genome-wide view of genetic diversity reveals paths of selection and cultivar differentiation in peach domestication. *DNA Research* 23, 271–282.

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19, 1655–1664.

Allaby, R. G., Ware, R. L., and Kistler, L. (2019). A re-evaluation of the domestication bottleneck from archaeogenomic evidence. *Evolutionary Applications* 12, 29–37.

Anderson, E. C., and Thompson, E. A. (2002). A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160, 1217–1229.

Baack, E. J., and Rieseberg, L. H. (2007). A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development* 17, 513–518.

Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A., et al. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Reviews Genetics* 12, 745–755.

Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics* 162, 2025–2035.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27, 573–580.

Bortiri, E., Oh, S.-H., Jiang, J., Baggett, S., Granger, A., Weeks, C., et al. (2001). Phylogeny and systematics of *Prunus* (Rosaceae) as determined by sequence analysis of ITS and the chloroplast *trnL-trnF* spacer DNA. *Systematic Botany* 26, 797–807.

Boutanaev, A. M., Moses, T., Zi, J., Nelson, D. R., Mugford, S. T., Peters, R. J., et al. (2015). Investigation of terpene diversification across multiple sequenced plant genomes. *Proceedings of the National Academy of Sciences, U. S. A.* 112, E81–E88.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635.

Brandenburg, J.-T., Mary-Huard, T., Rigaill, G., Hearne, S. J., Corti, H., Joets, J., et al. (2017). Independent introductions and admixtures have contributed to adaptation of European maize and its American counterparts. *PLOS Genetics* 13, e1006666.

Brantley, W. (2004). About Pluots. *Gastronomica* 4, 84–89.

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics* 103, 338–348.

Chen, E., Huang, X., Tian, Z., Wing, R. A., and Han, B. (2019). The genomics of *Oryza* species provides insights into rice domestication and heterosis. *Annual Review of Plant Biology* 70, 639–665.

Chen, J. (2017). *China mei flower (Prunus mume) cultivars in colour*. Beijing: China Forestry Publishing House.

Cheng, Y., Zhou, W., El Sheery, N. I., Peters, C., Li, M., Wang, X., et al. (2011). Characterization of the Arabidopsis glycerophosphodiester phosphodiesterase (GDPD) family reveals a role of the plastid-localized AtGDPD1 in maintaining cellular phosphate homeostasis under phosphate starvation. *The Plant Journal* 66, 781–795.

Choi, J. Y., and Purugganan, M. D. (2018). Multiple Origin but single domestication led to *Oryza sativa*. *G3: Genes, Genomes, Genetics* 8, 797–803.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* 6, 80–92.

Clark, L. V., and Jasieniuk, M. (2011). POLYSAT: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources* 11, 562–566.

Clark, R. M., Linton, E., Messing, J., and Doebley, J. F. (2004). Pattern of diversity in the genomic region near the maize domestication gene *tb1*. *Proceedings of the National Academy of Sciences, U. S. A.* 101, 700–707.

Cornille, A., Gladieux, P., Smulders, M. J. M., Roldán-Ruiz, I., Laurens, F., Le Cam, B., et al. (2012). New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS Genetics* 8, e1002703.

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., et al. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* 30,

1187–1189.

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.

Dardick, C., and Ronald, P. (2006). Plant and Animal Pathogen Recognition Receptors Signal through Non-RD Kinases. *PLoS Pathogens* 2, e2.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. *Nature* 418, 700–707.

Diez, C. M., Trujillo, I., Martinez-Urdiroz, N., Barranco, D., Rallo, L., Marfil, P., et al. (2015). Olive domestication and diversification in the Mediterranean Basin. *New Phytologist.* 206, 436–447.

Doebley, J. F., Gaut, B. S., and Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell* 127, 1309–1321.

Doebley, J., Stec, A., and Hubbard, L. (1997). The evolution of apical dominance in maize. *Nature* 386, 485–488.

Dvorak, J., Deal, K. R., Luo, M.-C., You, F. M., von Borstel, K., and Dehghani, H. (2012). The Origin of Spelt and Free-Threshing Hexaploid Wheat. *Journal of Heredity* 103, 426–441.

Earl, D. A., and vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4, 359–361.

Ebana, K., Kojima, Y., Fukuoka, S., Nagamine, T., and Kawase, M. (2008). Development of mini core collection of Japanese rice landrace. *Breeding Science* 58, 281–291.

Ellis, J., Dodds, P., and Pryor, T. (2000). Structure, function and evolution of plant disease resistance genes. *Current Opinion in Plant Biology* 3, 278–284.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14, 2611–2620.

Excoffier, L., Estoup, A., and Cornuet, J.-M. (2005). Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics* 169, 1727–1738.

Faust, M., Timon, B., Surányi, D., Nyujtó, F., and Gradziel, T. M. (2011). *Origin and dissemination of Prunus crops: peach, cherry, apricot, plum and almond.* , ed. J. Jules Belgium: ISHS.

Ferrer-Admetlla, A., Liang, M., Korneliussen, T., and Nielsen, R. (2014). On detecting incomplete soft or hard selective sweeps using haplotype structure. *Molecular Biology and Evolution* 31, 1275–1291.

Frantz, L. A. F., Mullin, V. E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri, A., et al. (2016). Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science* 352, 1228–1231.

Gao, Z. H., Shen, Z. J., Han, Z. H., Fang, J. G., Zhang, Y. M., and Zhang, Z. (2004). Microsatellite markers and genetic diversity in Japanese apricot (*Prunus mume*). *HortScience* 39, 1571–1574.

Gaunt, T. R., Rodriguez, S., Zapata, C., and Day, I. N. (2006). MIDAS: software for analysis and visualisation of interallelic disequilibrium between multiallelic markers. *BMC Bioinformatics* 7, 227.

Gaut, B. S., Díez, C. M., and Morrell, P. L. (2015). Genomics and the contrasting dynamics of annual and perennial domestication. *Trends in Genetics* 31, 709–719.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W., et al. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature Biotechnology* 27, 182–189.

Gros-Balthazard, M., Besnard, G., Sarah, G., Holtz, Y., Leclercq, J., Santoni, S., et al. (2019). Evolutionary transcriptomics reveals the origins of olives and the genomic changes associated with their domestication. *The Plant Journal* 100, 143–157.

Haas, M., Schreiber, M., and Mascher, M. (2019). Domestication and crop evolution of wheat and barley: Genes, genomics, and future directions. *Journal of Integrative Plant Biology* 61, 204–225.

Haasl, R. J., and Payseur, B. A. (2011). Multi-locus inference of population structure: a comparison between single nucleotide polymorphisms and microsatellites. *Heredity* 106, 158–171.

Harrison, R. G., and Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity* 105, 795–809.

Hayashi, K. (2009). Genetic diversity analysis in Japanese apricot using DNA markers (In Japanese). Ph.D. Thesis, University of Tsukuba, Tsukuba.

Hayashi, K., Shimazu, K., Yaegaki, H., Yamaguchi, M., Iketani, H., and Yamamoto, T. (2008). Genetic diversity in fruiting and flower-ornamental Japanese apricot (*Prunus mume*) germplasms assessed by SSR markers. *Breeding science* 58, 401–410.

He, F., Pasam, R., Shi, F., Kant, S., Keeble-Gagnere, G., Kay, P., et al. (2019). Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nature Genetics* 51, 896–904.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution* 59, 1633–1638.

Hijikata, S. (1984). "Ume. Hinsyu seitai to saibai," in *Nougyou gijyutsu taikei, Kaju-hen 6* (Tokyo: Nosan Gyoson Bunka Kyokai), 27–46 (In Japanese).

Horiuchi, S., Yoshida, M., Kariya, H., Nakamura, T., Hasebe, H., Suzaki, T., et al. (1996). Nihonnoume Sekainoume. Tokyo: Yokendo (In Japanese).

Htun, T. M., Inoue, C., Chhourn, O., Ishii, T., and Ishikawa, R. (2014). Effect of quantitative trait loci for seed shattering on abscission layer formation in Asian wild rice *Oryza rufipogon*. *Breeding Science* 64, 199–205.

Hudson, R. R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* 18, 337–338.

Hufford, M. B., Lubinksy, P., Pyhäjärvi, T., Devengenzo, M. T., Ellstrand, N. C., and Ross-Ibarra, J. (2013). The genomic signature of crop-wild introgression in maize. *PLoS Genetics*. 9, e1003477.

Iemura, H., Shimazu, K., and Natsumi, K. (1995). Occurrence of a graft-transmissible disease of mume (Abstract in Japanese). *Annals of the Phytopathological Society of Japan* 61, 631.

Ishii, T., Numaguchi, K., Miura, K., Yoshida, K., Thanh, P. T., Htun, T. M., et al. (2013). *OsLG1* regulates a closed panicle trait in domesticated rice. *Nature Genetics* 45, 462–465.

Iwata, H., Hayashi, T., Terakami, S., Takada, N., Sawamura, Y., and Yamamoto, T. (2013). Potential assessment of genome-wide association study and genomic selection in Japanese pear *Pyrus pyrifolia*. *Breeding Science* 63, 125–140.

Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806.

Jebanathirajah, J. A., Peri, S., and Pandey, A. (2002). Toll and interleukin-1 receptor (TIR) domain-containing proteins in plants: a genomic perspective. *Trends in Plant Science* 7, 388–391.

Jin, J., Huang, W., Gao, J.-P., Yang, J., Shi, M., Zhu, M.-Z., et al. (2008). Genetic control of rice plant

architecture under domestication. *Nature Genetics* 40, 1365–1369.

Jost, L. (2008). GST and its relatives do not measure differentiation. *Molecular Ecology* 17, 4015–4026.

Kaga, A., Shimizu, T., Watanabe, S., Tsubokura, Y., Katayose, Y., Harada, K., et al. (2012). Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. *Breeding Science* 61, 566–592.

Kansako, M., Iemura, H., Natsumi, K., and Shimazu, K. (2000). Studies on Japanese apricot [*Prunus mume*] leaf-edge necrosis disease (In Japanese). *Bulletin of the Wakayama Research Center of Agriculture, Forestry and Fisheries* 1, 67–78.

Kaur, P., and Gaikwad, K. (2017). From Genomes to GENE-omes: Exome Sequencing Concept and Applications in Crop Improvement. *Frontiers in Plant Science* 8, 2164.

Kistler, L., Maezumi, S. Y., Souza, J. G. de, Przelomska, N. A. S., Costa, F. M., Smith, O., et al. (2018). Multiproxy evidence highlights a complex evolutionary legacy of maize in South America. *Science* 362, 1309–1313.

Kitamura, Y., Habu, T., Yamane, H., Nishiyama, S., Kajita, K., Sobue, T., et al. (2018). Identification of QTLs controlling chilling and heat requirements for dormancy release and bud break in Japanese apricot (Prunus mume). *Tree Genetics & Genomes* 14, 33.

Kloosterman, A. D., Budowle, B., and Daselaar, P. (1993). PCR-amplification and detection of the human D1S80 VNTR locus. Amplification conditions, population genetics and application in forensic analysis. *International Journal of Legal Medicine* 105, 257–264.

Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., et al. (2009). VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25, 2283–2285

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 Research*. 22, 568–576.

Konishi, S., Izawa, T., Lin, S. Y., Ebana, K., Fukuta, Y., Sasaki, T., et al. (2006). An SNP caused loss of seed shattering during rice domestication. *Science* 312, 1392–1396.

Kosova, G., Pickrell, J. K., Kelley, J. L., McArdle, P. F., Shuldiner, A. R., Abney, M., et al. (2010). The CFTR Met 470 allele is associated with lower birth rates in fertile men from a population isolate.

PLOS Genetics 6, e1000974.

Kurihara, J., Tomaru, K., Otsubo, T., T. Natsuaki, K. T., and Arimoto, Y. (2006). Mume leaf margin necrosis, a new disease of Japanese apricot [*Prunus mume*] caused by multiple infection with *Cucumber mosaic virus* and prunus necrotic ringspot related *Ilarvirus*. *Journal of agriculture science, Tokyo University of Agriculture* 50, 95–105

Kyotani, H., Yoshida, M., Yamaguchi, M., Ishizawa, Y., Kozono, T., Nshida, T., et al. (1988). Breeding of plum-mume parental lines "PM-1-1" and "PM-1-4", interspecific hybrids of Japanese plum (*Prunus salicina* Lindl.) and mume (*P. mume* Sieb. et Zucc.). *Bulletin of the Fruit Tree Research Station* 15, 1–10.

Lee, T.-H., Guo, H., Wang, X., Kim, C., and Paterson, A. H. (2014). SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* 15, 162.

Lee, W., Ahn, S., Taye, M., Sung, S., Lee, H.-J., Cho, S., et al. (2016). Detecting positive selection of Korean native goat populations using next-generation sequencing. *Moleucles and Cells* 39, 862–868.

Li, C., Zhou, A., and Sang, T. (2006). Rice domestication by reducing shattering. *Science* 311, 1936–1939.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997 [q-bio]*.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.

MAFF (2017). Statistics on production of fruit classified by cultivar. Available at: https://www.maff.go.jp/j/tokei/kouhyou/tokusan_kazyu/ [Accessed May 5, 2020] (In Japanese).

MAFF (2019). The 93re statistical yearbook of Ministry of Agriculture, Forestry and Fisheries. Available at: https://www.maff.go.jp/e/data/stat/93th/index.html [Accessed May 5, 2020].

Mariette, S., Wong Jun Tai, F., Roch, G., Barre, A., Chague, A., Decroocq, S., et al. (2016). Genome-wide association links candidate genes to resistance to *Plum Pox Virus* in apricot (*Prunus armeniaca*). *New Phytologist.* 209, 773–784.

Mascher, M., Schuenemann, V. J., Davidovich, U., Marom, N., Himmelbach, A., Hübner, S., et al. (2016). Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of

barley. *Nature Genetics* 48, 1089–1093.

McClure, K. A., Sawler, J., Gardner, K. M., Money, D., and Myles, S. (2014). Genomics: a potential panacea for the perennial problem. *American Journal of Botany* 101, 1780–1790.

Mega, K., Tomita, E., Kitamura, S., Saito, S., and Mizukami, S. (1988). The grand dictionary of horticulture. , ed. T. Aoba Tokyo: Shogakukan (In Japanese).

Mehlenbacher, S., Cocin, V., and Hough, L. (1991). "Apricots (*Prunus*)," in *Genetic resources of temperate fruit and nut crops* (Wagenengen: ISHS), 65–107.

Miller, A. J., and Gross, B. L. (2011). From forest to field: perennial fruit crop domestication. *American Journal of Botany* 98, 1389–1414.

Molnár-Láng, M., Ceoloni, C., and Doležel, J. (2015). *Alien Introgression in Wheat: Cytogenetics, Molecular Biology, and Genomics*. New York: Springer.

Morimoto, T., Kitamura, Y., Numaguchi, K., Akagi, T., and Tao, R. (2019). Characterization of post-mating interspecific cross-compatibility in *Prunus* (Rosaceae). *Scientia Horticulturae* 246, 693–699.

Nadachowska-Brzyska, K., Burri, R., and Ellegren, H. (2019). Footprints of adaptive evolution revealed by whole Z chromosomes haplotypes in flycatchers. *Molecular Ecology.* 28, 2290–2304.

Nakaune, R., Takeda, T., Numaguchi, K., Nakazono-Nagaoka, E., and Fujikawa, T. (2018). Ampeloviruses associated with incomplete flower syndrome and leaf-edge necrosis in Japanese apricot. *Journal of General Plant Pathology* 84, 202–207.

Narasimhan, V. M., Patterson, N., Moorjani, P., Rohland, N., Bernardos, R., Mallick, S., et al. (2019). The formation of human populations in South and Central Asia. *Science* 365, eaat7487.

Nei, M. (1972). Genetic distance between populations. *The American Naturalist* 106, 283–292.

Nei, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.

Nei, M., Tajima, F., and Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. Journal of Molecular Evolution 19, 153–170.

Nielsen, R., Williamson, S., Kim, Y., Hubisz, M. J., Clark, A. G., and Bustamante, C. (2005). Genomic scans for selective sweeps using SNP data. *Genome Research* 15, 1566–1575.

Numaguchi, K., Takeda, T., Tsuchida, Y., and Nakaune, R. (2019). Large-scale field survey reveals overall

yield loss in Japanese apricot possibly caused by two ampeloviruses. *Journal of General Plant Pathology* 85, 116–121.

Ohnishi, T., Szatmari, A.-M., Watanabe, B., Fujita, S., Bancos, S., Koncz, C., et al. (2006). C-23 hydroxylation by *Arabidopsis* CYP90C1 and CYP90D1 reveals a novel shortcut in brassinosteroid biosynthesis. *Plant Cell* 18, 3275–3288.

Ohta, S., Hayashi, K., Yaegaki, H., Mitsui, N., Omura, M., Nishitani, C., et al. (2006). Genetic relationship among fruiting and flower-Japanese apricot characterized by chloroplast DNA markers. *DNA Polymorphism* 14, 138–140.

Ohtubo, T., Kurihara, J., Sakakibara, M., and Tomaru, K. (2002). Confirmation of symptom reproduction in mume leaf margin necrosis disease on *Prunus mume* (c.v. Nankou) by sap inoculation with causal two viruses, *Cucumber mosaic* and *Prunus necrotic ringspot*-like viruses. *Journal of agriculture science, Tokyo University of Agriculture* 47, 11–16.

Otsubo, T., Harano, H., Natuaki, K., and Tomaru, K. (1991). Chakasu disorder of Japanese apricot. *Journal of the Japanese Society for Horticultural Science* 60, 164–165.

Panaud, O., Chen, X., and McCouch, S. R. (1996). Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Molecular Genetics and Genomics* 252, 597–607.

Pankin, A., Altmüller, J., Becker, C., and von Korff, M. (2018). Targeted resequencing reveals genomic signatures of barley domestication. *New Phytologist.* 218, 1247–1259.

Patterson, N., Price, A. L., and Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetics.* 2, e190.

Pavlidis, P., Živković, D., Stamatakis, A., and Alachiotis, N. (2013). SweeD: Likelihood-based detection of selective sweeps in thousands of genomes. *Molecular Biology and Evolution* 30, 2224–2234.

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics* 28, 2537–2539.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 2825−2830.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* 81, 559–575.

Purugganan, M. D., and Fuller, D. Q. (2009). The nature of selection during plant domestication. *Nature* 457, 843–848.

Quesada, V. (2016). The roles of mitochondrial transcription termination factors (MTERFs) in plants. *Physiologia Plantarum* 157, 389–399.

R Development Core Team (2008). *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramasamy, R. K., Ramasamy, S., Bindroo, B. B., and Naik, V. G. (2014). STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springerplus* 3, 431.

Raymond, M., and Rousset, F. (1995). GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86, 248–249.

Ricci, W. A., Lu, Z., Ji, L., Marand, A. P., Ethridge, C. L., Murphy, N. G., et al. (2019). Widespread long-range cis-regulatory elements in the maize genome. *Nature Plants* 5, 1237–1249.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* 8, 103–106.

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18, 233–234.

Shi, T., Luo, W., Li, H., Huang, X., Ni, Z., Gao, H., et al. (2020). Association between blooming time and climatic adaptation in *Prunus mume*. *Ecology and Evolution* 10, 292–306.

Shimada, T. (Kobe U. (Japan) F. of A., Haji, T., Yamaguchi, M., Takeda, T., Nomura, K., and Yoshida, M. (1994). Classification of mume (*Prunus mume* Sieb. et Zucc.) by RAPD assay. *Journal of the Japanese Society for Horticultural Science* 63, 543–551.

Smith, O., Nicholson, W. V., Kistler, L., Mace, E., Clapham, A., Rose, P., et al. (2019). A domestication history of dynamic adaptation and genomic deterioration in Sorghum. *Nature Plants* 5, 369–379.

Stephan, W. (2019). Selective Sweeps. *Genetics* 211, 5–13.

Stitzer, M. C., and Ross-Ibarra, J. (2018). Maize domestication and gene interaction. *New Phytologist* 220, 395–408.

Suarez-Gonzalez, A., Hefer, C. A., Lexer, C., Douglas, C. J., and Cronk, Q. C. B. (2018). Introgression from Populus balsamifera underlies adaptively significant variation and range boundaries in P. trichocarpa. *New Phytologist* 217, 416–427.

Szpiech, Z. A., and Hernandez, R. D. (2014). selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. *Molecular Biology and Evolution* 31, 2824–2827.

Takezaki, N., Nei, M., and Tamura, K. (2010). POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Molecular Biology and Evolution* 27, 747–752.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 30, 2725–2729.

Tsuda, Y., Nakao, K., Ide, Y., and Tsumura, Y. (2015). The population demography of Betula maximowicziana, a cool-temperate tree species in Japan, in relation to the last glacial period: its admixture-like genetic structure is the result of simple population splitting not admixing. *Molecular Ecology* 24, 1403–1418.

Tzonev, R., and Yamaguchi, M. (1999). Investigation on same far-east *Prunus* species: Phenology. *Acta Horticulturae* 488, 239–241.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Research* 40, e115–e115.

Verde, I., Jenkins, J., Dondini, L., Micali, S., Pagliarani, G., Vendramin, E., et al. (2017). The Peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. *BMC Genomics* 18, 225.

Vij, S., and Tyagi, A. K. (2008). A20/AN1 zinc-finger domain-containing proteins in plants and animals represent common elements in stress response. *Functional & Integrative Genomics* 8, 301–307.

Voight, B. F., Kudaravalli, S., Wen, X., and Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLOS Biology* 4, e72.

Wada, T., Noguchi, Y., Isobe, S., Kunihisa, M., Sueyoshi, T., and Shimomura, K. (2017). Development of a core collection of strawberry cultivars based on SSR and CAPS marker polymorphisms. *The Horticulture Journal* 86, 365–378.

Wang, H., Nussbaum-Wagler, T., Li, B., Zhao, Q., Vigouroux, Y., Faller, M., et al. (2005). The origin of the naked grains of maize. *Nature* 436, 714–719.

Watanabe, N., Sugiyama, K., Yamagishi, Y., and Sakata, Y. (2002). Comparative telosomic mapping of homoeologous genes for brittle rachis in tetraploid and hexaploid wheats. *Hereditas* 137, 180–185.

Weitzel, C., and Simonsen, H. T. (2015). Cytochrome P450-enzymes involved in the biosynthesis of mono- and sesquiterpenes. *Phytochemistry Reviews* 14, 7–24.

Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., et al. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature Biotechnology* 32, 656–662.

Wu, G. A., Terol, J., Ibanez, V., López-García, A., Pérez-Román, E., Borredá, C., et al. (2018). Genomics of the origin and evolution of Citrus. *Nature* 554, 311–316.

Yaegaki, H., Haji, T., Nakamura, Y., Miyake, M., Nishimura, K., Kyotani, H., et al. (2003). Varietal and yearly variations in fruit and endocarp weights and their ratio in Japanese apricot (*Prunus mume* Sieb. et Zucc.) cultivars (In Japanese with English abstract). *Journal of the Japanese Society for Horticultural Science* 72, 546–550.

Yaegaki, H., Yamaguchi, M., Haji, T., Suesada, Y., Miyake, M., Kihara, T., et al. (2012). New Japanese apricot cultivar "Tsuyuakane" (In Japanese with English abstract). *Bulletin of the NARO Institute of Fruit Tree Science*, 1–6.

Yamaguchi, M., Yaegaki, H., Suesada, Y., Haji, T., and Miyake, M. (2018). Study on interspecific hybridization among the subgenera *Amygdalus*, *Prunophora*, and *Cerasus* (In Japanese). *Horticultural Research (Japan)* 17, 147–159.

Yamane, H. (2014). Regulation of Bud Dormancy and Bud Break in Japanese Apricot (*Prunus mume* Siebold & Zucc.) and Peach [*Prunus persica* (L.) Batsch]: A summary of recent studies. *Journal of the Japanese Society for Horticultural Science* 83, 187–202.

Yoshida, M. (1984). "Ume. Gensanchi to raireki," in *Nougyou gijyutsu taikei, Kaju-hen 6* (Tokyo: Nosan Gyoson Bunka Kyokai), 3–6 (In Japanese).

Yoshida, M., and Yamanishi, H. (1988). Apricot cultivars in Japan. *Acta Horticulturae* 209, 69–81.

Yu, Y., Fu, J., Xu, Y., Zhang, J., Ren, F., Zhao, H., et al. (2018). Genome re-sequencing reveals the evolutionary history of peach fruit edibility. *Nature Communications* 9, 5404.

Yuying, S., Xiajun, D., Fei, W., Binhua, C., Zhihong, G., and Zhen, Z. (2011). Analysis of genetic diversity in Japanese apricot (*Prunus mume* Sieb. et Zucc.) based on REMAP and IRAP molecular markers. *Scientia Horticulturae* 132, 50–58.

Zeder, M. A. (2015). Core questions in domestication research. *Proceedings of the National Academy of Sciences, U. S. A.* 112, 3191–3198.

Zeder, M. A., Emshwiller, E., Smith, B. D., and Bradley, D. G. (2006). Documenting domestication: the intersection of genetics and archaeology. *Trends in Genetics.* 22, 139–155.

Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M., and Yang, T.-L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35, 1786–1788.

Zhang, J., Zhang, Q., Cheng, T., Yang, W., Pan, H., Zhong, J., et al. (2015). High-density genetic map construction and identification of a locus controlling weeping trait in an ornamental woody plant (Prunus mume Sieb. et Zucc). *DNA Research* 22, 183–191.

Zhang, Q., Chen, W., Sun, L., Zhao, F., Huang, B., Yang, W., et al. (2012). The genome of *Prunus mume*. *Nature Communications* 3, 1318.

Zhang, Q., Zhang, H., Sun, L., Fan, G., Ye, M., Jiang, L., et al. (2018). The genetic architecture of floral traits in the woody plant *Prunus mume*. *Nature Communications* 9, 1702.

Appendix

Appendix 1 Supplementary figures

Fig. S1. Chromosomal patterns of genetic differentiation among, (a) Chinese, (b) Japanese, and (c) Taiwanese accessions of *Prunus mume* and *P. armeniaca*.

Fig. S2 Chromosomal patterns of genetic differentiation among, (a) Chinese, (b) Japanese, and (c) Taiwanese accessions of *Prunus mume* and *P. salicina*.

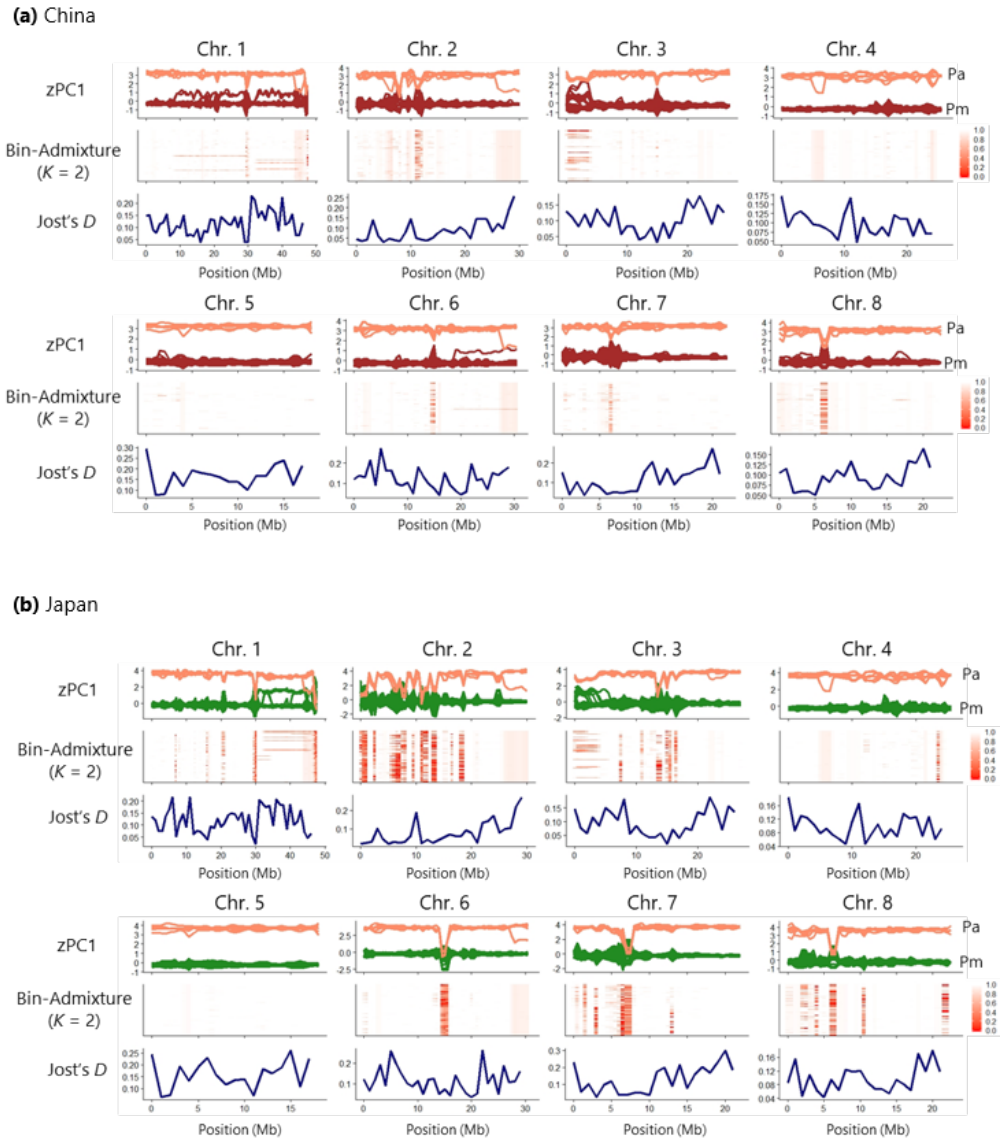

Fig. S1. Chromosomal patterns of genetic differentiation among, (a) Chinese, (b) Japanese, and (c) Taiwanese accessions of *Prunus mume* and *P. armeniaca*. All analyses were 1-Mb-binned. zPC1: Z-transformed PC1 calculated with Bin-PCA, Pa: *P. armeniaca*, Pm: *P. mume*. In Bin-Admixture, strength of red color (color scale, 0: highly introgressed–1: no introgression) indicates similarity to *P. armeniaca*.

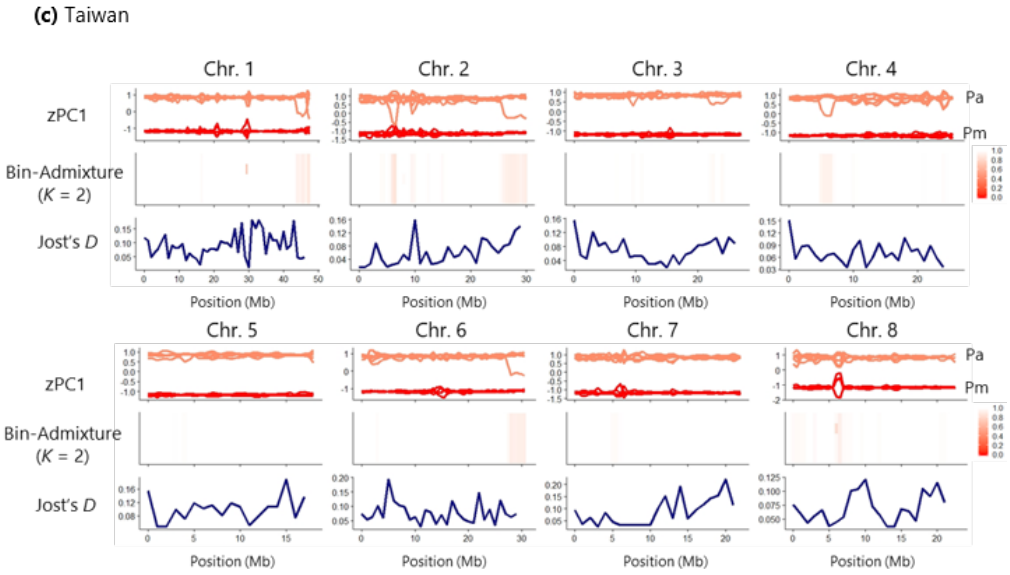

Fig. S1. Continued.

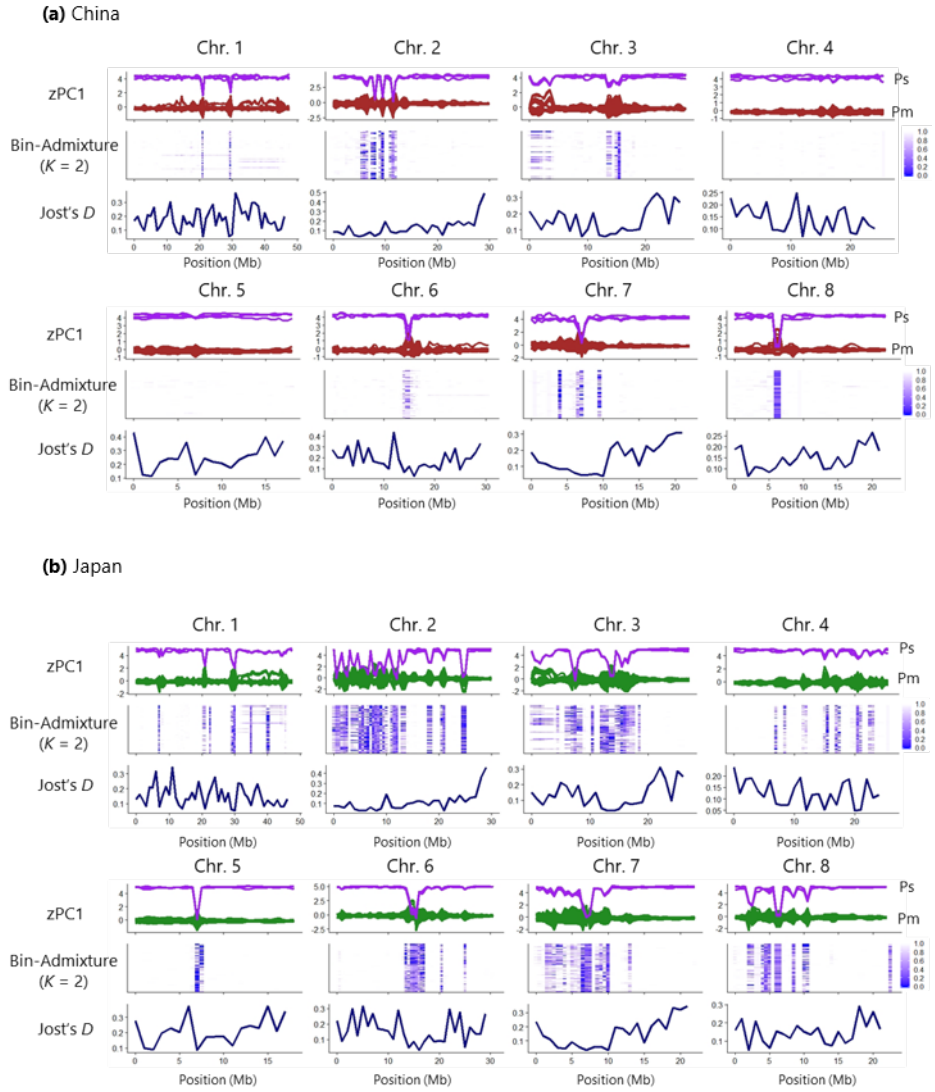

Fig. S2. Chromosomal patterns of genetic differentiation among, (a) Chinese, (b) Japanese, and (c) Taiwanese accessions of *Prunus mume* and *P. salicina*. All analyses were 1-Mb-binned. zPC1: Z-transformed PC1 calculated with Bin-PCA, Ps: *P. salicina*, Pm: *P. mume*. In Bin-Admixture, strength of blue color (color scale, 0: highly introgressed–1: no introgression) indicates similarity to *P. salicina.*

(c) Taiwan

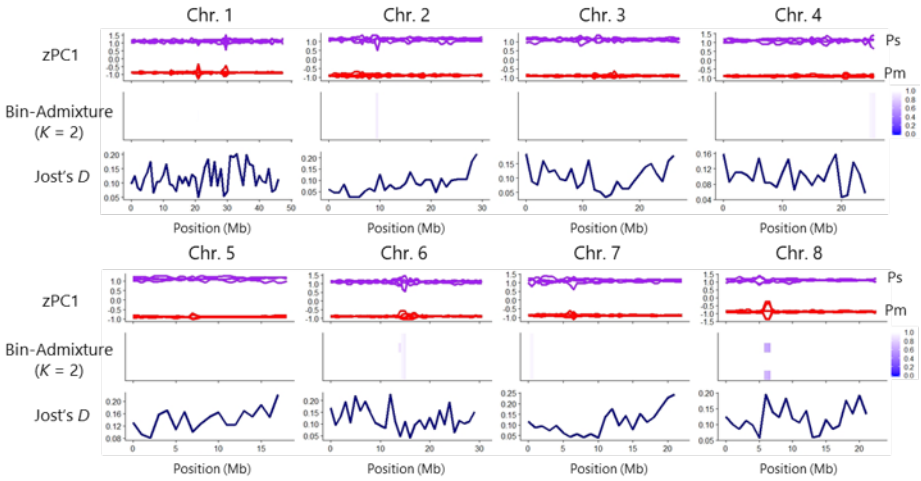

Fig. S2. Continued.

Appendix 2 Supplementary tables

Table S1. Detailed information on the *Prunus* accessions used in this study.

Table S2. Microsatellite markers used in this study.

Table S3. Amplified fragment allele sizes at 20 highly polymorphic microsatellite loci determined among 128 *Prunus* accessions (199 trees).

Table S4. Values (r^2) of linkage disequilibrium (LD) between the most major alleles (below diagonal) and among all combinations of alleles.

Table S5. The strongest candidates for selective sweep based on nSL test for selection.

Table S6. The strongest candidates for selective sweep based on XP-EHH test for selection.

Table S1. Detailed information on the *Prunus* accessions used in chapter 2.

Species	Code[z]	Acc. code	Name	Name in Japanese	Origin	Purpose[y]	No. trees[x]	JP acc. No.[w]		Additional information[v]
P. mume	F	F1	Ankoyabai	安行野梅	Unknown	F	2	-		
		F2	Aojiku	青軸	Nara	F, O	2	-	*	
		F3	Benisashi	紅サシ	Fukui	F	2	113065	*	
		F4_1[u]	Fudono_1	伏菟野	Unknown	F	1	170637		
		F4_2[u]	Fudono_2	伏菟野	Unknown	F	1	-		
		F5	Fukuju	福寿	Unknown	F	2	-		
		F6	Garyubai	臥竜梅	Unknown	F	1	-		
		F7	Gojiro	古城	Wakayama	F	2	172766	*	
		F8	Gyokuei	玉英	Tokyo	F	2	170659	*	
		F9	Hachiro	八郎	Ibaraki	F	2	-	*	Chance seedling of 'Jizoume'
		F10	Hanakami	花香実	Unknown	F	2	170639	*	
		F11	Jizoume	地蔵梅	Wakayama	F	3	172768	*	
		F12	Juro	十郎	Kanagawa	F	2	172769	*	
		F13	Kagajizo	加賀地蔵	Ibaraki	F	2	-	*	'Shirokaga'×'Jizoume'
		F14	Kahoku	河北	Unknown	F	2	-		
		F15	Kaidarewase	皆平早生	Wakayama	F	2	-	*	
		F16	Kairyouchida	改良内田	Wakayama	F	2	170661	*	
		F17	Kensaki	剣先	Fukui	F	2	170644	*	
		F18	Kinotakara	紀の宝	Mie	F	1	-		
		F19	Kinyuji	金熊寺	Osaka	F	2	-	*	
		F20	Kodama	児玉	Unknown	F	2	-	*	
		F21	Kotsubunanko	小粒南高	Wakayama	F	4	-	*	
		F22	Kushino	串野	Unknown	F	2	-		
		F23	Misato 1	美里一号	Wakayama	F	2	-		
		F24_1[u]	Naniwa_1	浪花	Unknown	F	2	172772		
		F24_2[u]	Naniwa_2	浪花	Unknown	F	1	-		
		F25	Naniwahitoe	難波一重	Unknown	F	2	-		
		F26	Nanko	南高	Wakayama	F	6	172773	*	
		F27	NK14	NK14	Wakayama	F	2	-		'Nanko'×'Kensaki'
		F28	Okunoume	奥野梅	Unknown	F	2	-		
		F29	Oshuku	鶯宿	Tokushima	F	2	172777	*	
		F30	Ozaki	尾崎	Unknown	F	1	-		
		F31	Rinshu-Fukui	林州（福井）	Fukui	F	2	-		
		F32	Rinshu-Nara	林州（奈良）	Nara	F	2	170647	*	
		F33	Sadayuume	佐太夫梅	Unknown	F	2	-		
		F34	Sakamoto	坂本	Unknown	F	1	-		
		F35	Shiratama	白玉	Wakayama	F	2	113054	*	
		F36	Shirokaga	白加賀	Unknown	F	2	172785	*	
		F37	Shisen	四川	Unknown	F	1	-		
		F38	Suiko	翠香	Ibaraki	F	3	-		'Gessekai'×'Baigo'
		F39	Taniguchikobai	谷口紅梅	Unknown	F	1	-		
		F40	Tenjin	天神	Unknown	F	2	-	*	
		F41	Tojikobai	東地紅梅	Unknown	F	3	-		
		F42	Toko	橙高	Wakayama	F	2	-		'Nanko'×'Jizoume'
		F43	Yakushi	薬師	Wakayama	F	2	174252	*	
		F44	Yogo1	四郷一号	Wakayama	F	2	-		
		F45	Yosei	養青	Wakayama	F	2	174255	*	
		F46	Zaronbai	座論梅	Unknown	F	1	-		
P. mume	FS	FS1	Benio	紅王	Wakayama	F	1	-	*	
		FS2	Hakuo	白王	Wakayama	F	2	-	*	Putative bud sport of 'Kosyu Saisyo'
		FS3	Kinugasa	衣笠	Wakayama	F	2	-	*	
		FS4	Koshusaisho	甲州最小	Nara	F	2	113057	*	
		FS5	Koyokoume	光陽小梅	Nara	F	2	-	*	
		FS6	Maezawakoume	前沢小梅	Nagano	F	2	-	*	
		FS7	Orihime	織姫	Saitama	F	2	172776	*	
		FS8	Purplequeen	パープルクイーン	Wakayama	F	1	-	*	Putative bud sport of 'Hakuo'
		FS9	Ryukyokoume	竜峡小梅	Nagano	F	2	172779	*	
		FS10	Shinanokoume	信濃小梅	Nagano	F	2	-	*	

Table S1. (Continued).

Species	Code[z]	Acc. code	Name	Name in Japanese	Origin	Purpose[y]	No. trees[x]	JP acc. No.[w]		Additional information
P. mume	O	O1	Akebono	曙	Unknown	O	1	172764		
		O2	Asahitaki	旭滝	Unknown	O	1	-	*	
		O3	Benichidori	紅千鳥	Unknown	O	1	-	*	
		O4	Benioshuku	紅鶯宿	Unknown	O	1	-		
		O5	Chasenbai	茶煎梅	Unknown	O	1	-	*	
		O6	Chinamume	中国梅	China	O	1	-	*	
		O7	Eikan	栄冠	Unknown	O	1	-	*	
		O8	Gekkyuden	月宮殿	Unknown	O	1	-	*	
		O9	Gofukushidare	呉服枝垂	Unknown	O	1	-	*	
		O10	Goshoko	御所紅	Unknown	O	1	-	*	
		O11	Hasegawashibori	長谷川絞り	Unknown	O	1	-	*	
		O12	Hitoeryokugaku	一重緑萼	Unknown	O	1	-	*	
		O13	Horyukaku	芳流閣	Unknown	O	1	-		Triploid variety
		O14	Ikuyonezame	幾夜寝覚	Unknown	O	1	-	*	
		O15	Jakobai	麝香梅	Unknown	O	1	-	*	
		O16	Kagoshimako	鹿児島紅	Unknown	O	1	-	*	
		O17	Kanbaishidare	寒梅枝垂	Unknown	O	1	-	*	
		O18	Kanseishidare	寒成枝垂	Unknown	O	1	-	*	
		O19	Kinko	錦光	Unknown	O	1	-	*	
		O20	Kurohikari	黒光	Unknown	O	1	-	*	
		O21	Kurokumo	黒雲	Unknown	O	1	-	*	
		O22	Mangetsushidare	満月枝垂	Unknown	O	1	170671	*	
		O23	Meotoshirare	夫婦枝垂	Unknown	O	1	-	*	
		O24	Mera	米良	Unknown	O	2	-	*	
		O25	Michishirube	道知辺	Unknown	O	1	170672	*	
		O26	Morinoseki	守の関	Unknown	O	1	-	*	
		O27	Morinoura	守の浦	Unknown	O	1	-		
		O28	Okinaume	翁梅	Unknown	O	1	-	*	
		O29	Omoinomama	思のまま	Unknown	O	1	-		
		O30	Osakazuki	大盃	Unknown	O	1	-	*	
		O31	Sabashiko	佐橋紅	Unknown	O	1	-		
		O32	Seiryushidare	青竜枝垂	Unknown	O	1	-	*	
		O33	Shinheike	新平家	Unknown	O	1	-	*	
		O34	Shirobotan	白牡丹	Unknown	O	1	174237	*	
		O35	Suishinbai	酔心梅	Unknown	O	1	-	*	
		O36	Suoume	栖鶯梅	Unknown	O	1	-		
		O37	Tagonotsuki	田子の月	Unknown	O	1	-		
		O38	Takasago	高砂	Unknown	O	1	-		Triploid variety
		O39	Tamabotan	玉牡丹	Unknown	O	1	174244	*	
		O40	Tanfun	淡粉	Unknown	O	1	-		
		O41	Toji	冬至	Unknown	O	1	174249	*	
		O42	Toyadenotaka	塒出の鷹	Unknown	O	1	-		
		O43	Tsukushiko	筑紫紅	Unknown	O	1	-	*	
		O44	Unryu	雲竜	Unknown	O	1	-		
		O45	Unryubai	雲龍梅	Unknown	O	1	-	*	
		O46	Utsushiroyama	映白山	Unknown	O	1	-		
		O47	Yaetoji	八重冬至	Unknown	O	1	-		
		O48	Yanagawashibori	柳川絞り	Unknown	O	1	-	*	
		O49	Yokihi	楊貴妃	Unknown	O	1	-		
P. mume	T	T1	Ellching	二青梅	Taiwan	F	2	-	*	
		T2	Hakufunbai	白粉梅	Unknown	F	1	-		Putative Taiwanese variety
		T3	ST	ST	Unknown	F	2	-		Putative Taiwanese variety
		T4	Taiwan	台湾野生梅	Taiwan	-	1	174242	*	
		T5	85486	85486	Taiwan	-	1	229937	*	Putative Taiwanese variety
P. mume	AM	AM1	Bungo	豊後	Oita	F	2	-	*	
		AM2	Fushida	節田	Unknown	F	2	-	*	
		AM3	Inabungo	伊那豊後	Nagano	F	1	-	*	
		AM4	Jumbotakada	ジャンボ高田	Fukushima	F	2	-		
		AM5	Kanshikobai	杆子紅梅	Unknown	O	1	-	*	
		AM6	Kurodaume	黒田梅	Unknown	O	1	-	*	
		AM7	Musashino	武蔵野	Unknown	O	1	-	*	
		AM8	Rinshibai	淋朱梅	Unknown	O	1	-	*	
		AM9	Seiyobai	西洋梅	Hokkaido	F	2	172782	*	
		AM10	Taihei	太平	Unknown	F	2	174241	*	
P. mume	SM	SM1	PM1-1	すももうめ中間母本農1号	Ibaraki	F	2	-		*P. salicina* 'Sordum'×'Jizoume'
		SM2	PM1-4	すももうめ中間母本農2号	Ibaraki	F	2	-		*P. salicina* 'Sordum'×'Jizoume'
		SM3	Sumomoume	李梅	Wakayama	F	2	174239		
		SM4	Tsuyuakane	露茜	Ibaraki	F	2	-		*P. salicina* 'Kasaharahatankyo'×'Yosei'
P. armeniaca	Pa	Pa_1[u]	Heiwa_1	平和	Nagano	F	1	174943		
		Pa_2[u]	Heiwa_2	平和	Nagano	F	1	-		
P. salicina	Ps	Ps	Oishiwase	大石早生	Fukushima	F	2	112962		
P. persica	Pp	Pp	Hakuho	白鳳	Kanagawa	F	1	112532		
P. dulcis	Pd	Pd	Almond Wakayama 1	アーモンドわかやま1号	Unknown	F	1	-		

[z] Code description is shown in Table 1.

[y] F: fruits, O: ornamental flowers, -: unknown.

[x] Number of tree entries.

[w] JP numbers from Genebank of the National Institute of Agrobiological Sciences.

[v] *: the same accession used by Hayashi et al. (2008).

[u] Accessions having multiple genotypes at more than two loci.

Table S2. Microsatellite markers used in chapter 2.

Marker name	Forward primer (5'-3')	Reverse primer (5'-3')	Motif	Repeats	Size (bp)[z]	Linkage group[z]	Position (bp)[z]	Reference[y]	Screening[x]
JAM1	AGTCAAACACGGTTGAAGGA	ACAGAGCAGCCTCTTGAACG	AG	29	190	LG1	2179120-2179309	1	3
JAM2	TGTGGGACTCTGGAGAAGTG	TGCTCGCCAAATGAAAATGCA	TC	31	202	LG1	3568514-3568715	1	3
JAM3	AGAGTGACGGAGCAAATGGT	CTGCCAAAAGTCACAAGGGC	GA	18	215	LG1	6201776-6201990	1	3
JAM4	GAGCGTCCAGTCCAATCCAA	CACGGCCATTTCACTTGCAA	AG	19	143	LG1	5670392-5670534	1	
JAM5	TGGCACGAAAATGAAACTGGA	GGTTTCACCGACACAACCCT	AG	20	172	LG1	7220660-7220831	1	3
JAM6	TGCCTAGGAGACGAGCTGAT	GGGGCTGATCAGTGGCTTTT	GA	21	233	LG1	8232652-8232884	1	1
JAM7	TGGGGATCCCTTATCGGACA	AGACCTCGGCGTCCATATTG	TC	33	219	LG1	9201447-9201665	1	
JAM8	CCGCATAGCCACTCATACCC	TGTAGTGCACCTAACAAACCCT	TC	20	247	LG1	10239448-10239694	1	3
JAM9	CTGCAGTGCTAAAAGCTTCCC	TCTCAGCACCCGAGAAGAGA	GA	24	243	LG1	11218630-11218872	1	3
JAM10	CAGTTGAGGACGAGGCAGAT	GGTTTGGTTAGGTTGAGTTGGG	TC	19	225	LG1	12247687-12247911	1	1
JAM11	CAAGGAAAGGGAGTTTACACCC	CGTCACTCTCCAAGCCATCA	GA	21	210	LG1	13523381-13523590	1	2
JAM12	AGAGTCCAAATGCAAGCCCA	TGTGCTAAATTTATTGCTGCCCA	CT	22	118	LG1	14354581-14354698	1	
JAM13	GTGGGCTTTGGATTGGCTTC	ACCCAACAGTCCTTCCCAAC	GA	18	250	LG1	15535904-15536153	1	2
JAM14	AGCCAGTTGATTGACCCAAGT	CCAAGACCAAGTGATGTGGGA	AG	22	200	LG1	16469165-16469364	1	2
PMKS15	ATGAGGTGGCACATGGTTGA	ACGGTGTGTATTCAGGCTGT	CT	20	169	LG1	17666857-17667025	1, 2	4
JAM16	GGAGCTCACTCACCTCAACC	GTTGACTCCAACTCCGATGGA	AG	21	114	LG1	18422693-18422806	1	1
JAM17	ACAGTTTGAACCCTCACTGTCA	TTTGCAAGCAGGTGATGGTG	GA	17	190	LG1	19344078-19344267	1	2
JAM18	TGTTCAATGTCTTAATGCAGCTCT	AGTCGTTTGGGGCCTTTCTT	GA	23	248	LG1	20421148-20421395	1	1
JAM19	CGAAGTGGTGAGGTTGGGAG	GGGTCATTTCCCATCTTCTTCC	AG	39	213	LG1	21252836-21253048	1	2
JAM20	TCGAATCGCTTGGACTCGAG	CGAAACGAGGACCTGGAGAG	TC	37	231	LG1	22305597-22305827	1	3
PMKS21	TATGACCACCACCGCAGAAA	AGGGAAGCGAAGATCTGGGA	CT	22	249	LG1	23254839-23255087	1, 2	4
JAM22	GAACCACTCGTGATGACGGT	TGACTCCATGCCAACCTGAG	TC	19	162	LG1	24390563-24390724	1	1
JAM23	AAAACGGTGTGGTGGGCTAT	ACCCTCTCCTTTGTTTGCTCA	GA	21	180	LG1	25362157-25362336	1	3
JAM24	GCTCCCTCTTGCTATGCTGT	CTGGTGTTTCTGGTGGTGGT	CT	17	175	LG1	26417187-26417361	1	1
JAM25	ACAGGTTCTGAATTCTGATGGCT	AAACCGGCGAAACTTTTGGG	TC	20	129	LG2	214274-214402	1	3
JAM26	TAAATCCAAACGGCAGGCCA	AGAGAGGGTGAGACTGGTGT	CT	21	196	LG2	1237481-1237676	1	2
JAM27	GGCTTTGGATGATTCGCGTG	TCACGTCACTCACTCACTCAC	AG	20	152	LG2	2256073-2256224	1	2
JAM28	TGCACGTAGCTTTCACTACCT	GGTACGCTGACCATCGATCA	TC	20	183	LG2	3307724-3307906	1	1
JAM29	CTGCTCTGATAGGTCCGTCC	AAGAACACCGAGGATCTGGC	AG	18	155	LG2	4321593-4321747	1	2
JAM30	TGTTTACAGCTCTCACCACGT	TGGGACACAGAACAGAGTCG	GA	18	141	LG2	5307576-5307716	1	1
JAM31	AACACACACACGCAAATGCA	CCTGAGGTGCTCCAGAGTTG	AG	21	169	LG2	6201988-6202156	1	2
JAM32	TGAAGCAGCTTCAAACCCTA	CCACTGTCCATGGCTCCATT	AG	22	189	LG2	7281145-7281333	1	3
JAM33	GATTCCCAATCTGCTTACGCA	CCAGACGGCGCTACCTTATT	GA	25	154	LG2	8342421-8342574	1	3
JAM34	TGTGCATGTCCACACACTCA	CTTTAGTTGGGCTGTGGGGT	AG	21	184	LG2	9348437-9348620	1	1
JAM35	TTTTCCCTGTCACTCGGTCG	GTGTGAATTTTTGGGGCTGCT	AG	21	147	LG2	10178044-10178190	1	3
JAM36	ACTAGAGGGAAAGATGTGGGA	CAGTCTCTCTGTATGCATCTGTG	CT	34	169	LG2	11186490-11186658	1	1
JAM37	CTTCAGCAGACTCAACCGGT	CAACCGAACTGCAGCATAGC	TC	18	130	LG2	12448508-12448637	1	3
JAM38	GGAAAGAAACAGCTGCGCAA	AGAAAGAGAATGGCAGCGCT	CT	26	128	LG2	13197827-13197954	1	3
JAM39	AGCGGGAATTGGAAAGCTCA	CCTGCTTAAATGACCCAGTGG	TC	20	176	LG2	14311509-14311684	1	2
JAM40	CCAGGAGACTTTTGCCTGCA	CCGCCCTGCAAAGCAAATTT	TC	17	174	LG2	15327390-15327563	1	1
JAM41	CCATTCATTTACCATTTCGTTTCGT	ACATCGGAGGGTGGTAGGTT	CT	24	184	LG2	16347880-16348063	1	3
JAM42	GGTTGGTGGCTTGAGTTTCG	TCTCACAAATAACCTCAATGGCAC	GA	20	154	LG2	17155024-17155177	1	1
JAM43	ACCTGAATTCCCTGTGAAAGT	GACAGACGCACAAAGACACG	AC/TC	18/14	172	LG2	18396542-18396713	1	3
JAM44	CAACGCTTCCAATCCCTTCA	ACCTACGCTAGAGTTCAGGT	CT	15	200	LG2	19290791-19290990	1	3
JAM45	TGATGCAGTTGAAGGCCACA	ACATTCAGTAGCCCAAAAGCA	CT	24	200	LG2	20182697-20182896	1	2
JAM46	ATGCATGCATGTCTCTGGAC	AATCATGTTCCAGCTGAGGC	CT	22	154	LG2	21346839-21346992	1	1
JAM47	GAGGATGGGGAGGAAGGACT	AGGTATTTGTGAGCACGGCA	AG	20	172	LG2	22285816-22285987	1	3
JAM48	TTACTCCCTCTCCCAGTCCG	ACATGAGGGCAAAGTGGTCA	GA	15	153	LG2	23497814-23497966	1	
PMKS49	TGACAGGTCATCATACCATTTGGT	GCATGCTGACTCCTCCGAAT	CT	20	181	LG2	24293774-24293954	1, 2	4
JAM50	ACCACTGAATTGGGCCCAAA	ACCCTTGTTCACCTTGAGAAACT	GA	22	242	LG2	25368436-25368677	1	2
JAM51	GCTGCCAAAACTCTGCGAAA	TGAACCAATCGACTGGAGCC	TC/TA	21/12	153	LG2	26412807-26412959	1	3
JAM52	TGCTGTTAAGATGAGGAGTCCG	CCATATTCCGCCTCTTCCCC	AG	27	157	LG2	27276081-27276237	1	1
JAM53	GTGATGTGCATGCCTGAAGC	GGGCGTCACTCCTTCATCAA	GA	18	136	LG2	28272822-28272957	1	2
JAM54	GGTGGAAGTAAAGGGGCACA	ACATGGCTTCTGCACGATAA	TC/TA	26/13	197	LG2	29183458-29183654	1	1
JAM55	TCCCCTTGTCTTTGATGGCC	CAGCAGCCAGGCTCAAGAAT	CT	18	164	LG2	30230224-30230387	1	3
JAM56	GCTGTTTCTGCATATGGGCA	AGCTGATTGTTGATTGGAGTGC	GA	21	184	LG2	31159318-31159501	1	2
JAM57	CCAAAAAGTTCCACGTCCACA	GACCTAAGCGGCTGAGGTTT	AG	20	145	LG2	32174264-32174408	1	2
JAM58	AGCTCCAAATGAGGAAGAGGTG	GAGTGCCTGCGTGTTTGTTT	TC/AC	14/16	155	LG2	33311970-33312124	1	1
PMKS59	GCCCTTTAATCCCAAGGAAGC	AACGAGCCCTAGGGTGTTTG	AG	20	197	LG2	34368689-34368885	1, 2	4
JAM60	CCCGGTTCCAGACACCTAAC	TGATGAGAAGGGGCTTGGTG	CT	23	191	LG2	35394151-35394341	1	1
JAM61	AAATCTCCCCTTGCCAACCA	TAGGCATGTGGTGGTTGGTT	AG	24	188	LG2	36425196-36425383	1	2
JAM62	TGGATCTACCCGTCATACTCCT	ATGGGCGAATGAGTGAGGTG	TC	18	232	LG2	37509480-37509711	1	2
JAM63	TCGATGCCTGCTGACATGTT	TCATTGCACCCAACAGTGGT	AG	18	156	LG2	38396727-38396882	1	3
JAM64	GCACTGCCAGGAGGAACATA	TTGGGACCAACAGAGCACAT	CT	19	144	LG2	39523431-39523574	1	1
JAM65	CCAGCGGATGAGACTTCAGG	AAATAGTACGCCGAGCTGGG	AG	18	144	LG2	40217724-40217867	1	3
JAM66	TCGGTTGTCATCACAAGTCCA	CCCAACAAACCACAACAACGT	CT	24	160	LG2	41203062-41203221	1	1
JAM67	GGACCAACCAAACAACGTCC	ACTCGCGAAACTGATCAGAA	CT	17	209	LG2	42052538-42052746	1	3
PMKS68	GCAACGTGAGGAAGAGAGGA	CCTTTCATGCAGTGGAGTGG	TC	23	210	LG3	267476-267685	1, 2	4
JAM69	AACCAAAGCCTACCCAACCC	GTGTGAGAGAGGAGCAACCA	CT	24	190	LG3	1298899-1299088	1	3
JAM70	GGACAGCAGCCTTGTTCCTA	TCCAAACCAGCCAAGCAGAA	CT	19	187	LG3	2326490-2326676	1	1
JAM71	TCATGTGACTCCTCTCTCCCT	GGAGTCTCCTCGGAGGTCAA	TC	33	205	LG3	3367061-3367265	1	3
JAM72	TGGCAAGAGACAGCTTCTAAG	CTTCCTTCCAACCGGTCACA	GA	22	192	LG3	4449388-4449579	1	1
JAM73	CCTCTGCGTGTGTCTCTCTG	ACTTGAGAGCAGGTGACACA	TC	23	187	LG3	5322827-5323013	1	2
JAM74	TTGCCAGATCCGTTTCCTCC	ACGAATCACCCACCCAACTC	AG	26	186	LG3	6299897-6300082	1	2
PMKS75	TGGGGAGTTCCTGTCCATGA	AGCACCAGTTAACACCAGCA	CT	20	163	LG3	7181901-7182063	1, 2	4
JAM76	TCGCGAAGTGGACTCCATAC	CCCCTTCTCAACGAGCAACT	TG/AG	11/18	246	LG3	8417458-8417703	1	1
JAM77	CCACAAAACCAAGGTGCCAA	GCTGATTTTGAGGATCGAAGGG	GA	28	150	LG3	9227802-9227951	1	3
JAM78	AGGCTTCTTTCCAATCCACTGT	TCTGCAATCCGCTCAGCTAG	TC	19	172	LG3	10158492-10158663	1	1
JAM79	CCTGGTTCCTGTATCCCCAC	GATCCTCTGCTACAGCCAGC	GA	21	184	LG3	11359706-11359889	1	3
JAM80	TAGGCATTTCCAGTGAGCCC	CGAACCCTGTTCATGTTGCA	AG	21	177	LG3	12359405-12359581	1	2
JAM81	AGACGTGGTGGTTGGATCAC	GCCAAGGCCCCTACATACTC	TC	17	189	LG3	13330809-13330997	1	3
JAM82	ATGTGTTCATGGCTGTGGGG	TGCAGCCACCATCCTCAATT	AG	17	167	LG3	14349547-14349713	1	1
JAM83	ACAATTATCATGGCATCCACTGT	AGTCTGGCCAACTGTTGTGT	AG	27	192	LG3	15138383-15138574	1	2
JAM84	ACCACCTCCCAAACCTGAAC	TGGCCTCTTAGTTGAGTGTCAG	AC	23	180	LG3	16214809-16214988	1	1
JAM85	CAGCTTTTGGGTGGGTCGTA	TCTCCTTACTTTGCCGGTGC	AG	18	178	LG3	17251679-17251856	1	2
JAM86	TGGAGTTTTACGGAGTAGCAGG	CGGGCTGGTACAACTGTGTA	AG	24	248	LG3	18353841-18354088	1	3
JAM87	TGCGTTTCTTTTGGGGTTTGA	CTTGCACTTCACGAGGTTGC	GA	18	148	LG3	19205677-19205824	1	2
JAM88	GCTCTTCCATATTCCCACCCC	GCAACACACCAAGAGCTTCG	AG	17	117	LG3	20465015-20465131	1	1
JAM89	CACAAAAACACGAGCACCCA	GGTCACCTCACCTCCTCCAT	AG	20	175	LG3	21294811-21294985	1	2
JAM90	TAATCATGCGTGCGTGTGAC	AGCAGAGCTTGTTCTGTTCA	AG	18	131	LG3	22242299-22242429	1	
JAM91	GCAAGGGAATTTCAGCTTTGC	GGTGTAGCTGCTGCATTTCG	AG	22	184	LG3	23248370-23248553	1	3
JAM92	AAGCCAAGCACAAGCACTGA	CCCGTCATGAAGCAGCCTAT	GA	13	172	LG3	24016459-24016630	1	2

Table S2. (Continued).

Marker name	Forward primer (5'-3')	Reverse primer (5'-3')	Motif	Repeats	Size (bp)[z]	Linkage group[z]	Position (bp)[z]	Reference[y]	Screening[x]
JAM93	GCTGCACTTCAAATTGGGACA	TCTGCACCTTGACGGACAAT	TC	22	242	LG4	136852-137093	1	2
JAM94	ATGCCAGCAGATCAGATGGG	ACATGCATGCAGGTAACCCA	GA	22	201	LG4	1322230-1322430	1	1
JAM95	ATCCGGCTTGAACTCAGCAA	GCACAGGCAATGTCATCTCT	AG	25	208	LG4	2434398-2434605	1	2
JAM96	GATTGGCTAGAGCGGCGTTA	CAACTTGGACCACACAATTCCA	GT	38	218	LG4	3391073-3391290	1	
JAM97	TACTCAGTGGGCTTCCAGGA	AGCTATCCATATTCGGCACCA	CT	20	182	LG4	4168587-4168768	1	2
JAM98	GGCCTATTTGTGCCTCCCTT	AAACAAATGGTGGGCTCCCA	AG	20	198	LG4	5252512-5252709	1	3
PMKS99	TCACCCCTCCTCTCTCTTGG	TGCAGAAACATTTGGTGGGA	TC	21	170	LG4	6286204-6286373	1, 2	4
JAM100	TACTGCTCGTCGCTCAATGT	TTTGCCCTGCTTCTAGTGCT	AG	23	131	LG4	7347329-7347459	1	1
JAM101	AGCTGACACCAAACCCAGTT	TCGATACCACAAGGAGCACAC	AG	18	114	LG4	8260126-8260239	1	3
JAM102	ACCCATGTCCTGTTTTCGACA	CACCTCCATTCACCAAACGC	GA	20	198	LG4	9138533-9138730	1	1
JAM103	TCTAGCTTCTTAGGGCCCGT	CTGAATTGGGTGCCATATGCG	CT	18	177	LG4	10223442-10223618	1	2
JAM104	CATGCGGCACCCTCAATTTC	GAGAGGGTGGCATGAGTGAG	TC	20	188	LG4	11603573-11603760	1	2
JAM105	ACAAACACTGTGGGTCTGAAGA	CTTCAACTGCTGCCATTTGCT	GA	22	122	LG4	12218018-12218139	1	2
JAM106	TGCCAAATATAGGTTGCCCGA	AGCTTGATATGCTGCCGTGT	GA	20	179	LG4	13177473-13177651	1	1
JAM107	GTGTTCCCTCTGATTTGCAACA	GCACCTCGCATGCTCAATG	TG	24	186	LG4	14231943-14232128	1	3
JAM108	ACGTCAGCAATAGAGCAAGCT	GGATGGTTGCTGGCAGAAGA	TC	27	166	LG4	15303819-15303984	1	1
JAM109	AGCTAGGCTGGGAGATCACA	TGCTTCAGTGTGGTGAACGT	AG	22	156	LG4	16212940-16213095	1	3
JAM110	GTGCCTGTTGTGCAAGTTTCT	CCTGAAACTGCTCTGAGGGG	TC	24	180	LG4	17189319-17189498	1	2
JAM111	ACAAGTTTCAGCGTAAATCGG	ATCCCTCTTTTCCGCCCAAG	GA	19	125	LG4	18178025-18178149	1	2
JAM112	AGAGGACTGCAGGGGTTAGT	GTTGTGTGCATCTCACTCGC	GA	18	115	LG4	19323110-19323224	1	1
PMKS113	TCTTCTAAGTCAAATGCCTGCCT	TAAGGGATACAGCGGGGTCA	GA	19	161	LG4	20315713-20315873	1, 2	4
JAM114	TGACAAACGCGAGGTAGGTT	CCGCAGTTGAGTCTTCACCA	CT	27	202	LG4	21226349-21226550	1	1
JAM115	GGAGGGCTGTCGGTTTCTTT	CTGAAATGCAGCTTCCGCTC	GA	20	172	LG4	22250612-22250783	1	3
JAM116	TGAAGAAGAGCAAGCTACTAGC	CACCACACACGTCTCCTTCT	AG	26	222	LG4	23202622-23202843	1	2
JAM117	TCCTTCGCACCAGCAATAGT	GGGATAATGGAGCCAATTGGGA	CT	19	148	LG5	171147-171294	1	2
JAM118	AGGTAAGAGAGGTGGTGGCA	TCCTCCCTCAATTGGAAAAGGA	GA	20	192	LG5	1338402-1338593	1	1
JAM119	CAAATGCTCGTCCCAAAGGC	GGCAACGACTTGACAGCTTG	GA	18	179	LG5	2489328-2489506	1	2
JAM120	CCATCCCTGACCCACTCTCT	GCGTGGCACATATGTGGAGT	TC	18	198	LG5	3419597-3419794	1	1
PMKS121	AGAAATGCGGAGGTGTAGTGT	GACCTGCAGACAAAATGAGCA	TC	26	202	LG5	4271799-4272000	1, 2	4
JAM122	AAAACCCTAGCCTCCCTCCT	TGAGCATGGAAATGGGGACT	TC	20	196	LG5	5296402-5296597	1	2
JAM123	CCCTCAGTGGCACAACTGTA	CAAAATGTGGCCGACTGTGG	TC	25	162	LG5	6182205-6182366	1	2
JAM124	GCTCCCTTGCTTTTCTTCTTCC	GTGAACCAAAGGGGCCTGTA	CT	19	123	LG5	7183602-7183724	1	1
JAM125	ATCTCACCACCCCATCCTCA	AGAAAAGAGGGGCGGATTGG	TC	18	175	LG5	8431352-8431526	1	2
JAM126	GAACGTAGTCGGCCATCCAT	CCTCCTTTTGCCTTCCCCAT	AG	20	129	LG5	9301008-9301136	1	1
JAM127	ACCCAAAGGTCCCAACTTGA	CTCTCTCCCGTTCGCAACTC	GA	18	142	LG5	10335534-10335675	1	2
JAM128	AATGTCAAAATGCCCCCAGC	CCAACCCTGAAGCCTTCCTC	GA	16	233	LG5	11440674-11440906	1	2
JAM129	TCTCACCTGGCTCATGTGTG	GCTCGCATTTGACGTTGTGT	CT	19	131	LG5	12085054-12085184	1	2
JAM130	AGTTTCTCATGTGCAGCCCA	CTGCAGTGCAAGATGTTCGG	CA	21	113	LG5	13211057-13211169	1	1
PMKS131	GCACCATTCTTTACCACCGC	ACAAGTCAAGTGAAACTCTGCA	CT	18	178	LG5	14199800-14199977	1, 2	4
JAM132	GGAGGAAAAGGTGAGCCACA	ATTCCATCCCCTACTGGCCT	GA	20	158	LG5	15161940-15162097	1	1
PMKS133	CCACTCACAGATCGACACGT	GTCAGGTTTGCTGCTGGTTG	TC	20	191	LG5	16105876-16106066	1, 2	4
JAM134	ACTGCATTAGCCATGGCAGT	TGCCAAATGCTGCAGATCAT	TC	21	162	LG5	17189988-17190149	1	3
JAM135	TCATAGTGCATACGGCTGCC	GGACAGATCAGATTCCCTTGTGT	CT	24	157	LG5	1825524-18255405	1	2
JAM136	ACACCCAAGAAGAGCATACTGA	AGTTGAAGAAGCTCCCCATCA	GA	19	186	LG5	19439133-19439318	1	1
JAM137	TTGCAAGGTCTACGGTTGGT	GGCCTGCAAATGGTTATCGT	CT	23	193	LG5	20374324-20374516	1	
JAM138	CACTCCTTCCTTGCACCACA	TGAAGCCAACTGCCGAGTAG	CT	17	137	LG5	21402184-21402320	1	1
JAM139	CACAGTAGTCCATGTCGCCA	TGGATGACAGATGTGCACAA	TC	24	246	LG5	22324945-22325190	1	2
JAM140	CTGCAGCACATGTACAACAGG	TGACCACCAGAAGAGGCTTC	CT	19	187	LG5	23351037-23351223	1	2
JAM141	TTCAGTGCACCAAACCACAC	GCCCCCAGTATACGTAAGGC	AG	25	154	LG5	24266401-24266554	1	2
JAM142	CAGCATATGCAACTGCCATGT	TGTGCTCTGATCAACCAGGTC	AG	20	169	LG5	25035249-25035417	1	1
JAM143	ATAGCAACTGTGGGCTGCTT	ACCGGTCTCTTCCATGATGC	CT	17	194	LG5	26099505-26099698	1	2
JAM144	CCTTCAGTGTGTCCTGCCTT	TGTACTAGCCATCCCCACCT	CT	20	165	LG6	432807-432971	1	1
JAM145	TTTGGAGTTTGGAGGCCACC	TGCGTTTTTGTTAGGGCCAA	CT	20	228	LG6	1082199-1082426	1	2
JAM146	AACCAGGAGTCAAGCGTCTG	CGCGGTAGATTTCCACGGAT	CT	20	189	LG6	2412435-2412623	1	3
JAM147	TCAAGTCATCACCTGCAAAGGT	TTATTGGGTTCCCAGGCCAC	CT	17	214	LG6	3128746-3128959	1	
JAM148	GCAAAGCTTTTCTGGCCCAT	TCTCTGGGAAATGCAAACCTG	AG	29	223	LG6	4080821-4081043	1	1
PMKS149	AGGATCATGGGCAGAAGTTGT	AAGAAGTTGGACTGGGTGCC	AG	19	165	LG6	5343510-5343674	1, 2	4
JAM150	CGAGTAGTTGTGGTGCGGAA	CCGGATAACCCACCCGAAAT	AG	19	150	LG6	6308340-6308489	1	1
JAM151	GTGGTGATCAAACGCATGCA	AGGACCTCTCCCTCGCTAAT	GA	24	198	LG6	7167905-7168102	1	2
JAM152	TAAGGTGGGAGCCATGCATG	AGGGCCTTGACCAAAACCAA	TC	29	228	LG6	8587400-8587627	1	2
JAM153	TGACCCGGTGTGAACTAAGC	GCACTCCTTGTCGGCTTCTT	GA	22	233	LG6	9077115-9077347	1	2
JAM154	GGCACTCATACGTATCACACCA	TACCGTGATTTCTCTCGCGC	TC	26	153	LG6	10203561-10203713	1	1
JAM155	ATGCTAGCTGCCCACAAGTC	GGTCTCACCCCAACCAGTTT	GA	18	164	LG6	11120581-11120744	1	3
JAM156	AGCACCACAGAGAAGAGCTT	GCTCTTTTTGGAAGGCGAGG	CA	17	103	LG6	12371465-12371567	1	1
JAM157	GGGTTGGAGATGGCCTCAAA	TTGCTAAGTTGCCCACCCAT	TC	25	182	LG6	13018248-13018429	1	2
JAM158	GCCATCTCCACCCGTCATAG	GGTTATGCCCCAAGTTCCCA	TC	21	156	LG6	14094784-14094939	1	2
JAM159	CACTACGAGTGTGCGGATGA	AGCAGGTGTTCTTGCACAGA	CT	26	143	LG6	15192734-15192876	1	
JAM160	ACGAATCGCCTTTCAGTGCT	TGCTGCAGATCCATCAGTCA	AG	20	184	LG6	16562749-16562932	1	1
JAM161	TGGGAATTAGGGTTGGGAATCA	TGAGATGACACCCCCATTGC	GA	19	155	LG6	17429614-17429768	1	2
JAM162	ACGAAGGATTGTTGTACGTTTCTG	TCCGTCAAGACTTGCAGCAT	TC	25	172	LG6	18025228-18025399	1	1
JAM163	TGTGAGGGAGAGGGAGAGTT	CGCCCAACCAGCTTTCAAAT	GA	21	176	LG6	19501830-19502005	1	2
PMKS164	AAGGGAGGGGACTTCGGTTA	CATCCCAGACGCAATCACCA	AG	22	111	LG6	20105981-20106091	1, 2	4
JAM165	GTGATGACACGAGTGGAGGG	GCAGTTGTGGAGTTGCTTCAG	TC	23	190	LG6	21033155-21033344	1	2
JAM166	AACTCGTGGCGCATCTTGTA	TGTAATGCGGGTCCAGTTGG	TC/AC	18/20	150	LG7	200229-200378	1	1
JAM167	GGAGCCCAACACATCGAAGA	AAGTGCAGGATGGAGGCAAA	CT	20	175	LG7	1138789-1138963	1	2
JAM168	CGGCAAGCCTCTCATCATCA	CCGGAAGTGCTTTCTTCTGC	AG	23	234	LG7	2353698-2353931	1	1
JAM169	GCCCACCTCAGATGCCATAT	GGGAGCCATGAAGAAGAGGA	TC	22	166	LG7	3285746-3285911	1	3
JAM170	CGAAAGCCATCCTATCAGGGA	TGCCCTAGATGTGCTTATCCT	AG	25	166	LG7	4320847-4321012	1	
JAM171	GGTAGCGCAGAAAAAGAGCAG	TCCAAGGAAACATTGGCCACT	TC	20	198	LG7	5224992-5225189	1	2
JAM172	CGTCCACCCCAACTTGATCA	AGCAGAGCAGCACAAGTGTA	TC	27	238	LG7	6092033-6092270	1	1
JAM173	CTCTTCCCCCTCACCAAACC	GCTAGCTTCTTGGTCACTTCCT	AG	20	181	LG7	7218802-7218982	1	2
JAM174	GCAGAAAAGCCACATTGCCA	CTGGGTATCAGCAGGTGGTC	GA	19	165	LG7	8248170-8248334	1	1
PMKS175	TGGTTAAGCCACCCATGAGAG	AGCTTTTCCAACCACAACTCA	GA	20	169	LG7	9257791-9257959	1, 2	4
JAM176	CTATGGTCATCGTCCCTGCC	CGAATCACGTCTGCTAGGGA	TC	22	183	LG7	10203265-10203447	1	2
JAM177	ATGGCCCACGTAACGTTGAT	ATCACGGGAATGTGGTGCAT	TC	23	108	LG7	11367769-11367876	1	2
JAM178	AGCACCGACAGTCTTGTGAG	GGCCTTAAATGAGCAAGTGTCC	CT	36	183	LG7	12413008-12413190	1	1
PMKS179	AAATCTCAGGTTGCGTCCCC	ACCCCAAAATGTCATCATCGA	TC	25	178	LG7	13259694-13259871	1, 2	4
JAM180	GCTCGTATGCCAGACCCTTT	TGACATGAAGTGATTGCGCA	GA	25	224	LG7	14302280-14302503	1	1
JAM181	TGATCTGTGGCCTGCTTAGC	ATTTTGCCTGCGTTGTGTGG	GA	22	153	LG7	15191329-15191481	1	2
JAM182	GCCCTCAACAAAACCCAGGT	ACTTCAAGGTCCTGCAATGGA	TC	25	196	LG7	16236524-16236719	1	2
JAM183	TCACCCAAATCCAGCCATCC	TTCGCTGCTTTTGGTGTTGG	TC	22	110	LG7	17008151-17008260	1	2

Table S2. (Continued).

Marker name	Forward primer (5'-3')	Reverse primer (5'-3')	Motif	Repeats	Size (bp)[z]	Linkage group[z]	Position (bp)[z]	Reference[y]	Screening[x]
JAM184	GCAGTTTGCTCTGTTGCCAA	GTTGCGGTTTTTGCTGAGGT	GA	21	180	LG8	243080-243259	1	1
JAM185	TATCGGAGGGTCTTCAGCCA	CACTTCCCTTTGGCAGGCTA	CT	19	132	LG8	1046459-1046590	1	2
JAM186	CAGTTGAAGCAGTCGCAAGC	ACCATGCTGAATCGATATCGGT	GA	21	163	LG8	2011234-2011396	1	1
PMKS187	TCCCAAACCACCCTTTTCCC	GGGTCTGCAGTCTACACAGG	CT	18	182	LG8	3127817-3127998	1, 2	4
JAM188	TGGATTGAATTACAACGACCGT	AACCAACTGCGGAGAAACTG	CT	23	153	LG8	4199626-4199778	1	
JAM189	GGTTTTTGGGCATGAGGATTGA	CTTCTCTCCACAGCCACCTC	GA	26	189	LG8	5186984-5187172	1	
JAM190	AGGCGTTTTGGTTACCGGAT	GCGCTGAATGGCACCCTATA	CT	26	173	LG8	6057770-6057942	1	
PMKS191	ATTGTGACCCAAGTCCGCAT	TGCACGGGGTGGTGTACTAT	AG	20	215	LG8	7245191-7245405	1, 2	4
JAM192	AGCATTCTAGGGCATCCTAAGA	ACTCACACTTAGCACCCACC	GA	24	232	LG8	8144650-8144881	1	1
PMKS193	TTTCTGTGCCTTGCATTGCC	GCTAGCTGGTCATTGTTGCG	CT	21	187	LG8	9038780-9038966	1, 2	4
JAM194	AAGCTGCCGAACAAGAAAGG	AAGGTCCAAGGCAGTGACAC	AG	27	196	LG8	10108041-10108236	1	2
JAM195	GCCAAATAAAACTCGCTGCCA	TTTGCCACCAGACCAGCTAG	GA	23	200	LG8	11296774-11296973	1	2
JAM196	TTGCAGGCCCAGTTCTAGAC	CGGTAAGAGGGCTGAGAACC	CT	22	183	LG8	12196580-12196762	1	1
PMKS197	CCTCAGGTTCACAAATTGCACA	ACACAGGTATCTGGTCCCGA	GA	20	175	LG8	13246995-13247169	1, 2	4
JAM198	CCCCTTGCCAAGCATCTAGT	TCCAGCAGGTTTTGTCTCCA	AG	25	179	LG8	14323081-14323259	1	1
JAM199	TGAGCCTCTGCAAATGAAGACA	ACACTCGCTGGTCCTGATTG	AG	22	152	LG8	15154914-15155065	1	2
JAM200	AGCCTCTTGCATTGGTCACA	AATTAGGGGCCTGTGCTGTG	AG	25	173	LG8	16080318-16080490	1	2
PMKS201	TCAACTTCTCTCTCGCGCTG	GTGCTTGCTTACACCAGTGC	AG	17	188	LG8	17079257-17079444	1, 2	4
UDP96-001[w]	AGTTTGATTTTCTGATGCATCC	TGCCATAAGGACCGGTATGT	CA	12	112	LG1	6460524-6460635	3	-
pchgms3[w]	ACGGTATGTCCGTACACTCTCCATG	CAACCTGTGATTGCTCCTATTAAAC	CT	15	188	LG2	16497572-16497759	4	-
MA007a[w]	GTGCATCGTTAGGAACTGCC	GCCCCTGAGATACAACTGCA	GA	20	112	LG5	19161124-19161235	5	-
MA017a[w]	AAGGCATATAGCGCAGGT	ATCTGAGGCCTTCAACACTT	Unknown	Unknown	Unknown	Unknown	Unknown	5	-
MA040a[w]	AGAAATTGGAGTGACGTAAC	ACGTGATGAGAAGTAGGGAG	GA	9	205	LG1	23029242-23029446	5	-
M6a[w]	AGAAGGGCAAGCCCAAGTGC	TGCAAAGCCAGAGCCCACAA	TC	12	182	LG6	7579781-7579962	5	-
M7a[w]	GAAGAAAGACTGAAACAACG	CCAGTTGAGAGTGTCTTTGA	TC	17	154	LG6	10518674-10518827	5	-
PaCITA4[w]	GTGAAAATGAAAGAATCGCTACC	TGTCCCTTGACGCCCAGATTTCTCC	GA	16	150	LG4	15967804-15967953	6	-
PaCITA7[w]	CTTTTGTGCCTCAGCTTCCCAACAC	CCTGGCCTGACCCTAAGCAATTCG	CT	21	236	LG2	14873383-14873618	6	-
PaCITA19[w]	GACAAATACAATCAAGAAGTGTCGC	GAACAGCTAGCCCCTTTGTCATAC	TC	15	115	LG5	10221597-10221711	6	-
PaCITA21[w]	GATTATATAAGTTGGTTTTTGTAAG	GTATTCTATAATGTATAAATGTACG	CT	14	222	LG7	8480253-8480474	6	-

[z] After *P. mume* reference genome sequences by Zhang et al. (2012).

[y] 1: the present study, 2: Ishio et al. (patent pending), 3: Testolin et al. (2000), 4: Sosinski et al. (2000), 5: Yamamoto et al. (2002), 6: Lopes et al. (2002).

[x] 1: amplified in the 1st screening, 2: examined until the 2nd screening, 3: examined until the 3rd screening, 4: selected as highly polymorphic markers (renamed as PMKS markers).

[w] Markers previously reported by Hayashi et al. (2008).

Table S3. Amplified fragment allele sizes at 20 highly polymorphic microsatellite loci determined among 128 *Prunus* accessions (199 trees).

Species	Code	Name	Identity[z]	PMKS15	PMKS21	PMKS49	PMKS59	PMKS68	PMKS75	PMKS99	PMKS113	PMKS121	PMKS131	PMKS133	PMKS149	PMKS164	PMKS175	PMKS179	PMKS187	PMKS191	PMKS193	PMKS197	PMKS201
P. mume	F1	Ankoyabai		172/178	266/272	199/206	226/239	214/216	212/226	178/194	183/199	208/224	202	205	179/203	127/131	199/205	135/149	199	239	218	206	204
	F2	Aojiku	A	166/168	258/270	213/215	231/257	226/231	198/214	182/188	199/205	220/226	214	203/209	189/205	111/131	171/185	147/159	213/228	214/235	196	212	194/204
	F3	Benisashi		166/172	252/266	211/214	237/239	233/237	210/226	194	199/207	208	194/202	205/223	189	111/131	173/185	147/151	199/207	214/237	216/218	206/210	206
	F4_1[y]	Fudono_1		184/209	266/282	211/213	239/241	212/233	210/228	194/196	199	214/224	212	199/227	189	111/127	173/185	147/153	199	214	196	206/247	204
	F4_2[y]	Fudono_2		166/184	266	213/214	239	237	199/226	190/196	197	214/224	206	199/221	189	111	173	137/151	199	214/251	204/216	210/212	204/206
	F5	Fukuju	B	162/168	270/272	213	239/257	212/224	214/216	188/194	183/199	198/230	200/214	203/209	189/205	129	171/185	133/137	205/230	214	204	210/247	204
	F6	Garyubai		166/209	270/286	199/214	226/253	231/237	199/214	178/188	191/199	214	194	221/223	189	111/127	185	137/151	199/207	239/263	196/218	200/212	204
	F7	Gojiro	B	162/168	270/272	213	239/257	212/224	214/216	188/194	183/199	198/230	200/214	203/209	189/205	129	171/185	133/137	205/230	214	204	210/247	204
	F8	Gyokuei	B	162/168	270/272	213	239/257	212/224	214/216	188/194	183/199	198/230	200/214	203/209	189/205	129	171/185	133/137	205/230	214	204	210/247	204
	F9	Hachiro		172/209	258/282	199/211	239	233	210	194/196	199	224	194/208	207/223	189	127/131	185/207	137/149	199/207	214	204/206	206/210	204/206
	F10	Hanakami	C	172/184	266/282	213	239	212/237	210/226	178/196	183/197	198/214	-	199/227	189/203	127/131	185	137/147	199/205	214/237	196	212/247	194/204
	F11	Jizoume		166/209	274/282	211	239/241	212/233	210/214	194	197/207	224	206/212	199/207	189	111	173/207	147	207	251	204/216	206	204/206
	F12	Juro		209	266/272	206/211	226/239	212/214	214/226	192/196	183/191	222/224	194	209/225	189/191	129	185	147/149	199	214	216	206/249	204
	F13	Kagajizo		166/168	272/274	211/213	239/257	212/233	214/216	194	183/207	198/224	200/206	207/209	189	111/129	171/207	133/147	205/207	214/251	204/216	206/247	204
	F14	Kahoku		166/184	258/276	204/215	226/231	212/226	198/214	184/188	191/199	208/226	202/214	203/217	195/205	111/131	173/185	147/159	207/228	214	190/196	206/212	194/204
	F15	Kaidarewase		166/209	266/282	214	239/243	212/237	199/214	194/196	183/199	224	194/202	205/224	189	111	185/195	147/151	199	214	204/216	206/210	194/204
	F16	Kairyouchida		166/209	254/270	211	241/253	212/233	210/214	184/194	199/207	198/224	206	199/221	189	111	173/197	147	199/205	237/263	188/196	206	204/206
	F17	Kensaki		166	252/282	199/214	253/255	212/233	194/226	194	199	198	202	205	189	111/131	185	147/151	199/207	214	204/216	210	204/206
	F18	Kinotakara		166/209	266/282	199/214	226/239	212/233	194/226	194	197/199	224	206	205/219	189	131	173/185	147	199/207	214	204	206/210	206
	F19	Kinyuji	D	166/172	258/282	211/215	239/257	212/231	214/226	188/196	199/205	224/226	214	203/207	189	127/131	171/207	137/159	207/228	235/251	188/196	206/212	194/204
	F20	Kodama	D	166/172	258/282	211/215	239/257	212/231	214/226	188/196	199/205	224/226	214	203/207	189	127/131	171/207	137/159	207/228	235/251	188/196	206/212	194/204
	F21	Kotsubunanko		166/209	266/268	211/214	239/241	224/237	199/226	194	197/199	208	202/212	199/227	189/193	111/127	195	137/147	207	237	196/216	206	204
	F22	Kushino		168/184	266/274	211/213	239	233	194/210	184/196	199/207	198/224	202	199	189	111/131	185/194/207	147/153	205/207	237	188/216	206	204
	F23	Misato1	B	162/168	270/272	213	239/257	212/224	214/216	188/194	183/199	198/230	200/214	203/209	189/205	129	171/185	133/137	205/230	214	204	210/247	204
	F24_1[y]	Naniwa_1		170/209	266/270	214	241/253	212/237	199/214	184/190	199	214/224	194/206	209/223	193/203	111	185/195	137/149	199/207	237/251	216/218	210/212	204
	F24_2[y]	Naniwa_2		170	256/272	206/211	226/231	214/216	212/226	188/190	191/199	224	206	203/223	189/205	127/131	171/185	133/149	230	214/237	190/218	206/208	194/198
	F25	Naniwahitoe		180/184	252/258	206/211	226/239	224/237	194/226	194/196	183/199	208/224	-	209/223	193/205	127/129	185/197	147	199/213	214/237	196/220	206	204
	F26	Nanko		209	266	211	239/241	233/237	210/226	194	199	208/224	210/212	199	189/199	111/131	173/185/194	147/153	199/207	214/251	196/216	206	204
	F27	NK14		166/209	266/282	199/211	239/255	212/233	210/228	194	199	198/224	202/212	199/205	189	111/131	185	151/153	199/207	214/251	196/204	206/210	204
	F28	Okunoume	B	162/168	270/272	213	239/257	212/224	214/216	188/194	183/199	198/230	200/214	203/209	189/205	129	171/185	133/137	205/230	214	204	210/247	204
	F29	Oshuku		166	258/266	211/215	226/231	226/233	198/228	178/182	199/205	198/226	206/214	203	203/205	111/131	173/185/198	147/153	213/228	214	196/204	210/212	194/204
	F30	Ozaki		184/209	266/282	199/213	239	212/237	212/228	190	199/207	224	202	205	189/203	127/131	177	147/149	199	214/237	196/218	210/247	206
	F31	Rinshu-Fukui		172/180	245/282	199/213	226/237	212	210/226	178/190	189/197	198/224	206	207/227	189	111/127	185	137/147	205/207	214/237	196/216	206/247	194
	F32	Rinshu-Nara	C	172/184	266/282	213	239	212/237	210/226	178/196	183/197	198/214	-	199/227	189/203	127/131	185	137/147	199/205	214/237	196	212/247	194/204
	F33	Sadayuume	E	209	266	211	239/241	233/237	210/226	194	199	208/224	210/212	199	189	111/131	173/185/194	147/153	199/207	214/251	196/216	206	204
	F34	Sakamoto		172	266/284	206/213	251	212/226	214/216	188/194	183/189	198/214	202/222	190/211	189/191	125/127	173	147/153	213	233/237	204/216	206/212	206
	F35	Shiratama		166/209	274/282	211	239	212/233	210	194	199	224	206	207/221	189	111	177/202	147	207	214/251	216	206	204
	F36	Shirokaga	B	162/168	270/272	213	239/257	212/224	214/216	188/194	183/199	198/230	200/214	203/209	189/205	129	171/185	133/137	205/230	214	204	210/247	204
	F37	Shisen		168/172	260	211/215	237/257	212/226	194/214	194/196	183/207	198/226	202	190/227	193/195	125/127	185	147/151	199/213	214	196/198	206/210	204/206
	F38	Suiko		166/170	256/266	211/213	226/257	212	214/216	194/196	199	198/226	214	203/221	205	111/129	177/185	147/149	205	214	190/204	206/212	194/204
	F39	Taniguchikobai		166/180	258/282	211	239	233	210	194	199/207	224	206	207	189	111	185/202/207	147	207/230	214/237	196/204	206/210	204
	F40	Tenjin		184/209	266	211/213	239/241	233/237	199/210	190/194	197/199	214/224	212	199/221	189	111/131	173	151/153	199/207	251	204/216	206/210	204
	F41	Tojikobai		209	258/268	211/214	239/251	233/241	194/210	184	199/207	224	202	205/221	189	131	173/195/200	137/147	205/207	237	204/216	206/210	204/206
	F42	Toko		209	266/282	211	239	233/237	214/226	194	197/199	208/224	206/212	199/207	189	111	173	147/153	199/207	214/251	196/216	206	204/206
	F43	Yakushi		166/184	266	214	239	212/224	194/210	190/194	199/207	214/224	202	205	193	111/129	185/197/200	137/147	199/205	214	196/204	210/212	204/206
	F44	Yogo1	B	162/168	270/272	213	239/257	212/224	214/216	188/194	183/199	198/230	200/214	203/209	189/205	129	171/185	133/137	205/230	214	204	210/247	204
	F45	Yosei		166	266/282	199/213	228/239	212	194/210	178/190	199/207	214/224	-	223/225	189/203	111	173	137/151	199	214/251	188/204	212	204
	F46	Zaronbai		166/190	252/276	208/211	239/243	231/241	194/214	196	197/207	222/226	202/206	205/208	193	111/135	185/195	133	203/207	214/233	204/220	200	204/206
P. mume	FS1	Benio		172	282	213	239	237	210	190	201	214/224	206	199/227	189/201	111	173/185	147	207	214	196/216	206/212	204/206
	FS2	Hakuo		172/209	270/284	206/213	231/239	212/224	216/226	178/196	199/201	224	194/210	217/225	183/189	127	185/207	133/139	199/207	214	188/216	206/249	194/204
	FS3	Kinugasa		170/172	258/282	211/213	239	214/224	228	178/190	201/205	224/226	206/210	199	183/189	127	185/194	147	207	214/251	216/218	206	204/206
	FS4	Koshusaisho	F	172/209	270/282	206/213	231/239	212/224	216/226	178/196	199/201	224	194/210	217/225	183/189	127	185/207	133/139	199/207	214	188/216	206/249	194/204
	FS5	Koyokoume		172/209	270/282	210/213	226	214/224	195/228	178/184	197/199	198	202/206	205	193	127/131	195/207	147	207	214/237	196/204	200/210	204
	FS6	Maezawakoume		170/209	272/282	199/213	239	212/214	194/214	190/196	183/199	198/214	200/206	207/225	189/193	127	173/207	137/151	199/207	214	188/216	206	204
	FS7	Orihime		170/209	268/282	199	226/239	212/224	199/226	178/196	189/191	198/214	202/210	213/217	189/193	111/131	177/185	133/149	207	237	204/222	200	204/206
	FS8	Purplequeen	F	172/209	270/282	206/213	231/239	212/224	216/226	178/196	199/201	224	194/210	217/225	183/189	127	185/207	133/139	199/207	214	188/216	206/249	194/204
	FS9	Ryukyokoume		172/174	258/282	199/206	243/253	212	214/228	190/196	191/201	214/224	206	203/221	189	127	177/185	147/153	199/207	214	188/196	206/247	194/204
	FS10	Shinanokoume		172/174	266/282	199/206	226/239	212/214	199/214	186/196	183/201	214/226	202	209/217	189	127	177/185	147	203/207	214/237	188	206	194/204

Table S3. (Continued).

Species	Code	Name	Identity[z]	PMKS15	PMKS21	PMKS49	PMKS59	PMKS68	PMKS75	PMKS99	PMKS113	PMKS121	PMKS131	PMKS133	PMKS149	PMKS164	PMKS175	PMKS179	PMKS187	PMKS191	PMKS193	PMKS197	PMKS201
P. mume	O1	Akebono		162/180	252/286	199/213	228/239	224/233	199/226	188/196	183/189	198/226	214	203/227	189/203	111/125	185	149	199/228	214/237	196/220	206/212	194
	O2	Asahitaki		166	266/272	200/208	226/253	212/224	199/214	190/196	183/191	198/224	206	205/221	179/189	111/127	173/185	147/149	205/207	214	204/218	206/249	194/204
	O3	Benichidori		168/172	270/282	199/212	231/239	212/224	175/199	196/198	183/207	226/230	194	223	179/203	111	195/199	133/147	205	214/241	198/204	202/247	196/204
	O4	Benioshuku	G	172/180	256/270	199	223/239	224/239	192/216	188/196	199/207	214/230	194/208	196/223	168/189	111/131	199/207	137/151	199/207	214/237	188/196	206/212	194/198
	O5	Chasenbai		172	282	206/211	226	231	199	178/190	199	208/214	194/200	217/225	193	127	185	149	199	214/233	196/220	247	194/204
	O6	Chinamume		170/209	266	199/206	239/253	216/224	212/226	178/190	199/207	224	202/206	205	179/193	127/131	185/199/200	135/147	199	237/239	188/218	206	204
	O7	Eikan	H	172/178	264/282	212	231/239	220	175/198	196/198	199/207	230	214	203/209	193/201	111	185/199	135/147	199	214/241	192/196	247	204
	O8	Gekkyuden		170/172	256/266	199	223/251	212/224	216/226	196	207	214/224	194/210	221/223	189/205	111/129	199	135/137	199/207	214/237	188/196	206/208	198
	O9	Gofukushidare		166/180	282/286	199	228/239	224/231	199/214	188/196	183/205	214/226	214	203/209	168/203	111/125	185	147/149	205/228	214	196/216	212/247	194/204
	O10	Goshoko		172/209	270/272	199/213	226/239	212/237	199/216	178/196	191	198/226	200	223/225	189/193	111/131	185/195	147	203/228	214/263	216	200/206	204
	O11	Hasegawashibori	G	172/180	256/270	199	223/239	224/239	192/216	188/196	199/207	214/230	194/208	196/223	168/189	111/131	199/207	137/151	199/207	214/237	188/196	206/212	194/198
	O12	Hitoeryokugaku	A	166/168	258/270	213/215	231/257	226/231	198/214	182/188	199/205	220/226	214	203/209	189/205	111/131	171/185	147/159	213/228	214/235	196	212	194/204
	O13	Horyukaku		180/209	256/270/286	199/206/213	223/231/239	212	212/214/216	188/194/196	183/199	198/208/230	208	196/199/223	179/189	111/127/131	177/185	135/147/149	207/228	214/241	188/190/216	206/212	194/204
	O14	Ikuyonezame		172/178	264/282	212	231/239	220/237	175/198	196/198	201/207	214/230	214	203/209	193/201	111	185/199	135/147	199	214/241	194/196	208/247	198/204
	O15	Jakobai		170/209	266	199/206	239/251	216/224	212/228	178/190	199/209	224	202/206	205	179/193	127/131	185/199/200	135/147	199	237/239	188/218	206	204
	O16	Kagoshimako	H	172/178	264/282	212	231/239	220	175/198	196/198	199/207	230	214	203/209	193/201	111	185/199	135/147	199	214/241	192/196	247	204
	O17	Kanbaishidare		180/209	284/286	199/213	239/253	212	175/214	190/198	191/207	224/230	214	203	189	127	185	149/153	213/230	214/235	196/220	212/247	204
	O18	Kanseishidare	I	162/168	266/270	199/206	231/239	214/224	214/230	188/194	183/191	224/230	210/214	199/203	193/205	111/129	185/187/194/198	147/149	213/230	235/251	196/216	206/212	204
	O19	Kinko		172/178	264/282	212	231/239	220/237	175/198	196/198	199/207	214/230	214	203/209	193/201	111	185/199	135/147	199	214/241	192/196	208/247	198/204
	O20	Kurohikari		172/178	264/282	212	231/239	220	175/198	198	199/207	230	214	203	193/201	111	185/198	147	199	214/241	192/196	208	198
	O21	Kurokumo		178	264/282	212	231/239	220/237	175/198	196/198	199	214/230	214	203/209	193	111	199	135	199	214/241	192/196/218	208	198
	O22	Mangetsushidare		168/209	270	199/211	226/239	212/224	194/228	188/190	197/199	222/230	214	203/223	183/205	127/129	185/195	147/149	213/228	233/235	188/196	206/212	204
	O23	Meotoshidare		166/172	266/282	199/211	226/239	214/231	214/216	178/188	191/205	214/224	214	203	168/201	125/131	185	149	199/228	214	216/220	212/247	204
	O24	Mera		209	258	206	226/243	212/214	199/214	180/188	183/189	224/226	194/212	199/219	189	127/131	185/197	137/147	199	214/237	196/220	200/206	204
	O25	Michishirube		180/209	264/284	199	223/253	212/239	192/214	190/198	191/209	214/230	206/208	196/221	189/191	111/127	185/199	135/147	207/228	237/263	188/198	202/249	196/204
	O26	Morinoseki		172	282	213	239	212	175	190	199/201	214	194/206	221/223	189/201	111/127	185	147	199	237	218	247	204
	O27	Morinoura		180	264/286	199/212	223/228	220/224	198/214	188/196	205/207	214/230	214	203	179/203	111	185/199	133/147	228	214/241	192/216	208/247	198/204
	O28	Okinaume		170/180	258/270	206/208	226	212	226	192/196	183	214/220	-	203/213	189/205	127/129	185	147	230	214	188/190	206	194/204
	O29	Omoinomama	J	180/209	256/286	199/211	253/257	224/231	214/216	188/196	183/199	214/230	214	203/223	168/203	111/125	185/207	137/147	207/213	214	196	206/212	194/204
	O30	Osakazuki		178/209	264/272	199/213	223/239	212/224	192/214	190	183/191	214	206	203/221	189/193	111/127	185/199	147	199/207	237/263	188/198	202/206	194/196
	O31	Sabashiko	K	170/209	266	199/206	237/251	216/224	212/226	178/190	199/207	224	202/206	205	179/193	127/131	185/199/200	135/147	199	237/239	188/218	206	204
	O32	Seiryushidare	I	162/168	266/270	199/206	231/239	214/224	214/230	188/194	183/191	224/230	210/214	199/203	193/205	111/129	185/187/194/198	147/149	213/230	235/251	196/216	206/212	204
	O33	Shinheike		172/180	282	212	231/239	212/220	175	190/198	199/207	214	194	223	179/189	111/127	185/199	133/153	230	214	218	247	196/204
	O34	Shirobotan		180/209	270/286	199/212	226/251	216/239	199/212	196	199/207	214/224	206	203/205	179/189	131	185	147/151	230	214/241	188/192	206/208	198/204
	O35	Suishinbai		166/172	256/266	199/213	239/259	224	175/216	188	199/207	230	208/214	203/223	189/201	111	171/199	133/151	205/207	214	196	206/212	198/204
	O36	Suoume		172/180	266/286	200/225	223/253	214/216	212/214	178/196	183/191	224	206	205/223	179/193	111/131	185	135/147	207/222	241/263	192/216	198/208	198/208
	O37	Tagonotsuki		178/184	252/272	199/206	226/239	214/224	226	190/194	183/199	208/224	202	205/209	193/203	127/129	197/205	147/149	199/213	237/239	196/218	206	204
	O38	Takasago		172/194/209	256/259/282	199/223	227/237/253	224/257	190/214/226	186/190/194	183/199	198/214/226	194/200	197/209/223	176/189	111	173/177/185	147/153	199/207	214/237/251	184/188	183/247	194/204/208
	O39	Tamabotan	L	170/209	266	199/206	239/251	216/224	212/226	178/190	199/207	224	202/206	205	179/193	127/131	185/199	135/145	199	237/239	188/218	206	204
	O40	Tanfun	E	209	266	211	239/241	233/237	210/226	194	199	208/224	210/212	199	189	111/131	173/185/194	147/153	199/207	214/251	196/216	206	204
	O41	Toji		172/184	256/266	199/213	226/239	214	216/228	196	189/191	224	200	205/213	189/193	111	185/195	137/149	199	214	216	206/247	204
	O42	Toyadenotaka		172	252/272	199/213	239/243	212/224	216	178/196	189/199	198/224	200/202	205/227	189/193	111/127	173/185	147	199/207	214/239	220	200/249	194/198
	O43	Tsukushiko		180	266/286	199/213	239	212/216	175/199	196/198	183/207	198/230	206/214	203/223	189/203	111	185/199	133/135	222	241	198/216	202/206	196/204
	O44	Unryu	K	170/209	266	199/206	239/251	216/224	212/226	178/190	199/207	224	202/206	205	179/193	127/131	185/199/200	135/147	199	237/239	188/218	206	204
	O45	Unryubai	L	170/209	266	199/206	239/251	216/224	212/226	178/190	199/207	224	202/206	205	179/193	127/131	185/199	135/147	199	237/239	188/218	206	204
	O46	Utsushiroyama	J	180/209	256/286	199/211	253/257	224/231	214/216	188/196	183/199	214/230	214	203/223	168/203	111/125	185/207	137/147	207/213	214	196	206/212	194/204
	O47	Yaetoji		170/209	266	199/206	239/251	216/224	212/226	178/190	199/209	224/226	202/206	205	179/193	127/131	185/199/200	135/147	199	237/239	188/218	206	204
	O48	Yanagawashibori		180/209	256/286	199/211	251/257	224/231	214/216	188/196	183/199	214/230	214	203/223	168/203	111/125	185/207	137/147	207/213	214	196	206/212	194/204
	O49	Yokihi	J	180/209	256/286	199/211	253/257	224/231	214/216	188/196	183/199	214/230	214	203/223	168/203	111/125	185/207	137/147	207/213	214	196	206/212	194/204

Table S3. (Continued).

Species	Code	Name	Identity[z]	PMKS15	PMKS21	PMKS49	PMKS59	PMKS68	PMKS75	PMKS99	PMKS113	PMKS121	PMKS131	PMKS133	PMKS149	PMKS164	PMKS175	PMKS179	PMKS187	PMKS191	PMKS193	PMKS197	PMKS201
P. mume	T1	Ellching		184/209	260/266	199	237/239	212	213/230	166	195/209	214/226	192/196	196/225	188/189	127/135	195/199	151	207/220	243/253	188	208/210	198
	T2	Hakufunbai		200/207	264/268	199/215	237	212	198/200	190	191/195	224/226	188	196	177/201	125/131	187/191/195	131/165	222/228	249	188	208/214	198/204
	T3	ST		184	258/264	199/206	229/237	212/224	198/201	184/188	199/201	186	192	225	188/193	129	175/191	131/151	209/211	243	188/192	206/208	198/204
	T4	Taiwan		192/202	268/272	206	233/237	224	191/198	184	195	208	190/196	196/203	179/200	127/129	175/195	131	220/228	231/243	188/192	208	198
	T5	85486		180/184	264/272	206	226/237	212/224	202/230	184	195/199	208/214	196/204	196	189	127	175/191	131	209	243/253	188	208	198/200
P. mume	AM1	Bungo		194/209	259/270	199/223	209/227	214/257	190/199	186/198	183	214/226	194	221/223	189/193	111/127	173/177	147	199	214/251	184/220	183/206	194/208
	AM2	Fushida		172/194	252/259	206/223	226/227	214/257	190/199	186/196	191	198/224	194/200	197/221	176/193	111	177/207	147	199	239/251	184/196	183/206	204/208
	AM3	Inabungo		172/194	259/282	199/223	227/253	224/257	190/226	186/190	199	198/214	194/200	197/223	176/189	111	173/177	147/153	199	214/251	184/188	183/247	194/208
	AM4	Jumbotakada		188/192	262/282	199	227/247	257	190	184/190	199	214	194/220	197/223	176/189	111/119	177/192	137/147	-	251	184/190	183/198	208
	AM5	Kanshikobai	M	172/192	271/282	213/223	225/239	212/257	190/199	186/190	197	198	200/206	197/221	166/189	111	173/185	145/147	207	214/251	202/218	198/206	194/247
	AM6	Kurodaume	N	172/192	258/274	206/223	223/239	224/257	181/216	186/196	191	198/226	200	197/223	166/189	111/129	173/177	137/151	203	239/251	184/196	198/200	204/208
	AM7	Musashino	N	172/192	258/274	206/223	223/239	224/257	181/216	186/196	191	198/226	200	197/223	166/189	111/129	173/177	137/151	203	239/251	184/196	198/200	204/208
	AM8	Rinshibai	M	172/192	271/282	213/223	225/239	212/257	190/199	186/190	197	198	200/206	197/221	166/189	111	173/185	145/147	207	214/251	202/218	198/206	194/247
	AM9	Seiyobai		166/194	259/266	213/223	226/227	212/257	190/214	186/196	197	198/224	194/200	197/209	189/201	127	177/185	145/147	199	214/251	184/196	183/206	206/208
	AM10	Taihei		172/188	261/282	213	239/247	212/224	190/216	178/198	207	214/224	194/210	213/227	189/193	131	173/195	147	199/205	214/251	190/216	212/247	194/204
P. mume	SM1	PM1-1	O	166	260/282	192/211	241	212/240	210/222	194	197/225	198/224	212	199/223	184/189	111	173	147/152	200/207	209/251	212/216	194/206	206/215
				166	260/282	192/211	241	212/240	210/222	194	197/227	198/224	212	199/223	184/189	111	173	147/152	200/207	209/251	212/216	194/206	206/215
	SM2	PM1-4	O	166	260/282	192/211	241	212/240	210/222	194	197/225	198/224	212	199/223	184/189	111	173	147/152	200/207	209/251	212/216	194/206	206/215
	SM3	Sumomoume		166	266/268	192/211	239/245	228/237	176/199	196	199/225	224	210	199/223	166/189	121/129	195	137/165	203/207/216	209/237	196/198	206	204/215
	SM4	Tsuyuakane		166	260/282	178/213	239/256	212/228	170/194	190	199/225	224	-	223/225	166/203	111/121	173	137/157	199/200/216	210/214	204	198/212	204/230
				166	260/282	178/213	239/256	212/228	170/194	190	199/227	224	-	223/225	166/203	111/121	173	137/157	199/200/216	210/214	204	198/212	204/230
P. armeniaca	Pa_1[y]	Heiwa_1		188/196	266/284	223	247	255	190/228	180	185	198	200	197	166/176	111	173/177	151	173/177	251	184/202	183/198	208
	Pa_2[y]	Heiwa_2		188/196	266/284	223	227/247	257	190/228	186	185/187	198	198/200	197	166/176	111	173/177	151	173/177	251	184/202	183/198	208
P. salicina	Ps	Oishiwase		-	270	192/203	243/270/272	228/240	202/251	176/186/190	183/187	-	-	210/225	166/184	143	-	152/187	210/216/224	209/210	196/198	202/206	227/238
P. persica	Pa	Hakuho		177/216	282/284	215/219	225/256/264	225	179	166/198	185/195	200/206	-	195	185/199	111	190/196	157	-	237	186	177	258
P. dulcis	Pd	Almond Wakayama 1		177/216	282	215	217/225	225	179	173	199	200	-	195	193	127	190	157	190	249	186	177	257

[z] The same letters indicate the accessions sharing identical genotypes at all loci.

[y] Accessions having multiple genotypes at more than two loci.

Table S4. Values (r^2) of linkage disequilibrium (LD) between the most major alleles (below diagonal) and among all combinations of alleles. (above diagonal) using 20 microsatellite loci.

	LG[z]	1	1	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8	8	8	8
LG	Marker	PMKS15	PMKS21	PMKS49	PMKS59	PMKS68	PMKS75	PMKS99	PMKS113	PMKS121	PMKS131	PMKS133	PMKS149	PMKS164	PMKS175	PMKS179	PMKS187	PMKS191	PMKS193	PMKS197	PMKS201
1	PMKS15		0.027[x]	0.023	0.013	0.016	0.018	0.021	0.023	0.018	0.015	0.017	0.023	0.024	0.018	0.010	0.022	0.016	0.013	0.033	0.026
1	PMKS21	0.051[y]		0.014	0.029	0.019	0.025	0.018	0.017	0.025	0.016	0.025	0.013	0.016	0.020	0.013	0.023	0.021	0.011	0.032	0.022
2	PMKS49	0.000	0.000		0.027	0.033	0.022	0.035	0.029	0.023	0.014	0.020	0.024	0.027	0.015	0.021	0.015	0.016	0.023	0.037	0.018
2	PMKS59	0.009	0.033	0.006		0.026	0.032	0.017	0.016	0.017	0.013	0.022	0.022	0.022	0.025	0.012	0.019	0.016	0.017	0.018	0.021
3	PMKS68	0.004	0.036	0.013	0.010		0.027	0.024	0.024	0.023	0.014	0.020	0.044	0.033	0.019	0.015	0.017	0.012	0.024	0.022	0.013
3	PMKS75	0.002	0.072	0.003	0.074	0.036		0.024	0.024	0.027	0.019	0.024	0.020	0.025	0.017	0.015	0.018	0.018	0.015	0.030	0.014
4	PMKS99	0.033	0.023	0.014	0.000	0.000	0.008		0.024	0.029	0.024	0.023	0.025	0.026	0.023	0.019	0.015	0.027	0.025	0.032	0.024
4	PMKS113	0.089	0.021	0.073	0.003	0.027	0.058	0.041		0.012	0.010	0.013	0.012	0.016	0.022	0.013	0.021	0.012	0.014	0.016	0.018
5	PMKS121	0.063	0.096	0.005	0.044	0.004	0.022	0.003	0.011		0.032	0.028	0.018	0.019	0.015	0.008	0.023	0.014	0.027	0.043	0.016
5	PMKS131	0.000	0.001	0.005	0.001	0.002	0.002	0.002	0.019	0.080		0.052	0.025	0.019	0.015	0.016	0.019	0.022	0.021	0.030	0.011
5	PMKS133	0.048	0.038	0.000	0.041	0.007	0.064	0.006	0.019	0.061	0.007		0.032	0.018	0.020	0.013	0.026	0.019	0.022	0.033	0.021
6	PMKS149	0.064	0.000	0.009	0.040	0.058	0.001	0.007	0.005	0.043	0.019	0.081		0.012	0.018	0.020	0.025	0.014	0.017	0.021	0.009
6	PMKS164	0.014	0.000	0.011	0.027	0.015	0.001	0.012	0.002	0.011	0.008	0.010	0.006		0.018	0.009	0.012	0.018	0.029	0.016	0.006
7	PMKS175	0.005	0.025	0.000	0.038	0.006	0.031	0.052	0.022	0.003	0.001	0.075	0.006	0.030		0.021	0.019	0.030	0.013	0.029	0.028
7	PMKS179	0.005	0.012	0.046	0.007	0.000	0.010	0.018	0.006	0.002	0.071	0.003	0.003	0.023	0.014		0.015	0.012	0.014	0.026	0.010
8	PMKS187	0.027	0.093	0.001	0.034	0.019	0.044	0.001	0.047	0.068	0.005	0.018	0.005	0.003	0.002	0.004		0.011	0.016	0.024	0.011
8	PMKS191	0.006	0.012	0.012	0.001	0.005	0.009	0.084	0.005	0.001	0.003	0.029	0.002	0.005	0.078	0.002	0.000		0.033	0.021	0.015
8	PMKS193	0.000	0.007	0.004	0.007	0.031	0.007	0.001	0.027	0.095	0.023	0.054	0.001	0.021	0.004	0.014	0.000	0.067		0.032	0.018
8	PMKS197	0.106	0.093	0.002	0.027	0.028	0.006	0.006	0.020	0.110	0.018	0.128	0.038	0.029	0.076	0.060	0.003	0.009	0.005		0.050
8	PMKS201	0.070	0.044	0.007	0.024	0.011	0.003	0.046	0.043	0.020	0.004	0.026	0.000	0.001	0.004	0.001	0.024	0.003	0.001	0.087	

[z] Linkage group.

[y] Values (below diagonal) were calculated between the most frequent alleles at both loci.

[x] Values (above diagonal) were calculated among all combinations of alleles at both loci and summed up with weight of allele frequencies.

Table S5. The strongest candidates for selective sweep based on nSL test for selection.

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)	*P* value ($-\log_{10} p$)[1]	Location[2]	Nearest gene[3]	Gene description (Phytozome v12.1)[4]
1	27821640	Taiwan	nSL	-4.89	5.59	exonic	Prupe.1G271100	Myb-like DNA-binding domain (Myb_DNA-bind_6)
1	27928643	Taiwan	nSL	-4.23	4.29	intronic	Prupe.1G272900	F14L17.7 PROTEIN
1	43365326	Taiwan	nSL	-4.52	4.84	intronic	Prupe.1G530500	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 5 [EC:2.1.1.-] (NDUFAF5)
2	14084678	China	nSL	4.42	4.64	intronic	Prupe.2G089100	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	14084693	China	nSL	5.47	6.89	exonic	Prupe.2G089100	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	14085185	China	nSL	5.03	5.88	exonic	Prupe.2G089100	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	14085205	China	nSL	5.03	5.88	exonic	Prupe.2G089100	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	24347512	China	nSL	-4.09	4.03	exonic	Prupe.2G210800	Non-specific serine/threonine protein kinase / Threonine-specific protein kinase
2	25058720	Japan	nSL	-4.39	4.59	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058752	Japan	nSL	-4.58	4.96	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058764	China	nSL	5.69	7.44	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058764	Japan	nSL	-4.67	5.14	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058787	Japan	nSL	-5.11	6.07	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058915	Japan	nSL	4.74	5.28	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058915	China	nSL	4.74	4.51	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058958	Japan	nSL	-4.61	5.01	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063416	Japan	nSL	-5.32	6.54	intronic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063416	China	nSL	-4.27	4.36	intronic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063421	Japan	nSL	-5.45	6.84	intronic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063421	China	nSL	-4.11	4.07	intronic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063518	Japan	nSL	-5.29	6.47	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063529	Japan	nSL	-4.87	5.54	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063561	Japan	nSL	-4.20	4.23	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063583	Japan	nSL	-4.40	4.60	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063584	Japan	nSL	-4.62	5.03	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063588	Japan	nSL	-4.09	4.03	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063591	Japan	nSL	-4.76	5.31	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063607	Japan	nSL	-4.88	5.57	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063640	Japan	nSL	-5.57	7.14	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063646	Japan	nSL	-4.31	4.44	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063723	Japan	nSL	-4.14	4.13	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25066646	Japan	nSL	-4.09	4.03	5'-UTR	Prupe.2G223400	Non-specific serine/threonine protein kinase / Threonine-specific protein kinase
2	29182499	Taiwan	nSL	-4.90	5.61	exonic	Prupe.2G306600	Phosphate-transporting ATPase / ABC phosphate transporter
2	29364194	Taiwan	nSL	-4.86	5.52	exonic	Prupe.2G310200	PPR repeat (PPR) // PPR repeat family (PPR_2)
3	3589546	Japan	nSL	4.32	4.45	exonic	Prupe.3G050900	3-epi-6-deoxocathasterone 23-monooxygenase (CYP90D1)
4	15533455	Japan	nSL	-4.44	4.68	exonic	Prupe.4G237900	Premnaspirodiene oxygenase / Hyoscymus muticus premnaspirodiene oxygenase
4	23440961	Taiwan	nSL	-4.78	5.36	exonic	Prupe.4G278500	BETA-GALACTOSIDASE 10
5	17802487	Taiwan	nSL, SweeD	-4.61	5.01	exonic	Prupe.5G234600	F15H18.19
6	15201611	China	nSL	-4.41	4.62	intergenic	-	-
6	15201645	China	nSL	-4.52	4.83	intergenic	-	-
6	15203924	China	nSL	-4.59	4.97	intergenic	-	-
6	15204586	China	nSL	-4.51	4.81	intergenic	-	-
6	15205213	China	nSL	-4.15	4.15	intergenic	-	-
6	15208354	Taiwan	nSL, SweeD	-4.22	4.26	intergenic	-	-
6	15208368	China	nSL	-4.21	4.25	intergenic	-	-
6	15208368	Taiwan	nSL, SweeD	-4.22	4.26	intergenic	-	-

Table S5. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)	*P* value ($-\log_{10} p$)[1]	Location[2]	Nearest gene[3]	Gene description (Phytozome v12.1)[4]
6	15208373	Taiwan	nSL, SweeD	-4.21	4.24	intergenic	-	-
6	15208387	China	nSL	-4.21	4.25	intergenic	-	-
6	15208387	Taiwan	nSL, SweeD	-4.21	4.24	intergenic	-	-
6	15208402	Taiwan	nSL, SweeD	-4.21	4.24	intergenic	-	-
6	15210295	China	nSL	-5.33	6.57	intergenic	-	-
6	15210295	Taiwan	nSL, SweeD	-4.18	4.19	intergenic	-	-
6	15210309	China	nSL	-5.17	6.20	intergenic	-	-
6	15210309	Taiwan	nSL, SweeD	-4.18	4.19	intergenic	-	-
6	15210314	Taiwan	nSL, SweeD	-4.18	4.19	intergenic	-	-
6	15210328	China	nSL	-4.86	5.52	intergenic	-	-
6	15210328	Taiwan	nSL, SweeD	-4.11	4.07	intergenic	-	-
6	15210330	China	nSL	-4.10	4.06	intergenic	-	-
6	15210343	China	nSL	-4.65	5.09	intergenic	-	-
6	15210343	Taiwan	nSL, SweeD	-4.14	4.11	intergenic	-	-
6	15210956	China	nSL	-4.73	5.25	intergenic	-	-
6	15210956	Taiwan	nSL, SweeD	-4.73	4.11	intergenic	-	-
6	15210970	China	nSL	-5.39	6.70	intergenic	-	-
6	15210970	Taiwan	nSL, SweeD	-4.16	4.16	intergenic	-	-
6	15210975	Taiwan	nSL, SweeD	-4.16	4.16	intergenic	-	-
6	15210989	China	nSL	-5.39	6.70	intergenic	-	-
6	15210989	Taiwan	nSL, SweeD	-4.16	4.16	intergenic	-	-
6	15211004	China	nSL	-4.95	5.72	intergenic	-	-
6	15211004	Taiwan	nSL, SweeD	-4.20	4.23	intergenic	-	-
6	15212276	China	nSL	-4.42	4.64	intergenic	-	-
6	15212290	China	nSL	-4.27	4.36	intergenic	-	-
6	15212309	China	nSL	-4.27	4.36	intergenic	-	-
6	15212324	China	nSL	-4.36	4.54	intergenic	-	-
6	15216857	China	nSL	-4.21	4.25	intergenic	-	-
6	15216890	China	nSL	-4.26	4.34	intergenic	-	-
6	15220832	China	nSL	-4.13	4.10	intergenic	-	-
6	15221597	China	nSL	-4.15	4.13	intergenic	-	-
7	2150169	Taiwan	nSL	-5.29	6.48	upstream	Prupe.7G014600	Polyribonucleotide nucleotidyltransferase / Polynucleotide phosphorylase
7	14227342	Taiwan	nSL	-4.68	5.16	upstream	Prupe.7G117500	-
7	17430048	Taiwan	nSL	-4.12	4.08	exonic	Prupe.7G174600	Non-specific serine/threonine protein kinase / Threonine-specific protein kinase
7	18710723	Taiwan	nSL	-5.72	7.49	exonic	Prupe.7G200000	MOB kinase activator 1 (MOB1, Mats)
8	961016	Taiwan	nSL, SweeD	-5.70	7.45	exonic	Prupe.8G012000	PH DOMAIN LEUCINE-RICH REPEAT-CONTAINING PROTEIN PHOSPHATASE 2
8	1063829	Taiwan	nSL, SweeD	-4.80	5.39	intronic	Prupe.8G012800	F-box and leucine-rich repeat protein 1 (S-phase kinase-associated protein 2) (SKP2, FBXL1)
8	6635251	Taiwan	nSL	-4.24	4.31	intergenic	-	-
8	6635296	Taiwan	nSL	-5.42	6.77	intergenic	-	-
8	6736751	Japan	nSL	-5.21	6.30	intergenic	-	-
8	6741779	China	nSL	-4.88	5.56	intergenic	-	-
8	6741779	Japan	nSL	-4.11	4.06	intergenic	-	-
8	6742605	Japan	nSL	-4.53	4.86	intergenic	-	-
8	6742652	Japan	nSL	-4.57	4.94	intergenic	-	-
8	6743451	China	nSL	-5.40	6.73	intergenic	-	-

Table S5. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)	*P* value ($-\log_{10}p$)[1]	Location[2]	Nearest gene[3]	Gene description (Phytozome v12.1)[4]
8	6743451	Japan	nSL	-6.61	9.88	intergenic	-	-
8	6743498	Japan	nSL	-4.46	4.71	intergenic	-	-
8	6747688	Japan	nSL	-4.52	4.83	intergenic	-	-
8	6747712	Taiwan	nSL	-4.78	5.36	intergenic	-	-
8	6747712	Japan	nSL	-4.55	4.89	intergenic	-	-
8	6747735	Japan	nSL	-4.51	4.82	intergenic	-	-
8	6749382	China	nSL	-4.65	5.10	intergenic	-	-
8	6749382	Japan	nSL	-4.55	4.90	intergenic	-	-
8	6749406	Taiwan	nSL	-4.89	5.58	intergenic	-	-
8	6750746	China	nSL	-5.68	7.41	intergenic	-	-
8	6750746	Japan	nSL	-5.17	6.20	intergenic	-	-
8	6753331	China	nSL	-4.16	4.15	intergenic	-	-
8	6753331	Japan	nSL	-4.47	4.74	intergenic	-	-
8	6756087	China	nSL	-4.56	4.92	intergenic	-	-
8	6756087	Japan	nSL	-4.52	4.83	intergenic	-	-
8	6756134	China	nSL	-4.30	4.42	intergenic	-	-
8	6756980	Japan	nSL	-5.08	6.00	intergenic	-	-
8	6759474	Taiwan	nSL	-4.71	5.22	intergenic	-	-
8	6759474	China	nSL	-4.58	4.94	intergenic	-	-
8	6759474	Japan	nSL	-4.40	4.61	intergenic	-	-
8	6759521	Japan	nSL	-4.69	5.17	intergenic	-	-
8	6759521	Taiwan	nSL	-4.82	5.45	intergenic	-	-
8	6759521	China	nSL	-4.33	4.47	intergenic	-	-
8	6760320	China	nSL	-4.70	5.20	intergenic	-	-
8	6760320	Japan	nSL	-4.10	4.05	intergenic	-	-
8	6762857	Japan	nSL	-5.30	6.49	intergenic	-	-
8	6762857	China	nSL	-4.40	4.61	intergenic	-	-
8	6762881	Japan	nSL	-4.69	5.17	intergenic	-	-
8	6762904	Japan	nSL	-4.46	4.72	intergenic	-	-
8	6764413	China	nSL	-4.91	5.63	intergenic	-	-
8	6764550	China	nSL	-4.76	5.33	intergenic	-	-
8	6764550	Japan	nSL	-4.52	4.83	intergenic	-	-
8	6767936	Japan	nSL	-5.07	5.98	intergenic	-	-
8	6767983	China	nSL	-4.39	4.59	intergenic	-	-
8	6767983	Japan	nSL	-4.20	4.23	intergenic	-	-
8	6770475	Japan	nSL	-4.30	4.42	intergenic	-	-
8	6770489	China	nSL	-4.28	4.37	intergenic	-	-
8	6770522	Japan	nSL	-4.37	4.55	intergenic	-	-
8	6771321	Japan	nSL	-4.08	4.01	intergenic	-	-
8	6771335	China	nSL	-4.75	5.30	intergenic	-	-
8	6771345	Japan	nSL	-4.47	4.75	intergenic	-	-
8	6771368	Taiwan	nSL	-5.53	7.03	intergenic	-	-
8	6771368	Japan	nSL	-4.49	4.77	intergenic	-	-
8	6772168	Japan	nSL	-5.20	6.27	intergenic	-	-
8	6772168	Taiwan	nSL	-5.58	7.17	intergenic	-	-
8	6772182	China	nSL	-5.30	6.50	intergenic	-	-
8	6772215	Japan	nSL	-4.50	4.80	intergenic	-	-
8	6773861	Taiwan	nSL	-4.30	4.41	intergenic	-	-
8	6773875	China	nSL	-4.60	4.99	intergenic	-	-
8	6773885	Taiwan	nSL	-5.73	7.54	intergenic	-	-

Table S5. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)	*P* value ($-\log_{10} p$)[1]	Location[2]	Nearest gene[3]	Gene description (Phytozome v12.1)[4]
8	6773908	Taiwan	nSL	-4.68	5.16	intergenic	-	-
8	6776414	China	nSL	-4.13	4.10	intergenic	-	-
8	6776424	Taiwan	nSL	-5.35	6.62	intergenic	-	-
8	6776447	China	nSL	-4.88	5.57	intergenic	-	-
8	6776447	Japan	nSL	-4.81	5.42	intergenic	-	-
8	6777246	China	nSL	-4.81	6.43	intergenic	-	-
8	6777246	Japan	nSL	-5.25	6.38	intergenic	-	-
8	6777246	Taiwan	nSL	-4.33	4.47	intergenic	-	-
8	6780631	Taiwan	nSL	-5.50	6.98	intergenic	-	-
8	6780631	Japan	nSL	-4.16	4.15	intergenic	-	-
8	6780645	China	nSL	-4.25	4.32	intergenic	-	-
8	6780678	Taiwan	nSL	-5.50	6.96	intergenic	-	-
8	6782323	China	nSL	-4.60	5.00	intergenic	-	-
8	6782323	Japan	nSL	-4.10	4.05	intergenic	-	-
8	6782337	China	nSL	-4.70	5.20	intergenic	-	-
8	6782347	China	nSL	-4.11	4.07	intergenic	-	-
8	6782370	Japan	nSL	-4.16	4.15	intergenic	-	-
8	6783169	China	nSL	-4.17	4.18	intergenic	-	-
8	6783183	China	nSL	-4.91	5.64	intergenic	-	-
8	6783193	Taiwan	nSL	-4.84	5.49	intergenic	-	-
8	6783216	Japan	nSL	-4.74	5.28	intergenic	-	-
8	6783216	China	nSL	-4.26	4.35	intergenic	-	-
8	6784862	Japan	nSL	-4.22	4.26	intergenic	-	-
8	6784876	China	nSL	-4.76	5.31	intergenic	-	-
8	6784909	Japan	nSL	-4.80	5.41	intergenic	-	-
8	6786554	China	nSL	-4.80	5.41	intergenic	-	-
8	6786554	Japan	nSL	-4.80	5.46	intergenic	-	-
8	6786568	China	nSL	-4.73	5.27	intergenic	-	-
8	6786578	Taiwan	nSL	-5.37	6.65	intergenic	-	-
8	6786578	Japan	nSL	-4.10	4.05	intergenic	-	-
8	6786601	Japan	nSL	-4.96	5.74	intergenic	-	-
8	6787401	Japan	nSL	-4.28	4.38	intergenic	-	-
8	6787448	Japan	nSL	-4.82	5.44	intergenic	-	-
8	6788095	China	nSL	-4.20	4.23	intergenic	-	-
8	6788095	Japan	nSL	-4.33	4.48	intergenic	-	-
8	6789105	China	nSL	-4.31	4.44	intergenic	-	-
8	6789152	China	nSL	-4.08	4.02	intergenic	-	-
8	6789152	Japan	nSL	-4.33	4.48	intergenic	-	-
8	6789951	Japan	nSL	-4.96	5.75	intergenic	-	-
8	6791481	China	nSL	-4.56	4.91	intergenic	-	-
8	6791644	Japan	nSL	-4.57	4.93	intergenic	-	-
8	6791668	China	nSL	-4.32	4.45	intergenic	-	-

Table S5. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)	*P* value ($-\log_{10} p$)[1]	Location[2]	Nearest gene[3]	Gene description (Phytozome v12.1)[4]
8	6791691	Japan	nSL	-4.20	4.24	intergenic	-	-
8	6792327	Japan	nSL	-5.24	6.36	intergenic	-	-
8	6795054	China	nSL	-4.61	5.02	intergenic	-	-
8	6795077	China	nSL	-4.41	4.61	intergenic	-	-
8	6795876	Japan	nSL	-4.66	5.12	intergenic	-	-
8	6795876	China	nSL	-4.08	4.01	intergenic	-	-
8	6795900	China	nSL	-4.43	4.67	intergenic	-	-
8	6795900	Japan	nSL	-4.11	4.07	intergenic	-	-
8	6795923	China	nSL	-4.51	4.82	intergenic	-	-
8	6796722	Japan	nSL	-4.82	5.45	intergenic	-	-
8	6796722	China	nSL	-4.20	4.24	intergenic	-	-
8	6796746	China	nSL	-4.14	4.12	intergenic	-	-
8	6796769	Japan	nSL	-4.80	5.40	intergenic	-	-
8	6796769	China	nSL	-4.80	4.34	intergenic	-	-
8	6797406	Japan	nSL	-5.92	8.01	intergenic	-	-
8	6797569	Japan	nSL	-4.59	4.97	intergenic	-	-
8	6797593	China	nSL	-4.18	4.19	intergenic	-	-
8	6797593	Taiwan	nSL	-4.32	4.44	intergenic	-	-
8	6797616	China	nSL	-4.89	5.60	intergenic	-	-
8	6797627	Taiwan	nSL	5.62	7.26	intergenic	-	-
8	6879826	Japan	nSL	-4.09	4.02	intergenic	-	-
8	6879826	Taiwan	nSL	-4.39	4.59	intergenic	-	-
8	12838917	Taiwan	nSL	-4.31	4.43	downstream	Prupe.8G096200	-
8	14751706	China	nSL	-4.11	4.07	exonic	Prupe.8G121500	CARBOXYLESTERASE 12-RELATED
8	17899604	China	nSL	4.20	4.24	intronic	Prupe.8G177300	Cytochrome P450 CYP4/CYP19/CYP26 subfamilies

[1] *P* values based on long haplotype tests: nSL or XP-EHH. The strongest ($-\log_{10} p > 4$) candidates are listed.

[2] Upstream and downstream indicate the regions within 5 kb from the coding sequence.

[3] Nearest gene detected for SNPs (excluding intergenic locations) based on *P. persica* genome.

[4] Gene description is from Phytozome v12.1 (https://phytozome.jgi.doe.gov/pz/portal.html).

Table S6. The strongest candidates for selective sweep based on XP-EHH test for selection.

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)[1]	*P* value ($-\log_{10} p$)[2]	Location[3]	Nearest gene[4]	Gene description (Phytozome v12.1)[5]
1	2273856	Japan (ref) vs Taiwan	XP-EHH	-4.65	5.09	exonic	Prupe.1G032400	Potato inhibitor I family (potato_inhibit)
1	2273857	Japan (ref) vs Taiwan	XP-EHH	-4.65	5.10	exonic	Prupe.1G032400	Potato inhibitor I family (potato_inhibit)
1	2273872	Japan (ref) vs Taiwan	XP-EHH	-4.65	5.10	exonic	Prupe.1G032400	Potato inhibitor I family (potato_inhibit)
1	2275951	China (ref) vs Japan	XP-EHH	4.18	4.19	exonic	Prupe.1G032500	Potato inhibitor I family (potato_inhibit)
1	2275984	China (ref) vs Japan	XP-EHH	4.16	4.16	exonic	Prupe.1G032500	Potato inhibitor I family (potato_inhibit)
1	2275986	China (ref) vs Japan	XP-EHH	4.16	4.16	exonic	Prupe.1G032500	Potato inhibitor I family (potato_inhibit)
1	2276022	China (ref) vs Japan	XP-EHH	4.16	4.16	exonic	Prupe.1G032500	Potato inhibitor I family (potato_inhibit)
1	2276031	China (ref) vs Japan	XP-EHH	4.13	4.10	exonic	Prupe.1G032500	Potato inhibitor I family (potato_inhibit)
1	2276032	China (ref) vs Japan	XP-EHH	4.15	4.14	exonic	Prupe.1G032500	Potato inhibitor I family (potato_inhibit)
1	2480877	Japan (ref) vs Taiwan	XP-EHH	-7.60	12.95	exonic	Prupe.1G035500	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG 10-RELATED
1	2480878	Japan (ref) vs Taiwan	XP-EHH	-6.21	8.77	exonic	Prupe.1G035500	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG 10-RELATED
1	2480890	Japan (ref) vs Taiwan	XP-EHH	-5.12	6.08	exonic	Prupe.1G035500	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG 10-RELATED
1	2480898	Japan (ref) vs Taiwan	XP-EHH	-5.11	6.07	exonic	Prupe.1G035500	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG 10-RELATED
1	2480902	Japan (ref) vs Taiwan	XP-EHH	-4.40	4.60	exonic	Prupe.1G035500	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG 10-RELATED
1	2480908	Japan (ref) vs Taiwan	XP-EHH	-4.40	4.60	exonic	Prupe.1G035500	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG 10-RELATED
1	2480909	Japan (ref) vs Taiwan	XP-EHH	-4.40	4.61	exonic	Prupe.1G035500	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG 10-RELATED
1	2480943	Japan (ref) vs Taiwan	XP-EHH	-4.26	4.34	exonic	Prupe.1G035500	1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE HOMOLOG 10-RELATED
1	8601166	Japan (ref) vs Taiwan	XP-EHH	-4.09	4.03	intronic	Prupe.1G107200	Speckle-type POZ protein SPOP and related proteins with TRAF, MATH and BTB/POZ domains
1	8601182	Japan (ref) vs Taiwan	XP-EHH	-4.10	4.05	intronic	Prupe.1G107200	Speckle-type POZ protein SPOP and related proteins with TRAF, MATH and BTB/POZ domains
1	12958144	Japan (ref) vs Taiwan	XP-EHH	-6.67	10.05	exonic	Prupe.1G161800	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
1	12958147	Japan (ref) vs Taiwan	XP-EHH	-6.41	9.33	exonic	Prupe.1G161800	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
1	12958151	Japan (ref) vs Taiwan	XP-EHH	-5.70	7.46	exonic	Prupe.1G161800	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
1	13539405	Japan (ref) vs Taiwan	XP-EHH	-6.75	10.30	exonic	Prupe.1G166700	PPR repeat (PPR) // PPR repeat family (PPR_2)
1	22886475	Japan (ref) vs Taiwan	XP-EHH	-4.54	4.87	exonic	Prupe.1G215900	DNA2/NAM7 HELICASE FAMILY // SUBFAMILY NOT NAMED
1	27567308	Japan (ref) vs Taiwan	XP-EHH	-5.64	7.30	intronic	Prupe.1G268300	4-alpha-D-((1->4)-alpha-D-glucano)trehalose trehalohydrolase / Maltooligosyl trehalose trehalohydrolase
1	28656511	Japan (ref) vs Taiwan	XP-EHH	-4.66	5.12	exonic	Prupe.1G285700	PPR repeat (PPR) // PPR repeat family (PPR_2)
1	28656518	Japan (ref) vs Taiwan	XP-EHH	-4.66	5.12	exonic	Prupe.1G285700	PPR repeat (PPR) // PPR repeat family (PPR_2)
1	28656530	Japan (ref) vs Taiwan	XP-EHH	-4.66	5.12	exonic	Prupe.1G285700	PPR repeat (PPR) // PPR repeat family (PPR_2)
1	28656548	Japan (ref) vs Taiwan	XP-EHH	-4.66	5.12	exonic	Prupe.1G285700	PPR repeat (PPR) // PPR repeat family (PPR_2)
1	28656551	Japan (ref) vs Taiwan	XP-EHH	-4.67	5.14	exonic	Prupe.1G285700	PPR repeat (PPR) // PPR repeat family (PPR_2)
1	28798325	Japan (ref) vs Taiwan	XP-EHH	-4.62	5.03	exonic	Prupe.1G288400	CELL DIVISION CONTROL PROTEIN 48 HOMOLOG B
1	28798347	Japan (ref) vs Taiwan	XP-EHH	-4.65	5.10	exonic	Prupe.1G288400	CELL DIVISION CONTROL PROTEIN 48 HOMOLOG B
1	28798370	Japan (ref) vs Taiwan	XP-EHH	-4.88	5.58	exonic	Prupe.1G288400	CELL DIVISION CONTROL PROTEIN 48 HOMOLOG B
1	28798371	Japan (ref) vs Taiwan	XP-EHH	-4.90	5.61	exonic	Prupe.1G288400	CELL DIVISION CONTROL PROTEIN 48 HOMOLOG B
1	28798381	Japan (ref) vs Taiwan	XP-EHH	-5.39	6.71	exonic	Prupe.1G288400	CELL DIVISION CONTROL PROTEIN 48 HOMOLOG B
1	33939943	Japan (ref) vs Taiwan	XP-EHH	4.10	4.05	intronic	Prupe.1G373100	THIOREDOXIN H1
1	33939985	Japan (ref) vs Taiwan	XP-EHH	4.70	5.20	intronic	Prupe.1G373100	THIOREDOXIN H1
1	33940005	Japan (ref) vs Taiwan	XP-EHH	4.91	5.64	intronic	Prupe.1G373100	THIOREDOXIN H1
1	33940006	Japan (ref) vs Taiwan	XP-EHH	4.36	4.52	intronic	Prupe.1G373100	THIOREDOXIN H1
1	33940017	Japan (ref) vs Taiwan	XP-EHH	4.16	4.15	intronic	Prupe.1G373100	THIOREDOXIN H1
1	33940024	Japan (ref) vs Taiwan	XP-EHH	4.10	4.05	intronic	Prupe.1G373100	THIOREDOXIN H1
1	46787155	China (ref) vs Japan	XP-EHH	2.15	4.40	upstream	Prupe.1G574400	B3 DNA binding domain (B3)
1	46787156	China (ref) vs Japan	XP-EHH	4.62	5.03	upstream	Prupe.1G574400	B3 DNA binding domain (B3)

Table S6. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)[1]	*P* value ($-\log_{10} p$)[2]	Location[3]	Nearest gene[4]	Gene description (Phytozome v12.1)[5]
1	46787159	China (ref) vs Japan	XP-EHH	2.23	4.68	upstream	Prupe.1G574400	B3 DNA binding domain (B3)
1	46787185	China (ref) vs Japan	XP-EHH	2.30	4.97	exonic	Prupe.1G574400	B3 DNA binding domain (B3)
1	46787187	China (ref) vs Japan	XP-EHH	4.93	5.67	upstream	Prupe.1G574400	B3 DNA binding domain (B3)
1	46787199	China (ref) vs Japan	XP-EHH	4.82	5.44	upstream	Prupe.1G574400	B3 DNA binding domain (B3)
1	46787252	China (ref) vs Japan	XP-EHH	2.23	4.70	exonic	Prupe.1G574400	B3 DNA binding domain (B3)
2	1711983	Ornamental (ref) vs Fruit	XP-EHH	4.48	4.76	intronic	Prupe.2G018300	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
2	1712006	Ornamental (ref) vs Fruit	XP-EHH	4.12	4.09	intronic	Prupe.2G018300	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
2	4011035	Japan (ref) vs Taiwan	XP-EHH	-4.09	4.03	exonic	Prupe.2G037400	MITOCHONDRIAL TRANSCRIPTION TERMINATION FACTOR FAMILY PROTEIN-RELATED
2	4011069	Japan (ref) vs Taiwan	XP-EHH	-4.21	4.25	exonic	Prupe.2G037400	MITOCHONDRIAL TRANSCRIPTION TERMINATION FACTOR FAMILY PROTEIN-RELATED
2	4011127	Japan (ref) vs Taiwan	XP-EHH	-5.66	7.36	exonic	Prupe.2G037400	MITOCHONDRIAL TRANSCRIPTION TERMINATION FACTOR FAMILY PROTEIN-RELATED
2	4011141	Japan (ref) vs Taiwan	XP-EHH	-7.29	11.95	exonic	Prupe.2G037400	MITOCHONDRIAL TRANSCRIPTION TERMINATION FACTOR FAMILY PROTEIN-RELATED
2	4011143	Japan (ref) vs Taiwan	XP-EHH	-7.20	11.66	exonic	Prupe.2G037400	MITOCHONDRIAL TRANSCRIPTION TERMINATION FACTOR FAMILY PROTEIN-RELATED
2	4011158	Japan (ref) vs Taiwan	XP-EHH	-8.71	16.87	exonic	Prupe.2G037400	MITOCHONDRIAL TRANSCRIPTION TERMINATION FACTOR FAMILY PROTEIN-RELATED
2	5310774	Japan (ref) vs Taiwan	XP-EHH	-4.72	5.25	exonic	Prupe.2G046300	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
2	5310782	Japan (ref) vs Taiwan	XP-EHH	-4.39	4.59	exonic	Prupe.2G046300	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
2	5310822	Japan (ref) vs Taiwan	XP-EHH	-4.43	4.67	exonic	Prupe.2G046300	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
2	5310829	Japan (ref) vs Taiwan	XP-EHH	-4.73	5.25	exonic	Prupe.2G046300	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
2	13765751	Japan (ref) vs Taiwan	XP-EHH	-4.09	4.04	exonic	Prupe.2G087200	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	13884649	Ornamental (ref) vs Fruit	XP-EHH	-4.57	4.94	exonic	Prupe.2G087900	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	13884658	Ornamental (ref) vs Fruit	XP-EHH	-4.46	4.72	exonic	Prupe.2G087900	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	13884694	Ornamental (ref) vs Fruit	XP-EHH	-4.19	4.20	exonic	Prupe.2G087900	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	14006632	Japan (ref) vs Taiwan	XP-EHH	-7.23	11.74	exonic	Prupe.2G088800	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	14006671	Japan (ref) vs Taiwan	XP-EHH	-6.31	9.04	exonic	Prupe.2G088800	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	14006688	Japan (ref) vs Taiwan	XP-EHH	-4.96	5.75	exonic	Prupe.2G088800	Glycerophosphodiester phosphodiesterase / Glycerophosphoryl diester phosphodiesterase
2	20206925	Fruit (ref) vs Small-fruit	XP-EHH	-4.38	4.57	exonic	Prupe.2G145200	ZINC FINGER A20 AND AN1 DOMAIN-CONTAINING STRESS-ASSOCIATED PROTEIN 10-RELATED
2	20206930	Fruit (ref) vs Small-fruit	XP-EHH	-4.30	4.42	exonic	Prupe.2G145200	ZINC FINGER A20 AND AN1 DOMAIN-CONTAINING STRESS-ASSOCIATED PROTEIN 10-RELATED
2	25058795	Japan (ref) vs Taiwan	XP-EHH	-4.95	5.72	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058828	Japan (ref) vs Taiwan	XP-EHH	-5.09	6.02	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058836	Japan (ref) vs Taiwan	XP-EHH	-4.56	4.91	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058841	Japan (ref) vs Taiwan	XP-EHH	-4.56	4.91	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058849	Japan (ref) vs Taiwan	XP-EHH	-4.56	4.91	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058865	Japan (ref) vs Taiwan	XP-EHH	-6.48	9.52	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058870	Japan (ref) vs Taiwan	XP-EHH	-7.01	11.08	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058874	Japan (ref) vs Taiwan	XP-EHH	-6.97	10.94	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058882	Japan (ref) vs Taiwan	XP-EHH	-6.92	10.80	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058900	Japan (ref) vs Taiwan	XP-EHH	-5.82	7.76	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058904	Japan (ref) vs Taiwan	XP-EHH	-5.82	7.76	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058909	Japan (ref) vs Taiwan	XP-EHH	-5.45	6.84	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058915	Japan (ref) vs Taiwan	XP-EHH	-5.01	5.84	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058927	Japan (ref) vs Taiwan	XP-EHH	-5.61	7.24	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25058952	Japan (ref) vs Taiwan	XP-EHH	-4.17	4.17	exonic	Prupe.2G223200	Interleukin-1 receptor-associated kinase 1 (IRAK1)
2	25063519	Japan (ref) vs Taiwan	XP-EHH	-4.47	4.74	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED

Table S6. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)[1]	*P* value ($-\log_{10} p$)[2]	Location[3]	Nearest gene[4]	Gene description (Phytozome v12.1)[5]
2	25063522	Japan (ref) vs Taiwan	XP-EHH	-5.12	6.10	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063524	Japan (ref) vs Taiwan	XP-EHH	-5.46	6.87	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063525	Japan (ref) vs Taiwan	XP-EHH	-5.46	6.87	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063528	Japan (ref) vs Taiwan	XP-EHH	-4.44	4.67	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063529	Japan (ref) vs Taiwan	XP-EHH	-4.43	4.66	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063540	Japan (ref) vs Taiwan	XP-EHH	-4.43	4.67	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25063561	Japan (ref) vs Taiwan	XP-EHH	-4.61	5.02	exonic	Prupe.2G223300	VOLTAGE AND LIGAND GATED POTASSIUM CHANNEL // SUBFAMILY NOT NAMED
2	25688670	China (ref) vs Japan	XP-EHH	-4.31	4.43	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688671	China (ref) vs Japan	XP-EHH	-4.37	4.55	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688676	China (ref) vs Japan	XP-EHH	-4.35	4.51	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688686	China (ref) vs Japan	XP-EHH	-4.29	4.40	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688707	China (ref) vs Japan	XP-EHH	-4.22	4.27	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688710	China (ref) vs Japan	XP-EHH	-4.26	4.34	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688711	China (ref) vs Japan	XP-EHH	-4.11	4.07	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688718	China (ref) vs Japan	XP-EHH	-4.11	4.07	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688723	China (ref) vs Japan	XP-EHH	-4.12	4.08	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	25688724	China (ref) vs Japan	XP-EHH	-4.10	4.06	exonic	Prupe.2G234300	MEDIATOR OF RNA POLYMERASE II TRANSCRIPTION SUBUNIT 37C-RELATED
2	26487535	Ornamental (ref) vs Fruit	XP-EHH	-4.10	4.05	upstream	Prupe.2G249600	-
3	2447435	Ornamental (ref) vs Fruit	XP-EHH	4.62	5.04	intronic	Prupe.3G033100	GDSL ESTERASE/LIPASE 5-RELATED
3	2447493	Ornamental (ref) vs Fruit	XP-EHH	5.13	6.12	intronic	Prupe.3G033100	GDSL ESTERASE/LIPASE 5-RELATED
3	2447536	Ornamental (ref) vs Fruit	XP-EHH	5.11	6.06	intronic	Prupe.3G033100	GDSL ESTERASE/LIPASE 5-RELATED
3	2447557	Ornamental (ref) vs Fruit	XP-EHH	5.11	6.06	intronic	Prupe.3G033100	GDSL ESTERASE/LIPASE 5-RELATED
3	2447600	Ornamental (ref) vs Fruit	XP-EHH	5.46	6.86	intronic	Prupe.3G033100	GDSL ESTERASE/LIPASE 5-RELATED
3	2447626	Ornamental (ref) vs Fruit	XP-EHH	5.65	7.33	intronic	Prupe.3G033100	GDSL ESTERASE/LIPASE 5-RELATED
3	2447684	Ornamental (ref) vs Fruit	XP-EHH	5.65	7.33	intronic	Prupe.3G033100	GDSL ESTERASE/LIPASE 5-RELATED
3	2479866	China (ref) vs Japan	XP-EHH	4.44	4.68	exonic	Prupe.3G033600	GDSL ESTERASE/LIPASE 5-RELATED
3	2479867	China (ref) vs Japan	XP-EHH	4.45	4.70	exonic	Prupe.3G033600	GDSL ESTERASE/LIPASE 5-RELATED
3	2479917	China (ref) vs Japan	XP-EHH	4.22	4.27	exonic	Prupe.3G033600	GDSL ESTERASE/LIPASE 5-RELATED
3	6368060	Japan (ref) vs Taiwan	XP-EHH	-4.17	4.18	intergenic	-	-
3	6368107	Japan (ref) vs Taiwan	XP-EHH	-5.21	6.30	intergenic	-	-
3	6368116	Japan (ref) vs Taiwan	XP-EHH	-5.72	7.50	intergenic	-	-
3	6368117	Japan (ref) vs Taiwan	XP-EHH	-6.32	9.07	intergenic	-	-
3	9035425	China (ref) vs Japan	XP-EHH	-4.42	4.63	exonic	Prupe.3G110300	Salt stress response/antifungal (Stress-antifung)
3	9035429	China (ref) vs Japan	XP-EHH	-4.56	4.91	exonic	Prupe.3G110300	Salt stress response/antifungal (Stress-antifung)
3	9035437	China (ref) vs Japan	XP-EHH	-4.32	4.45	exonic	Prupe.3G110300	Salt stress response/antifungal (Stress-antifung)
3	9035458	China (ref) vs Japan	XP-EHH	-4.24	4.30	exonic	Prupe.3G110300	Salt stress response/antifungal (Stress-antifung)
3	10861769	Japan (ref) vs Taiwan	XP-EHH	-5.85	7.84	exonic	Prupe.3G122100	Fruit bromelain
3	10861774	Japan (ref) vs Taiwan	XP-EHH	-5.91	7.97	exonic	Prupe.3G122100	Fruit bromelain
3	10861776	Japan (ref) vs Taiwan	XP-EHH	-5.94	8.07	exonic	Prupe.3G122100	Fruit bromelain
3	10861792	Japan (ref) vs Taiwan	XP-EHH	-6.36	9.18	exonic	Prupe.3G122100	Fruit bromelain
3	10861803	Japan (ref) vs Taiwan	XP-EHH	-6.41	9.32	exonic	Prupe.3G122100	Fruit bromelain
3	16941327	Fruit (ref) vs Small-fruit	XP-EHH	-5.48	6.91	exonic	Prupe.3G154000	ANKYRIN REPEAT FAMILY PROTEIN-RELATED
3	16941332	Fruit (ref) vs Small-fruit	XP-EHH	-5.54	7.05	exonic	Prupe.3G154000	ANKYRIN REPEAT FAMILY PROTEIN-RELATED
3	17332394	Japan (ref) vs Taiwan	XP-EHH	-4.79	5.39	intergenic	-	-

Table S6. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)[1]	*P* value ($-\log_{10} p$)[2]	Location[3]	Nearest gene[4]	Gene description (Phytozome v12.1)[5]
4	306307	China (ref) vs Japan	XP-EHH	4.38	4.57	upstream	Prupe.4G005100	COPPER TRANSPORT PROTEIN ATOX1-RELATED // SUBFAMILY NOT NAMED
4	1959168	China (ref) vs Japan	XP-EHH	4.22	4.26	exonic	Prupe.4G041500	Catechol oxidase / Tyrosinase
4	1959185	China (ref) vs Japan	XP-EHH	4.08	4.01	exonic	Prupe.4G041500	Catechol oxidase / Tyrosinase
4	5621253	Japan (ref) vs Taiwan	XP-EHH	-4.24	4.31	downstream	Prupe.4G107700	Phloem protein 2 (PP2)
4	5621275	Japan (ref) vs Taiwan	XP-EHH	-5.36	6.63	downstream	Prupe.4G107700	Phloem protein 2 (PP2)
4	5621295	Japan (ref) vs Taiwan	XP-EHH	-5.41	6.76	downstream	Prupe.4G107700	Phloem protein 2 (PP2)
4	5621301	Japan (ref) vs Taiwan	XP-EHH	-6.30	9.02	downstream	Prupe.4G107700	Phloem protein 2 (PP2)
4	5621323	Japan (ref) vs Taiwan	XP-EHH	-6.39	9.26	exonic	Prupe.4G107700	Phloem protein 2 (PP2)
4	5621333	Japan (ref) vs Taiwan	XP-EHH	-4.45	4.69	downstream	Prupe.4G107700	Phloem protein 2 (PP2)
4	8401394	Ornamental (ref) vs Fruit	XP-EHH	-4.53	4.86	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401421	Fruit (ref) vs Small-fruit	XP-EHH	-4.77	5.34	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401421	Ornamental (ref) vs Fruit	XP-EHH	-4.77	4.28	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401426	Ornamental (ref) vs Fruit	XP-EHH	-4.86	5.53	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401426	Fruit (ref) vs Small-fruit	XP-EHH	-4.83	5.46	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401471	Ornamental (ref) vs Fruit	XP-EHH	-4.72	5.24	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401471	Fruit (ref) vs Small-fruit	XP-EHH	-4.85	5.50	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401493	Ornamental (ref) vs Fruit	XP-EHH	-5.14	6.14	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401493	Fruit (ref) vs Small-fruit	XP-EHH	-4.11	4.07	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	8401507	Ornamental (ref) vs Fruit	XP-EHH	-5.48	6.93	exonic	Prupe.4G146400	GLUTATHIONE S-TRANSFERASE, GST, SUPERFAMILY, GST DOMAIN CONTAINING // SUBFAMILY NOT NAMED
4	9062970	Ornamental (ref) vs Fruit	XP-EHH	4.84	5.48	exonic	Prupe.4G157900	Protein tyrosine kinase (Pkinase_Tyr) // Di-glucose binding within endoplasmic reticulum (Malectin) // Leucine rich repeat (LRR_8)
4	9062983	Ornamental (ref) vs Fruit	XP-EHH	4.93	5.67	exonic	Prupe.4G157900	Protein tyrosine kinase (Pkinase_Tyr) // Di-glucose binding within endoplasmic reticulum (Malectin) // Leucine rich repeat (LRR_8)
4	9063070	Ornamental (ref) vs Fruit	XP-EHH	5.55	7.08	intronic	Prupe.4G157900	Protein tyrosine kinase (Pkinase_Tyr) // Di-glucose binding within endoplasmic reticulum (Malectin) // Leucine rich repeat (LRR_8)
4	9063109	Ornamental (ref) vs Fruit	XP-EHH	6.71	10.18	intronic	Prupe.4G157900	Protein tyrosine kinase (Pkinase_Tyr) // Di-glucose binding within endoplasmic reticulum (Malectin) // Leucine rich repeat (LRR_8)
4	9063121	Ornamental (ref) vs Fruit	XP-EHH	4.98	5.78	intronic	Prupe.4G157900	Protein tyrosine kinase (Pkinase_Tyr) // Di-glucose binding within endoplasmic reticulum (Malectin) // Leucine rich repeat (LRR_8)
4	9063122	Ornamental (ref) vs Fruit	XP-EHH	5.02	5.86	intronic	Prupe.4G157900	Protein tyrosine kinase (Pkinase_Tyr) // Di-glucose binding within endoplasmic reticulum (Malectin) // Leucine rich repeat (LRR_8)
4	10038561	Ornamental (ref) vs Fruit	XP-EHH	-4.09	4.03	exonic	Prupe.4G170000	NUMOD3 motif (2 copies) (NUMOD3)
4	10038612	Ornamental (ref) vs Fruit	XP-EHH	-4.13	4.10	exonic	Prupe.4G170000	NUMOD3 motif (2 copies) (NUMOD3)
4	10038677	Ornamental (ref) vs Fruit	XP-EHH	-5.60	7.21	exonic	Prupe.4G170000	NUMOD3 motif (2 copies) (NUMOD3)
4	10038686	Ornamental (ref) vs Fruit	XP-EHH	-4.39	4.59	exonic	Prupe.4G170000	NUMOD3 motif (2 copies) (NUMOD3)
4	15529247	Japan (ref) vs Taiwan	XP-EHH	-8.39	15.67	intronic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //
4	15529286	Japan (ref) vs Taiwan	XP-EHH	-8.49	16.07	exonic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //
4	15529291	Japan (ref) vs Taiwan	XP-EHH	-8.80	17.21	exonic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //
4	15529294	Japan (ref) vs Taiwan	XP-EHH	-8.60	16.47	exonic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //
4	15529313	Japan (ref) vs Taiwan	XP-EHH	-8.60	16.47	exonic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //
4	15529326	Japan (ref) vs Taiwan	XP-EHH	-7.75	13.45	exonic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //

Table S6. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)[1]	*P* value ($-\log_{10} p$)[2]	Location[3]	Nearest gene[4]	Gene description (Phytozome v12.1)[5]
4	15529333	Japan (ref) vs Taiwan	XP-EHH	-6.81	10.46	exonic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //
4	15529334	Japan (ref) vs Taiwan	XP-EHH	-6.78	10.39	exonic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //
4	15529362	Japan (ref) vs Taiwan	XP-EHH	-6.77	10.34	exonic	Prupe.4G237800	Protein tyrosine kinase (Pkinase_Tyr) // Domain of unknown function (DUF3403) (DUF3403) //
4	15542556	Ornamental (ref) vs Fruit	XP-EHH	4.82	5.44	exonic	Prupe.4G238000	Cytochrome P450 CYP2 subfamily
4	15542564	Ornamental (ref) vs Fruit	XP-EHH	4.70	5.19	exonic	Prupe.4G238000	Cytochrome P450 CYP2 subfamily
4	15648565	Japan (ref) vs Taiwan	XP-EHH	-4.15	4.15	exonic	Prupe.4G238600	Premnaspirodiene oxygenase / Hyoscymus muticus premnaspirodiene oxygenase
4	15648589	Japan (ref) vs Taiwan	XP-EHH	-5.80	7.71	exonic	Prupe.4G238600	Premnaspirodiene oxygenase / Hyoscymus muticus premnaspirodiene oxygenase
5	1902477	Japan (ref) vs Taiwan	XP-EHH	-4.13	4.10	exonic	Prupe.5G017800	Methanol O-anthraniloyltransferase
5	7735854	Japan (ref) vs Taiwan	XP-EHH	-4.12	4.09	exonic	Prupe.5G063500	Cullin 1 (CUL1, CDC53)
5	16825308	Fruit (ref) vs Small-fruit	XP-EHH	-5.50	6.97	exonic	Prupe.5G212400	STRICTOSIDINE SYNTHASE-RELATED // SUBFAMILY NOT NAMED
5	16825397	Fruit (ref) vs Small-fruit	XP-EHH	-4.71	5.21	exonic	Prupe.5G212400	STRICTOSIDINE SYNTHASE-RELATED // SUBFAMILY NOT NAMED
6	677264	Fruit (ref) vs Small-fruit	XP-EHH	4.36	4.53	exonic	Prupe.6G008400	GLUCOSYLTRANSFERASE-LIKE PROTEIN-RELATED
6	681698	Fruit (ref) vs Small-fruit	XP-EHH	4.75	5.31	exonic	Prupe.6G008600	GLUCOSYLTRANSFERASE-LIKE PROTEIN-RELATED
6	681806	Fruit (ref) vs Small-fruit	XP-EHH	4.40	4.61	exonic	Prupe.6G008600	GLUCOSYLTRANSFERASE-LIKE PROTEIN-RELATED
6	683508	Fruit (ref) vs Small-fruit	XP-EHH	4.32	4.46	exonic	Prupe.6G008700	GLUCOSYLTRANSFERASE-LIKE PROTEIN-RELATED
6	683514	Fruit (ref) vs Small-fruit	XP-EHH	4.32	4.46	exonic	Prupe.6G008700	GLUCOSYLTRANSFERASE-LIKE PROTEIN-RELATED
6	683632	Fruit (ref) vs Small-fruit	XP-EHH	4.64	5.06	exonic	Prupe.6G008700	GLUCOSYLTRANSFERASE-LIKE PROTEIN-RELATED
6	683705	Fruit (ref) vs Small-fruit	XP-EHH	4.28	4.38	exonic	Prupe.6G008700	GLUCOSYLTRANSFERASE-LIKE PROTEIN-RELATED
6	683790	Fruit (ref) vs Small-fruit	XP-EHH	4.16	4.15	exonic	Prupe.6G008700	GLUCOSYLTRANSFERASE-LIKE PROTEIN-RELATED
6	2485807	Ornamental (ref) vs Fruit	XP-EHH	-4.10	4.04	intergenic	-	-
6	2485818	Ornamental (ref) vs Fruit	XP-EHH	-4.37	4.55	intergenic	-	-
6	2485825	Ornamental (ref) vs Fruit	XP-EHH	-4.42	4.65	intergenic	-	-
6	3909678	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.72	5.23	3'-UTR	Prupe.6G055300	GIBBERELLIN OXIDASE-LIKE PROTEIN-RELATED
6	3913852	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.79	5.38	intronic	Prupe.6G055400	PHD finger-like domain-containing protein 5A (PHF5A)
6	3913856	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.93	5.67	intronic	Prupe.6G055400	PHD finger-like domain-containing protein 5A (PHF5A)
6	3913867	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	5.13	6.12	intronic	Prupe.6G055400	PHD finger-like domain-containing protein 5A (PHF5A)
6	3921361	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.77	5.34	exonic	Prupe.6G055600	NADH:ubiquinone reductase (non-electrogenic) / Ubiquinone reductase // NADH dehydrogenase / Type I dehydrogenase
6	3921439	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	5.09	6.02	exonic	Prupe.6G055600	NADH:ubiquinone reductase (non-electrogenic) / Ubiquinone reductase // NADH dehydrogenase / Type I dehydrogenase
6	3921514	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.86	5.54	3'-UTR	Prupe.6G055600	NADH:ubiquinone reductase (non-electrogenic) / Ubiquinone reductase // NADH dehydrogenase / Type I dehydrogenase
6	3953600	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.22	4.27	exonic	Prupe.6G056200	DUF761-associated sequence motif (VARLMGL)
6	3953608	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.30	4.41	exonic	Prupe.6G056200	DUF761-associated sequence motif (VARLMGL)
6	3975087	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.45	4.71	exonic	Prupe.6G056200	DUF761-associated sequence motif (VARLMGL)
6	3975093	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.34	4.50	intronic	Prupe.6G056500	30S RIBOSOMAL PROTEIN S18
6	4063377	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.12	4.08	exonic	Prupe.6G058100	DEHYDRODOLICHYL DIPHOSPHATE SYNTHASE 1-RELATED
6	4091457	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.45	4.70	3'-UTR	Prupe.6G058600	RNA-binding motif protein, X-linked 2 (RBMX2, IST3)
6	4091487	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.50	4.79	3'-UTR	Prupe.6G058600	RNA-binding motif protein, X-linked 2 (RBMX2, IST3)
6	4099299	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.21	4.25	exonic	Prupe.6G058900	UDP-GLYCOSYLTRANSFERASE 71C4
6	4099305	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.32	4.44	exonic	Prupe.6G058900	UDP-GLYCOSYLTRANSFERASE 71C4
6	4271775	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.18	4.19	downstream	Prupe.6G061700	EMB
6	4271932	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.68	5.16	exonic	Prupe.6G061700	EMB
6	4271965	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.65	5.09	exonic	Prupe.6G061700	EMB
6	4278744	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	5.01	5.85	exonic	Prupe.6G061900	Acyl-coenzyme A thioesterase 13 [EC:3.1.2.-] (ACOT13)
6	4287542	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.77	5.34	exonic	Prupe.6G062000	Acyl-coenzyme A thioesterase 13 [EC:3.1.2.-] (ACOT13)
6	4287646	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.83	5.47	exonic	Prupe.6G062000	Acyl-coenzyme A thioesterase 13 [EC:3.1.2.-] (ACOT13)
6	4287698	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.87	5.55	3'-UTR	Prupe.6G062000	Acyl-coenzyme A thioesterase 13 [EC:3.1.2.-] (ACOT13)
6	4287781	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.84	5.49	4'-UTR	Prupe.6G062000	Acyl-coenzyme A thioesterase 13 [EC:3.1.2.-] (ACOT13)

Table S6. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)[1]	*P* value ($-\log_{10} p$)[2]	Location[3]	Nearest gene[4]	Gene description (Phytozome v12.1)[5]
6	4313513	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.55	4.90	intronic	Prupe.6G062300	Acyl-coenzyme A thioesterase 13 [EC:3.1.2.-] (ACOT13)
6	4314118	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.71	5.21	exonic	Prupe.6G062300	Acyl-coenzyme A thioesterase 13 [EC:3.1.2.-] (ACOT13)
6	4360894	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	4.97	5.76	intronic	Prupe.6G063200	MITOCHONDRIAL UNCOUPLING PROTEIN 2
6	4360952	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	5.09	6.03	intronic	Prupe.6G063200	MITOCHONDRIAL UNCOUPLING PROTEIN 2
6	4361093	Fruit (ref) vs Small-fruit	XP-EHH, SweeD	5.12	6.09	intronic	Prupe.6G063200	MITOCHONDRIAL UNCOUPLING PROTEIN 2
6	8020807	China (ref) vs Japan	XP-EHH	4.96	5.75	intronic	Prupe.6G111800	-
6	8020820	China (ref) vs Japan	XP-EHH	5.68	7.41	intronic	Prupe.6G111800	-
6	8020832	China (ref) vs Japan	XP-EHH	5.27	6.43	intronic	Prupe.6G111800	-
6	8020837	China (ref) vs Japan	XP-EHH	4.77	5.34	intronic	Prupe.6G111800	-
6	8020840	China (ref) vs Japan	XP-EHH	4.74	5.28	intronic	Prupe.6G111800	-
6	8021028	China (ref) vs Japan	XP-EHH	5.97	8.14	intronic	Prupe.6G111800	-
6	8021034	China (ref) vs Japan	XP-EHH	6.06	8.37	intronic	Prupe.6G111800	-
6	8021063	China (ref) vs Japan	XP-EHH	6.47	9.48	intronic	Prupe.6G111800	-
6	8021166	China (ref) vs Japan	XP-EHH	6.92	10.80	intronic	Prupe.6G111800	-
6	8021201	China (ref) vs Japan	XP-EHH	6.99	11.02	intronic	Prupe.6G111800	-
6	8021203	China (ref) vs Japan	XP-EHH	6.99	11.00	intronic	Prupe.6G111800	-
6	8021222	China (ref) vs Japan	XP-EHH	6.55	9.72	intronic	Prupe.6G111800	-
6	8021247	China (ref) vs Japan	XP-EHH	6.53	9.67	intronic	Prupe.6G111800	-
6	8021268	China (ref) vs Japan	XP-EHH	6.43	9.38	intronic	Prupe.6G111800	-
6	8021281	China (ref) vs Japan	XP-EHH	6.03	8.29	intronic	Prupe.6G111800	-
6	8021284	China (ref) vs Japan	XP-EHH	6.52	9.63	intronic	Prupe.6G111800	-
6	8021299	China (ref) vs Japan	XP-EHH	5.81	7.73	exonic	Prupe.6G111800	-
6	8021337	China (ref) vs Japan	XP-EHH	5.62	7.25	intronic	Prupe.6G111800	-
6	8021374	China (ref) vs Japan	XP-EHH	4.76	5.33	intronic	Prupe.6G111800	-
6	13410619	Japan (ref) vs Taiwan	XP-EHH	-6.32	9.08	downstream	Prupe.6G154000	-
6	19101213	Japan (ref) vs Taiwan	XP-EHH	-8.02	14.37	intronic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101239	Japan (ref) vs Taiwan	XP-EHH	-7.48	12.55	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101246	Japan (ref) vs Taiwan	XP-EHH	-6.79	10.42	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101249	Japan (ref) vs Taiwan	XP-EHH	-6.77	10.35	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101251	Japan (ref) vs Taiwan	XP-EHH	-6.32	9.07	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101272	Japan (ref) vs Taiwan	XP-EHH	-6.53	9.65	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101276	Japan (ref) vs Taiwan	XP-EHH	-5.76	7.60	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101278	Japan (ref) vs Taiwan	XP-EHH	-5.76	7.60	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101284	Japan (ref) vs Taiwan	XP-EHH	-5.35	6.61	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101320	Japan (ref) vs Taiwan	XP-EHH	-5.02	5.88	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	19101325	Japan (ref) vs Taiwan	XP-EHH	-4.40	4.60	exonic	Prupe.6G183600	CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 28-RELATED
6	20213387	China (ref) vs Japan	XP-EHH	-4.12	4.08	exonic	Prupe.6G194900	Disease resistance protein RPM1 (RPM1, RPS3)
6	20213399	China (ref) vs Japan	XP-EHH	-4.17	4.17	exonic	Prupe.6G194900	Disease resistance protein RPM1 (RPM1, RPS3)
6	20213467	Ornamental (ref) vs Fruit	XP-EHH	-4.14	4.13	exonic	Prupe.6G194900	Disease resistance protein RPM1 (RPM1, RPS3)
6	22336225	China (ref) vs Japan	XP-EHH	4.25	4.32	intronic	Prupe.6G215800	PECTINESTERASE 10-RELATED
6	22336248	China (ref) vs Japan	XP-EHH	4.31	4.43	intronic	Prupe.6G215800	PECTINESTERASE 10-RELATED
6	23183155	Japan (ref) vs Taiwan	XP-EHH	-4.22	4.26	exonic	Prupe.6G227800	Protein tyrosine kinase (Pkinase_Tyr) // Di-glucose binding within endoplasmic reticulum (Malectin)
6	25071959	China (ref) vs Japan	XP-EHH	-4.09	4.03	exonic	Prupe.6G259900	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED
6	25164308	China (ref) vs Japan	XP-EHH	-5.02	5.87	exonic	Prupe.6G261400	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED
6	25164314	China (ref) vs Japan	XP-EHH	-5.57	7.15	exonic	Prupe.6G261400	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED
6	25164317	China (ref) vs Japan	XP-EHH	-5.36	6.63	exonic	Prupe.6G261400	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED
6	25164324	China (ref) vs Japan	XP-EHH	-5.47	6.89	exonic	Prupe.6G261400	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED
6	25164328	China (ref) vs Japan	XP-EHH	-5.43	6.79	exonic	Prupe.6G261400	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED
6	25164343	China (ref) vs Japan	XP-EHH	-6.11	8.49	exonic	Prupe.6G261400	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED
6	25164364	China (ref) vs Japan	XP-EHH	-6.13	8.55	exonic	Prupe.6G261400	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED
6	25164367	China (ref) vs Japan	XP-EHH	-4.17	4.17	exonic	Prupe.6G261400	L-TYPE LECTIN-DOMAIN CONTAINING RECEPTOR KINASE IV.1-RELATED

Table S6. (Continued).

Chr.	Position (bp)	Population	Test for selection	Value for EHH test (normalized)[1]	*P* value ($-\log_{10} p$)[2]	Location[3]	Nearest gene[4]	Gene description (Phytozome v12.1)[5]
6	25384222	Fruit (ref) vs Small-fruit	XP-EHH	-4.19	4.20	exonic	Prupe.6G264900	GLUTATHIONE S-TRANSFERASE U1-RELATED
7	9757769	Ornamental (ref) vs Fruit	XP-EHH	-4.22	4.27	intronic	Prupe.7G058100	ATP-binding cassette, subfamily A (ABC1), member 3 (ABCA3)
7	9757801	Ornamental (ref) vs Fruit	XP-EHH	-5.73	7.54	intronic	Prupe.7G058100	ATP-binding cassette, subfamily A (ABC1), member 3 (ABCA3)
7	9757803	Ornamental (ref) vs Fruit	XP-EHH	-5.62	7.25	intronic	Prupe.7G058100	ATP-binding cassette, subfamily A (ABC1), member 3 (ABCA3)
7	9757817	Ornamental (ref) vs Fruit	XP-EHH	-5.37	6.66	intronic	Prupe.7G058100	ATP-binding cassette, subfamily A (ABC1), member 3 (ABCA3)
7	15413121	Ornamental (ref) vs Fruit	XP-EHH	-4.89	5.59	intronic	Prupe.7G137200	Galactolipase
7	15413124	Ornamental (ref) vs Fruit	XP-EHH	-5.45	6.86	intronic	Prupe.7G137200	Galactolipase
7	15413130	Ornamental (ref) vs Fruit	XP-EHH	-5.01	5.84	intronic	Prupe.7G137200	Galactolipase
7	15413140	Ornamental (ref) vs Fruit	XP-EHH	-4.83	5.46	intronic	Prupe.7G137200	Galactolipase
7	15413165	Ornamental (ref) vs Fruit	XP-EHH	-4.54	4.88	intronic	Prupe.7G137200	Galactolipase
7	15413166	Ornamental (ref) vs Fruit	XP-EHH	-4.59	4.98	intronic	Prupe.7G137200	Galactolipase
7	15413178	Ornamental (ref) vs Fruit	XP-EHH	-4.59	4.97	intronic	Prupe.7G137200	Galactolipase
7	15413195	Ornamental (ref) vs Fruit	XP-EHH	-4.42	4.65	intronic	Prupe.7G137200	Galactolipase
7	15413199	Ornamental (ref) vs Fruit	XP-EHH	-4.47	4.73	intronic	Prupe.7G137200	Galactolipase
7	15413206	Ornamental (ref) vs Fruit	XP-EHH	-4.26	4.34	intronic	Prupe.7G137200	Galactolipase
7	15413252	Ornamental (ref) vs Fruit	XP-EHH	-4.19	4.21	exonic	Prupe.7G137200	Galactolipase
7	15413256	Ornamental (ref) vs Fruit	XP-EHH	-4.48	4.76	exonic	Prupe.7G137200	Galactolipase
7	18890311	Ornamental (ref) vs Fruit	XP-EHH	4.09	4.03	exonic	Prupe.7G203800	Ankyrin repeats (3 copies) (Ank_2) // Ankyrin repeats (many copies) (Ank_4)
7	21626124	China (ref) vs Japan	XP-EHH	4.29	4.39	intronic	Prupe.7G259500	EXPRESSED PROTEIN
7	21626135	China (ref) vs Japan	XP-EHH	4.23	4.29	intronic	Prupe.7G259500	EXPRESSED PROTEIN
8	4922862	China (ref) vs Japan	XP-EHH	4.13	4.10	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922863	China (ref) vs Japan	XP-EHH	4.13	4.09	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922874	China (ref) vs Japan	XP-EHH	4.39	4.59	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922877	China (ref) vs Japan	XP-EHH	4.35	4.50	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922882	China (ref) vs Japan	XP-EHH	4.35	4.50	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922886	China (ref) vs Japan	XP-EHH	4.70	5.19	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922898	China (ref) vs Japan	XP-EHH	4.52	4.84	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922948	China (ref) vs Japan	XP-EHH	4.50	4.79	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922956	China (ref) vs Japan	XP-EHH	4.56	4.91	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922959	China (ref) vs Japan	XP-EHH	4.54	4.88	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922961	China (ref) vs Japan	XP-EHH	4.64	5.07	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922966	China (ref) vs Japan	XP-EHH	4.29	4.40	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922967	China (ref) vs Japan	XP-EHH	4.29	4.40	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	4922978	China (ref) vs Japan	XP-EHH	4.50	4.80	exonic	Prupe.8G046600	LEUCINE-RICH REPEAT-CONTAINING PROTEIN // SUBFAMILY NOT NAMED
8	9209192	Ornamental (ref) vs Fruit	XP-EHH	-4.91	5.62	intergenic	-	-
8	9209208	Ornamental (ref) vs Fruit	XP-EHH	-4.92	5.66	intergenic	-	-
8	9209219	Ornamental (ref) vs Fruit	XP-EHH	-4.70	5.20	intergenic	-	-
8	10643051	Japan (ref) vs Taiwan	XP-EHH	-6.06	8.37	exonic	Prupe.8G072900	HVA22/DP1 gene product-related proteins
8	10643053	Japan (ref) vs Taiwan	XP-EHH	-6.92	10.80	exonic	Prupe.8G072900	HVA22/DP1 gene product-related proteins
8	10643056	Japan (ref) vs Taiwan	XP-EHH	-6.92	10.80	exonic	Prupe.8G072900	HVA22/DP1 gene product-related proteins
8	10643072	Japan (ref) vs Taiwan	XP-EHH	-5.65	7.32	exonic	Prupe.8G072900	HVA22/DP1 gene product-related proteins
8	13885783	Ornamental (ref) vs Fruit	XP-EHH	4.88	5.57	intronic	Prupe.8G108400	LONG CHAIN ACYL-COA SYNTHETASE 2
8	13885832	Ornamental (ref) vs Fruit	XP-EHH	4.88	5.57	exonic	Prupe.8G108400	LONG CHAIN ACYL-COA SYNTHETASE 2
8	17899889	Japan (ref) vs Taiwan	XP-EHH	-4.63	5.05	exonic	Prupe.8G177300	Cytochrome P450 CYP4/CYP19/CYP26 subfamilies

[1] In XP-EHH test, negative value means that reference (ref) population has longer haplotype than the other population at the core SNP.

[2] *P* values based on long haplotype tests: nSL or XP-EHH. The strongest ($-\log_{10} p > 4$) candidates are listed.

[3] Upstream and downstream indicate the regions within 5 kb from the coding sequence.

[4] Nearest gene detected for SNPs (excluding intergenic locations) based on *P. persica* genome.

[5] Gene description is from Phytozome v12.1 (https://phytozome.jgi.doe.gov/pz/portal.html).

Acknowledgements

I deeply grateful to Dr. Takashige Ishii, Professor, Kobe University, for courteous supervision and providing valuable comments and suggestions, as well as Dr. Takeshi Yasuda, Professor, Kobe University, and Dr. Hironori Katayama, Associate Professor, Kobe University, for their careful reading of my thesis and giving critical comments. I would like to express my gratitude to Dr. Ryo Ishikawa, Associate Professor, Kobe University, for invaluable discussions and continuous encouragement.

I sincerely acknowledge Dr. Takashi Akagi, Associate Professor, Okayama University, for his technical advice regarding exon-targeted resequencing and subsequent population genetic analyses. I deeply thank Dr. Shogo Ishio, Chief Researcher, Sumitomo Forestry Co. Ltd., for his technical advice on microsatellite marker development and analyses. I also thank all past and present members of the Laboratory of Plant Breeding, Kobe University, for their helpful discussions.

I offer special thanks to Mr. Yoshihisa Tanimoto, Mr. Yasuhiro Katayama, Mr. Akihiro Nobata and Mr. Masayoshi Takenaka, the successive chiefs of the Japanese Apricot Laboratory, Wakayama Fruit Experiment Station, for encouraging me to pursue doctoral studies and their continuous support. I thank all past and present members of the Japanese Apricot Laboratory, for their management of plant materials, helpful discussions and supports.

Finally, I deeply appreciate my wife and parents for their continuous encouragement and support.

2020

Koji Numaguchi