

RESEARCH ARTICLE

Fine mapping of the male-sterile genes (*MS1*, *MS2*, *MS3*, and *MS4*) and development of SNP markers for marker-assisted selection in Japanese cedar (*Cryptomeria japonica* D. Don)

Yoichi Hasegawa¹, Saneyoshi Ueno², Asako Matsumoto², Tokuko Ujino-Ihara², Kentaro Uchiyama², Satoko Totsuka³, Junji Iwai³, Tetsuji Hakamata⁴, Yoshinari Moriguchi^{1*}

1 Graduate School of Science and Technology, Niigata University, Niigata, Niigata, Japan, **2** Department of Forest Molecular Genetics and Biotechnology, Forestry and Forest Products Research Institute, Forest Research and Management Organization, Tsukuba, Ibaraki, Japan, **3** Niigata Prefectural Forest Research Institute, Murakami, Niigata, Japan, **4** Forestry and Forest Products Research Center, Shizuoka Prefectural Research Institute of Agriculture and Forestry, Hamamatsu, Shizuoka, Japan

* chimori@agr.niigata-u.ac.jp



OPEN ACCESS

Citation: Hasegawa Y, Ueno S, Matsumoto A, Ujino-Ihara T, Uchiyama K, Totsuka S, et al. (2018) Fine mapping of the male-sterile genes (*MS1*, *MS2*, *MS3*, and *MS4*) and development of SNP markers for marker-assisted selection in Japanese cedar (*Cryptomeria japonica* D. Don). PLoS ONE 13(11): e0206695. <https://doi.org/10.1371/journal.pone.0206695>

Editor: Chunxian Chen, USDA/ARS, UNITED STATES

Received: July 24, 2018

Accepted: October 17, 2018

Published: November 15, 2018

Copyright: © 2018 Hasegawa et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: We uploaded the microarray probe data set to Gene Expression Omnibus (GEO) and cited the accession number (GPL25617, P7 L162).

Funding: This work was supported by a Science and Technology Research Promotion Program for Agriculture, Forestry, Fisheries, and Food Industry and the NARO Bio-oriented Technology Research

Abstract

Pollinosis caused by Japanese cedar (*Cryptomeria japonica*) is a widespread social problem in Japan. To date, 23 male-sterile *C. japonica* trees have been selected in Japan to address pollinosis, from which four male-sterility loci (*MS1*, *MS2*, *MS3*, and *MS4*) have been identified from test crossing results. For efficient breeding of male-sterile *C. japonica* trees, more male-sterile individuals and individuals heterozygous for male-sterile genes are required. Therefore, we aimed to develop DNA markers for marker-assisted selection of four types of male-sterile genes from populations without a family structure. First, for four families exhibiting segregation of each male-sterile locus (*MS1*, *MS2*, *MS3*, and *MS4*), genome-wide single-nucleotide polymorphism and insertion/deletion (indel) genotyping was performed using the Axiom myDesign Targeted Genotyping Array method. Four high-density linkage maps for mapping the *MS1*, *MS2*, *MS3*, and *MS4* families were constructed, which included 4923, 1722, 1896, and 2247 markers, respectively. In these maps, 15, 4, 2, and 2 markers were located 0.0, 3.3, 1.1, and 0.0 cM from the *MS1*, *MS2*, *MS3*, and *MS4* loci, respectively. Second, for the markers located 0.0 cM from a male-sterile locus (i.e., *MS1* and *MS4*), to clarify the most tightly linked markers, we calculated the prediction rate of male-sterile gene genotypes from marker genotypes for 78 trees. The markers with the highest prediction rates were AX-174127446 (0.95) for *MS1* and AX-174121522 (1.00) for *MS4*. The AX-174121522 marker was considered to be suitable for selecting trees homozygous or heterozygous for the *MS4* gene from plus-trees without a pollination test, which requires a large amount of time and effort. The nearest markers to the male-sterile loci found in this study may facilitate the isolation of male-sterile genes in *C. japonica* via combination with the draft genomic sequence that is currently being collated.

Advancement Institution (Research program on development of innovative technology).

Competing interests: The authors have declared that no competing interests exist.

Introduction

Marker-assisted selection (MAS) is a breeding approach used to select individuals with desirable traits based on the genotypes of markers tightly linked to target genes, as has been reported for barley [1] and grape [2]. The main advantage of MAS is that it reduces the time required for breeding cycles. This is especially attractive for tree species with longer generation times than other crop species. MAS application is considered to be difficult in forest trees (especially conifers) due to the rapid decay of linkage disequilibrium [3]. However, in recent years, the construction of high-density linkage maps that facilitate the development of MAS markers has become more efficient and economical with the help of next-generation sequencing and high-throughput array technologies, even in non-model species.

Japanese cedar (*Cryptomeria japonica* D. Don) is an allogamous, monoecious, and wind-pollinated conifer with a haploid chromosome number (n) of 11 ($2n = 22$). The DNA content of the haploid cells was estimated to be 11.045 pg/C using flow cytometry, corresponding to a 10.8 Gb haploid genome size [4]. *C. japonica* is the most important forestry species in Japan, with outstanding features such as easy processes for cutting/propagation and artificial promotion of flowering, wide adaptability to the environment, rapid growth, straight bole, and soft wood with a pleasant color and scent. It covers an area of 4.5 million hectares, accounting for 44% of all Japanese artificial forests [5].

An increase in the area of *C. japonica* planted forests has resulted in widespread *C. japonica* pollinosis, which is one of the most serious allergic diseases in Japan. In 2008, 26.5% of Japanese residents had an allergy to *C. japonica* pollen [6]. Genetically male-sterile *C. japonica* trees are expected to play an important role in reducing the amount of pollen dispersed by the breeding program. The first genetically male-sterile *C. japonica* tree, Toyama-1, the sterility of which is controlled by the *MS1* locus [7,8], was discovered in 1992 [9]. There has since been considerable effort to characterize male sterility in *C. japonica*, for example, by identifying male-sterile individuals in plus-tree populations, making artificial matings between male-sterile and plus-trees, and propagating male-sterile trees. To date, 23 genetically male-sterile trees have been identified from seven prefectures [8], and the frequency of male-sterile trees in artificial forests has been estimated as one in several thousand [10,11]. Based on the results of test crossings, four recessive male-sterile genes, *MS1*, *MS2*, *MS3*, and *MS4*, were identified [7,8,12,13].

Male-sterile trees are currently identified by observing the release of pollen or by directly examining male strobili with a magnifying glass or microscope. Heterozygotes of male-sterile genes have been identified by examining the segregation data of the progeny arising from artificial crosses. Such trees are important breeding material for the production of male-sterile *C. japonica* trees. Most plus-trees heterozygous for a male-sterile gene found in Japan (e.g., Suzu-2, Ooi-7, Naka-4, and Ohara-13) are controlled by *MS1* because test crossing has been performed using Toyama-1. Therefore, heterozygous plus-trees for a male-sterile gene controlled by *MS2*, *MS3*, and *MS4* are required as breeding materials to avoid problems associated with inbreeding depression. Because conventional methods require time, labor, and space, it is desirable to localize the male-sterile genes on the linkage map and develop tightly linked DNA markers for MAS. Numerous DNA markers have been developed to promote the molecular breeding of *C. japonica* [14–23]. In addition, linkage maps were constructed with 1262 markers (average marker interval: 1.1 cM [24]), 2560 markers (0.49 cM [25]), and 3205 markers (0.47 cM [23]). Using information from these linkage maps, the male-sterile genes *MS1*, *MS2*, *MS3*, and *MS4* have been mapped onto the ninth linkage group (hereafter denoted as LG9 [24]), the fifth linkage group (LG5 [19]), the first linkage group (LG1 [25]), and the fourth linkage group (LG4 [25]), respectively.

In most high-density maps, male-sterile loci are not mapped due to the use of families unrelated to male sterility. In this study, we developed a large number of new single-nucleotide polymorphism (SNP) and insertion/deletion (indel) markers and constructed four high-density linkage maps exhibiting consistent marker order including the *MS1*, *MS2*, *MS3*, and *MS4* genes. Furthermore, we calculated the prediction rate of male-sterile gene genotypes from marker genotypes in a population without a family structure. These results allowed us to develop markers for MAS of male-sterile genes from plus-trees with superior traits (i.e., outside the mapping population).

Materials and methods

Ethics statement

All needle tissues used in this study were collected from scion gardens in Forestry and Forest Products Research Institute, Niigata University and five Prefectural Forest Experiment Stations (Niigata, Shizuoka, Ishikawa, Kanagawa, and Fukushima). Permission is obtained from all Institutes. In addition, plant tissues of natural forests were collected from the scion garden established using cutting with permission.

Reference transcript construction

Reference transcripts were constructed by clustering three sets of transcriptomes. Two sets were from RNA sequencing experiments (Ueno et al. [26] and this study) and one from full-length cDNA libraries [27]. Three sets of transcript sequences (totaling 573,944 sequences) were clustered with EvidentialGene [28], resulting in 81,510 “okay” and “okalt” transcripts (EviGene transcripts). We added 16 male sterility candidate gene homologues in *C. japonica* (Futamura, personal communication), because they were not completely represented in the EviGene transcripts (HSP for candidate gene homolog/HSP for EviGene transcript ≤ 0.6 , where HSP indicates the high-scoring segment pair in a BLAST search). In addition, 49 candidate genes detected in a genome-wide association study by Uchiyama et al. [29] were added, because they were not completely represented in the EviGene transcripts (HSP for candidate gene homolog/HSP for EviGene transcript ≤ 0.6). We used these 81,575 sequences with a mean length of 1128 bp as a reference for SNP and indel discovery.

RNA extraction, sequencing, and SNP and indel discovery

Four male-fertile trees with heterozygous male-sterile genes (Ooi-7 [*Ms1/ms1*], S1NK4 [*Ms2/ms2*], S5HK7 [*Ms3/ms3*], and S8HK5 [*Ms4/ms4*]) were used to develop SNP and indel markers using RNA sequencing. Ooi-7 was a plus-tree heterozygous for a male-sterile gene (*MS1*) found in Shizuoka Prefecture [30]. S1NK4 was an F₁ of Shindai-1 (male-sterile, *ms2/ms2*) and Nakakubiki-4 (male-fertile, *Ms2/Ms2*), where male sterility in Shindai-1 was controlled by the *MS2* locus [12]. S5HK7 was an F₁ of Shindai-5 (male-sterile, *ms3/ms3*) and Higashikanbara-7 (male-fertile, *Ms3/Ms3*), where male sterility in Shindai-5 was controlled by the *MS3* locus [12]. S8HK5 was an F₁ of Shindai-8 (male-sterile, *ms4/ms4*) and Higashikanbara-5 (male-fertile, *Ms4/Ms4*), where male sterility of Shindai-8 was controlled by the *MS4* locus [13]. For S1NK4 and S5HK7, RNA was extracted from the male strobili, inner bark, and needles. However, for Ooi-7 and S8HK5, RNA was extracted from only the inner bark and needles. Total RNA was extracted based on the method of Le Provost et al. [31]. Approximately ten spoonfuls of powder were added to 1 mL of CTAB extraction buffer. Extracted RNA was further purified with an SV total RNA purification system (Promega, Madison, WI, USA). Sequencing libraries were constructed and sequenced using a HiSeq 2500 (Illumina, San Diego, CA, USA) in

2 × 100-bp paired-end sequencing by Hokkaido System Science Co., Ltd (Sapporo, Japan). SNP and indel discovery was performed using Genedata Expressionist Genomic Profiling 9.1.4a software (Genedata, Lexington, MA, USA) at Takara Bio Inc. (Otsu, Shiga, Japan). First, 441 million reads were obtained by removing low-quality regions and sequence adaptor-derived regions using Trimmomatic 0.32 [32]. Following removal, sequence reads were mapped to the reference sequences described in the previous section using BWA-MEM 0.7.12 [33]. Low-quality and duplicate reads were removed from the mapping results and realigned. The differences between the reference sequence and read sequence were detected as genetic mutations. Of the 996,871 SNP and 91,825 indel mutations, 775,550 SNP and 64,807 indel mutations were heterozygous. The 69,199 heterozygous SNPs and 4235 heterozygous indels were used for Axiom myDesign Targeted Genotyping Array (Affymetrix; Thermo Fisher Scientific, Waltham, MA, USA).

DNA extraction and axion genotyping

High-density linkage maps of *C. japonica* were constructed for the F1O7 (84 individuals), S1-2 (94 individuals [19,25]), S5HK7 (94 individuals [25]), and S8HK5 (96 individuals [25]) families (Table 1). The F1O7 family was derived from a cross between Fukushima-1 (male-sterile, *ms1/ms1*) and Ooi-7, where male sterility in Fukushima-1 is controlled by the *MS1* locus [8,10]. The S1-2 family was derived from a backcross between Shindai-1 and the F₁ plant S1NK4. The S5HK7 family was derived from a backcross between Shindai-5 and the F₁ plant S5HK7. Finally, the S8HK5 family was derived from a backcross between Shindai-8 and the F₁ plant S8HK5.

In addition to the progeny and parents of the four mapping populations, four male-sterile trees, seven trees heterozygous for male-sterile genes, parents of the previous linkage map, seven trees from breeding materials such as plus-trees and cutting cultivars, and 50 trees from natural forests were used to calculate the prediction rate of male-sterile gene genotypes from marker genotypes (for a total of 78 trees; S1 Table).

Needle tissue was collected from all 446 trees (i.e., four mapping populations and 78 trees previously mentioned). Genomic DNA was extracted from these needles using a modified CTAB method [34]. Using Axiom myDesign Targeted Genotyping Array with 73,434 markers (GPL25617), information on the genotypes of the 446 trees was obtained according to the manufacturer's instructions. Genotyping was performed using Axiom Analysis Suite ver. 2.0.0.35 software. Markers with an off-target variant or Fisher's linear discriminant < 7 were not used for subsequent analysis.

Linkage map construction

Chi-squared tests were performed for each locus to assess the deviation from the expected Mendelian segregation ratio. Loci showing extreme segregation distortion ($P < 0.001$) and

Table 1. Summary of the families used to construct the linkage maps in this study.

Family name	Number of individuals	Maternal plant	Paternal plant
F1O7	84	Fukushima-1 [<i>ms1/ms1</i>]	Ooi-7 [<i>Ms1/ms1</i>]
S1-2	94	Shindai-1 [<i>ms2/ms2</i>]	S1NK4 (Shindai-1 × Nakakubiki-4) F ₁ [<i>Ms2/ms2</i>]
S5HK7	94	Shindai-5 [<i>ms3/ms3</i>]	S5HK7 (Shindai-5 × Higashikanbara-7) F ₁ [<i>Ms3/ms3</i>]
S8HK5	96	Shindai-8 [<i>ms4/ms4</i>]	S8HK5 (Shindai-8 × Higashikanbara-5) F ₁ [<i>Ms4/ms4</i>]

<https://doi.org/10.1371/journal.pone.0206695.t001>

with many missing data points (more than three individuals) were excluded from further linkage analysis.

The linkage analyses were performed using the maximum likelihood mapping algorithm of JoinMap ver. 4.1 software (Kyazma, Wageningen, The Netherlands) with a cross pollination (CP)-type population (i.e., $hk \times hk$, $lm \times ll$, and $nn \times np$) for the F1O7 family and backcross (BC)-type population (i.e., $nn \times np$) for the S1-2, S5HK7, and S8HK5 families. The resulting 4923, 1722, 1897, and 2247 markers were used for linkage analysis of the F1O7, S1-2, S5HK7, and S8HK5 families, respectively. In all mapping populations, the trait data of male-sterile and male-fertile trees were recorded as homozygous and heterozygous genotypes, respectively, according to the method of Moriguchi et al. [19,24,25] (S2–S5 Tables). Markers were initially assigned to tentative linkage groups using logarithm of odds ratio (LOD) thresholds of 3.0–9.0 in increments of 1.0. LOD thresholds of 6.0, 4.0, 4.0, and 5.0 for groups of markers were defined for the F1O7, S1-2, S5HK7, and S8HK5 families, respectively. Map distance was calculated using the Kosambi mapping function. Default settings were used for the recombination frequency threshold and ripple value. The linkage group numbers (LG1–LG11) were defined following the linkage maps published by Tani et al. [35] and Moriguchi et al. [24]. Images of the linkage groups were drawn using MapChart ver. 2.0 [36].

Prediction rate of male-sterile gene genotypes from marker genotypes

We found 15 and 2 markers located 0.0 cM from the *MS1* and *MS4* genes, respectively (see Results and discussion). To clarify the most tightly linked markers, we calculated the prediction rate of male-sterile gene genotypes from marker genotypes for 78 trees. The prediction rate was defined as the ratio of the number of individuals with no inconsistencies between marker genotype and the genotype of the male-sterility locus to the number of all individuals.

Results and discussion

High-density linkage maps

In total, 10,808 (F1O7 family), 11,442 (S1-2 family), 10,882 (S5HK7 family), and 10,913 (S8HK5 family) of the 73,434 markers on the Axiom array were successfully genotyped in this study (Table 2). In the F1O7 family (CP type), 4959 markers showed polymorphisms. In S1-2, S5HK7, and S8HK5 families (BC types), 1742, 2118, and 2363 markers showed polymorphisms, respectively. After excluding markers with large amounts of missing data and high distortions in the segregation ratio, 4923 (F1O7 family), 1719 (S1-2 family), 1897 (S5HK7 family), and 2247 (S8HK5 family) markers were ultimately used for the linkage analysis. The total length of the linkage maps constructed in this study for the F1O7, S1-2, S5HK7, and S8HK5 families was 1575.5, 1290.1, 1168.6, and 2350.9 cM and the average interval between markers was 0.32, 0.75, 0.62, and 1.05 cM, respectively (S1–S4 Figs and S2–S5 Tables). These maps had the shortest mean distance between adjacent markers in the *C. japonica* linkage map, including

Table 2. Numbers of markers used to construct the four linkage maps.

Family	Number of	Number of markers	Number of polymorphic markers between parents		Number of markers	Number of
name	analyzed markers	successfully genotyped	CP type	BC type	used for linkage analyses	positioned markers
F1O7	73,434	10,808	4,959	-	4,923	4,923
S1-2	73,434	11,442	-	1,742	1,719	1,719
S5HK7	73,434	10,882	-	2,118	1,897	1,896
S8HK5	73,434	10,913	-	2,363	2,247	2,247

<https://doi.org/10.1371/journal.pone.0206695.t002>

the male sterility locus, reported to date. In addition, these values were similar to the linkage maps of other conifers (e.g., 0.58 cM/marker in *Pinus taeda* [37], 0.60 cM/marker in a consensus map for *P. taeda* and *Pinus elliottii* [38], 0.93 cM/marker in *Pinus pinaster* [39], and 0.92 cM/marker in a consensus map for *Picea glauca* and *Pinus marina* [40]). Since the high-density linkage maps were constructed using markers derived from the cDNA libraries of *C. japonica*, they are useful not only for quantitative trait locus (QTL) analysis of this species, but also for chromosomal synteny analysis across related species of Cupressaceae.

In total, 15, 4, 2, and 2 markers were mapped to the vicinities of the *MS1*, *MS2*, *MS3*, and *MS4* loci, respectively (Fig 1). These markers were separated from their associated male-sterility genes by 0.0, 3.3, 1.1, and 0.0 cM, respectively, in the F1O7, S1-2, S5HK7, and S8HK5 families. Several other studies have attempted to develop markers closely linked to male-sterile genes in *C. japonica*. Moriguchi et al. [25] reported the closely linked markers for four male-sterile genes, in which the closest markers to the *MS1*, *MS2*, *MS3*, and *MS4* male-sterile loci were estSNP04188 (1.8 cM in the T5 family of 173 individuals), estSNP00695 (7.0 cM in the S1-2 family of 128 individuals), gSNP05415 (3.1 cM in the S5HK7 family of 167 individuals), and estSNP01408 (7.0 cM in the S8HK5 family of 108 individuals), respectively. In addition, Mishima et al. [23] reported markers that overlapped in the QTL region for *MS1*. Although some markers reported in this study were closer than those reported previously (see below), they are difficult to compare directly since map distance depends on the number of individuals of a mapping family. Therefore, all of these closely linked markers should be mapped using the same families with a large number of individuals to investigate their accurate map positions in

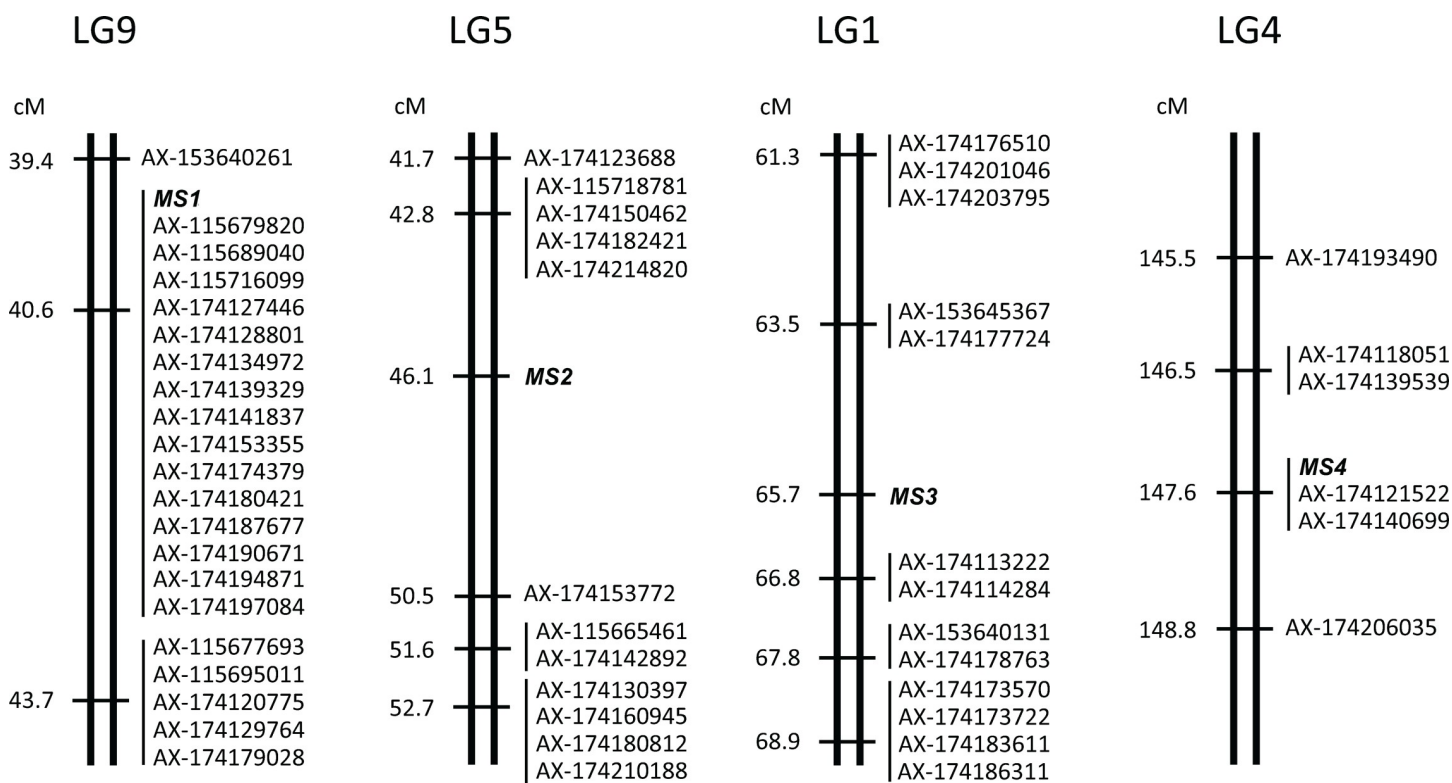


Fig 1. Partial linkage maps around the male-sterile genes *MS1*, *MS2*, *MS3*, and *MS4*. Marker names are indicated to the right of the linkage groups. Centimorgan distances are indicated to the left of each linkage group. Full linkage maps are shown in S1–S4 Figs.

<https://doi.org/10.1371/journal.pone.0206695.g001>

future studies. Determining the precise map distance of closely linked markers could enable more precise MAS selection by using two markers sandwiching a male-sterile locus.

Prediction rate of male-sterile gene genotypes from marker genotypes

We identified markers located 0.0 cM from the *MS1* and *MS4* genes and calculated the prediction rates of the male-sterile gene genotype based on the genotypes of these markers using 78 individual trees without a family structure (S1 Table). The prediction rates for the 15 markers 0 cM from the *MS1* gene were 0.29–0.95 (Table 3). Meanwhile, the prediction rates of the two markers 0 cM from the *MS4* gene were 0.69 and 1.00. The highest prediction rates for the *MS1* and *MS4* genes were 0.95 (AX-174127446) and 1.00 (AX-174121522), respectively (S5 and S6 Figs), and the prediction rates of these markers were the highest among all 73,434 markers in this study (S6 Table).

We did not calculate the prediction rates for the *MS2* and *MS3* genes, because the markers are located 0.0 cM from the *MS2* and *MS3* genes. To develop markers that are more tightly linked to the *MS2* and *MS3* genes used for MAS, it is necessary to map more markers in the S1-2 and S5HK7 families.

Table 3. Positions of the markers closest to the male-sterile genes and prediction rate of male-sterile gene genotypes from the marker genotypes of 78 *Cryptomeria japonica* trees.

Family name	Locus name		Linkage maps	Populations without family structure
(Target male-sterile gene)		Position	Map distance from	Prediction rate of male-sterile gene
		(cM)	<i>MS</i> gene (cM)	genotypes from the marker genotypes
F1O7 (<i>MS1</i>)	AX-115679820	40.6	0.0	0.32
	AX-115689040	40.6	0.0	0.82
	AX-115716099	40.6	0.0	0.65
	AX-174127446	40.6	0.0	0.95
	AX-174128801	40.6	0.0	0.53
	AX-174134972	40.6	0.0	0.35
	AX-174139329	40.6	0.0	0.69
	AX-174141837	40.6	0.0	0.82
	AX-174153355	40.6	0.0	0.78
	AX-174174379	40.6	0.0	0.33
	AX-174180421	40.6	0.0	0.32
	AX-174187677	40.6	0.0	0.58
	AX-174190671	40.6	0.0	0.41
	AX-174194871	40.6	0.0	0.47
	AX-174197084	40.6	0.0	0.29
S1-2 (<i>MS2</i>)	AX-115718781	43.9	3.3	-
	AX-174150462	43.9	3.3	-
	AX-174182421	43.9	3.3	-
	AX-174214820	43.9	3.3	-
S5HK7 (<i>MS3</i>)	AX-174113222	66.8	1.1	-
	AX-174114284	66.8	1.1	-
S8HK5 (<i>MS4</i>)	AX-174121522	147.6	0.0	1.00
	AX-174140699	147.6	0.0	0.69

The 78 trees included eight parents of the four mapping populations, four male-sterile trees, seven trees heterozygous for a male-sterile gene, two parents of the previous linkage map, seven trees from breeding materials, and 50 trees from natural forests (see S1 Table).

<https://doi.org/10.1371/journal.pone.0206695.t003>

The AX-174121522 marker was considered to be suitable for selection of trees homozygous or heterozygous for the *MS4* gene from plus-trees without test crossing, which requires a large amount of time and effort. The BLAST search revealed that cDNA containing a SNP of AX-174121522 was similar to the uroporphyrinogen-III synthase gene. With respect to the nucleotide substitution of the SNP, the amino acid substitution from threonine to asparagine was confirmed. Therefore, it is possible that the male-sterility of *MS4* is caused by nucleotide substitution of the SNP (AX-174121522), and further experiments are necessary to clarify whether this SNP mutation is responsible for male-sterility in *MS4*. For AX-17412746 (the closest marker to *MS1* in this study), the rate of prediction was not 1.00. Mishima et al. [23] reported two markers closely linked to the *MS1* locus (reCj19250_1927 and reCj19250_2335) in the sosyun family. The genotypes of these markers (G/G in reCj19250_1927 and C/C in reCj19250_2335) were not consistent with the expected genotype from the phenotype (G/A in reCj19250_1927 and C/T in reCj19250_2335) in the Ooi-7 used in this study. Therefore, these three markers appeared to be separated by some distance from the *MS1* gene. We are currently generating the draft genome sequence of *C. japonica*, and cumulative information for the high-quality draft genome sequence and closely linked markers to male-sterile genes might facilitate the isolation of male-sterile genes in *C. japonica*.

Supporting information

S1 Fig. Linkage map for *Cryptomeria japonica* derived from the F1O7 family (LG1–LG11). Marker names are indicated to the right of the linkage groups. Centimorgan distances are indicated to the left of each linkage group.
(PDF)

S2 Fig. Linkage map for *Cryptomeria japonica* derived from the S1-2 family (LG1–LG11). Marker names are indicated to the right of the linkage groups. Centimorgan distances are indicated to the left of each linkage group.
(PDF)

S3 Fig. Linkage map for *Cryptomeria japonica* derived from the S5HK7 family (LG1–LG11). Marker names are indicated to the right of the linkage groups. Centimorgan distances are indicated to the left of each linkage group.
(PDF)

S4 Fig. Linkage map for *Cryptomeria japonica* derived from the S8HK5 family (LG1–LG11). Marker names are indicated to the right of the linkage groups. Centimorgan distances are indicated to the left of each linkage group.
(PDF)

S5 Fig. Genotypes of the male-sterile gene (*MS1*) and markers obtained using Affymetrix microarray. Genotypes with red backgrounds showed inconsistencies between the marker genotypes and genotype for the *MS1* locus.
(PDF)

S6 Fig. Genotypes of the male-sterile gene (*MS4*) and markers obtained using Affymetrix microarray. Genotypes with red backgrounds showed inconsistencies between the marker genotypes and genotype for the *MS4* locus.
(PDF)

S1 Table. Summary of the *Cryptomeria japonica* trees used to calculate the prediction rate of male-sterile gene genotypes from the marker genotypes.

(XLSX)

S2 Table. Positions of mapped markers and the MS1 gene and these genotypes of the F107 family. Individual name with ms indicates the male sterility.

(XLSX)

S3 Table. Positions of mapped markers and the MS2 gene and these genotypes of the S1-2 family. Individual name with ms indicates the male sterility.

(XLSX)

S4 Table. Positions of mapped markers and the MS3 gene and these genotypes of the S5HK7 family. Individual name with ms indicates the male sterility.

(XLSX)

S5 Table. Positions of mapped markers and the MS4 gene and these genotypes of the S8HK5 family. Individual name with ms indicates the male sterility.

(XLSX)

S6 Table. Prediction rate of male-sterile gene genotypes from the marker genotypes of 78 *Cryptomeria japonica* trees. The 78 trees included eight parents of the four mapping populations, four male-sterile trees, seven trees heterozygous for male-sterile gene, two parents of the previous linkage maps, seven trees from breeding materials, and 50 trees from natural forests (see [S1 Table](#)).

(XLSX)

Acknowledgments

The authors would like to thank M. Koshiba and Y. Komatsu for their assistance with the laboratory work and N. Futamura for the donation of 16 male sterility candidate gene homologues. We also thank Ishikawa, Kanagawa, and Fukushima Prefectural Forest Experiment Stations for providing breeding materials for this study.

Author Contributions

Conceptualization: Saneyoshi Ueno, Asako Matsumoto, Tokuko Ujino-Ihara, Kentaro Uchiyama, Yoshinari Moriguchi.

Data curation: Yoichi Hasegawa, Saneyoshi Ueno, Tokuko Ujino-Ihara, Yoshinari Moriguchi.

Formal analysis: Yoichi Hasegawa, Saneyoshi Ueno, Tokuko Ujino-Ihara, Yoshinari Moriguchi.

Funding acquisition: Yoshinari Moriguchi.

Investigation: Saneyoshi Ueno, Asako Matsumoto, Satoko Totsuka, Junji Iwai, Tetsuji Hakamata, Yoshinari Moriguchi.

Methodology: Yoichi Hasegawa, Saneyoshi Ueno, Tokuko Ujino-Ihara, Yoshinari Moriguchi.

Project administration: Yoshinari Moriguchi.

Resources: Saneyoshi Ueno, Asako Matsumoto, Tokuko Ujino-Ihara, Kentaro Uchiyama, Satoko Totsuka, Junji Iwai, Tetsuji Hakamata, Yoshinari Moriguchi.

Supervision: Yoshinari Moriguchi.

Writing – original draft: Yoichi Hasegawa, Saneyoshi Ueno, Yoshinari Moriguchi.

Writing – review & editing: Yoichi Hasegawa, Saneyoshi Ueno, Tokuko Ujino-Ihara, Yoshinari Moriguchi.

References

1. Zhong S, Toubia-Rahme H, Steffenson BJ, Smith KP. Molecular mapping and marker-assisted selection of genes for septoria speckled leaf blotch resistance in barley. *Phytopathology*. 2006; 96: 993–999. <https://doi.org/10.1094/PHYTO-96-0993> PMID: 18944055
2. Karaagac E, Vargas AM, de Andrés MT, Carreño I, Ibáñez J, Carreño J et al. Marker assisted selection for seedlessness in table grape breeding. *Tree Genet Genomes*. 2012; 8: 1003–1015.
3. Neale DB, Savolainen O. Association genetics of complex traits in conifers. *Trends Plant Sci*. 2004; 9: 325–330. <https://doi.org/10.1016/j.tplants.2004.05.006> PMID: 15231277
4. Hizume M, Shibata F, Matsusaki Y, Garajova Z. Chromosome identification and comparative karyotypic analyses of four *Pinus* species. *Theor Appl Genet*. 2001; 105: 491–497.
5. Agency Forestry. Statistical handbook of forest and forestry. Forestry Agency, Ministry of Agriculture, Forestry and Fisheries; 2014. pp. 8–9. (in Japanese)
6. Baba K, Nakae K. The national epidemiological survey of allergic rhinitis in 2008—comparison between 1998 and 2008. *Prog Med*. 2008; 28: 2001–2012. (in Japanese)
7. Taira H, Saito M, Furuta Y. Inheritance of the trait of male sterility in *Cryptomeria japonica*. *J For Res*. 1999; 4: 271–273.
8. Saito M. Breeding strategy for the pollinosis preventive cultivars of *Cryptomeria japonica* D Don. *J Jpn For Soc*. 2010; 92: 316–323. (in Japanese with English abstract)
9. Taira H, Teranishi H, Kenda Y. A case study of male sterility in sugi (*Cryptomeria japonica*). *J Jpn For Soc*. 1993; 75: 377–379. (in Japanese with English abstract)
10. Igarashi M, Watanabe J, Ozawa H, Saito Y, Taira H. The male sterile sugi (*Cryptomeria japonica* D. Don) was found in Fukushima Prefecture (I): Selection of the search ground and identification of the male sterility. *Tohoku J For Sci*. 2004; 9: 86–89. (in Japanese with English abstract)
11. Igarashi M, Watanabe J, Saito Y, Ozawa H, Saito N, Furukawa S et al. Breeding of *Cryptomeria japonica* for the countermeasure of pollinosis. *Bull Fukushima Prefect For Res Centre*. 2006; 39: 1–9. (in Japanese)
12. Yoshii E, Taira H. Cytological and genetical studies on male sterile sugi (*Cryptomeria japonica* D. Don), Shindai 1 and Shindai 5. *J Jpn For Soc*. 2007; 89: 26–30. (in Japanese with English abstract)
13. Miyajima D, Yoshii E, Hosoo Y, Taira H. Cytological and genetic studies on male sterility in *Cryptomeria japonica* D. Don (Shindai 8). *J Jpn For Soc*. 2010; 92: 106–109. (in Japanese with English abstract)
14. Tsumura Y, Suyama Y, Yoshimura K, Shirato N, Mukai Y. Sequence-tagged-sites (STSs) of cDNA clones in *Cryptomeria japonica* and their evaluation as molecular markers in conifers. *Theor Appl Genet*. 1997; 94: 764–772.
15. Nikaido AM, Ujino T, Iwata H, Yoshimura K, Yoshimura H, Suyama Y et al. AFLP and CAPS linkage maps of *Cryptomeria japonica*. *Theor Appl Genet*. 2000; 100: 825–831.
16. Iwata H, Ujino-Ihara T, Yoshimura K, Nagasaka K, Mukai Y, Tsumura Y. Cleaved amplified polymorphic sequence markers in sugi, *Cryptomeria japonica* D. Don, and their locations on a linkage map. *Theor Appl Genet*. 2001; 103: 881–895.
17. Moriguchi Y, Iwata H, Ihara T, Yoshimura K, Taira H, Tsumura Y. Development and characterization of microsatellite markers for *Cryptomeria japonica* D. Don. *Theor Appl Genet*. 2003; 106: 751–758. <https://doi.org/10.1007/s00122-002-1149-0> PMID: 12596006
18. Moriguchi Y, Ueno S, Ujino-Ihara T, Futamura N, Matsumoto A, Shinohara K et al. Characterization of EST-SSRs from *Cryptomeria japonica*. *Conserv Genet Resour*. 2009; 1: 373–376.
19. Moriguchi Y, Ueno S, Higuchi Y, Miyajima D, Ito S, Futamura N et al. Establishment of a microsatellite panel covering the sugi (*Cryptomeria japonica*) genome, and its application for localization of a male sterile gene (ms-2). *Mol Breed*. 2014; 33: 315–325.
20. Tani N, Takahashi T, Ujino-Ihara T, Iwata H, Yoshimura K, Tsumura Y. Development and characteristics of microsatellite markers for sugi (*Cryptomeria japonica* D. Don) from microsatellite enriched libraries. *Ann For Sci*. 2004; 61: 569–575.

21. Ueno S, Moriguchi Y, Uchiyama K, Ujino-Ihara T, Futamura N, Sakurai T et al. A second generation framework for the analysis of microsatellites in expressed sequence tags and the development of EST-SSR markers for a conifer, *Cryptomeria japonica*. BMC Genomics. 2012; 13: 136. <https://doi.org/10.1186/1471-2164-13-136> PMID: 22507374
22. Uchiyama K, Ujino-Ihara T, Ueno S, Taguchi Y, Futamura N, Shinohara K et al. Single nucleotide polymorphisms in *Cryptomeria japonica*: their discovery and validation for genome mapping and diversity studies. Tree Genet Genomes. 2012; 8: 1213–1222.
23. Mishima K, Hirao T, Tsubomura M, Tamura M, Kurita M, Nose M et al. Identification of novel putative causative genes and genetic marker for male sterility in Japanese cedar (*Cryptomeria japonica* D. Don). BMC Genomics. 2018; 19: 277. <https://doi.org/10.1186/s12864-018-4581-5> PMID: 29685102
24. Moriguchi Y, Ujino-Ihara T, Uchiyama K, Futamura N, Saito M, Ueno S et al. The construction of a high-density linkage map for identifying SNP markers that are tightly linked to a nuclear-recessive major gene for male sterility in *Cryptomeria japonica* D. Don. BMC Genomics. 2012; 13: 95. <https://doi.org/10.1186/1471-2164-13-95> PMID: 22424262
25. Moriguchi Y, Uchiyama K, Ueno S, Ujino-Ihara T, Matsumoto A, Iwai J et al. A high-density linkage map with 2,560 markers and its application for the localization of the male-sterile genes ms3 and ms4 in *Cryptomeria japonica* D. Don. Tree Genet Genomes. 2016; 12: 57.
26. Ueno S, Uchiyama K, Moriguchi Y, Ujino-Ihara T, Matsumoto A, Wei FJ et al. Scanning RNA-Seq and RAD-Seq approach to develop SNP markers closely linked to *MALE STERILITY 1 (MS1)* in *Cryptomeria japonica* D. Don. Breed Sci. Accepted.
27. Futamura N, Totoki Y, Toyoda A, Igasaki T, Nanjo T, Seki M et al. Characterization of expressed sequence tags from a full-length enriched cDNA library of *Cryptomeria japonica* male strobili. BMC Genomics. 2008; 9: 383. <https://doi.org/10.1186/1471-2164-9-383> PMID: 18691438
28. Gilbert D. Gene-omes built from mRNA seq not genome DNA. In: 7th Annual Arthropod Genomics Symposium, Notre Dame. 2013.
29. Uchiyama K, Iwata H, Moriguchi Y, Ujino-Ihara T, Ueno S, Taguchi Y et al. Demonstration of genome-wide association studies for identifying markers for wood property and male strobili traits in *Cryptomeria japonica*. PLoS ONE. 2013; 8: e79866. <https://doi.org/10.1371/journal.pone.0079866> PMID: 24260312
30. Saito M, Aiura H, Kato A, Matsuura T. Characterization of the cutting cultivar of *Cryptomeria japonica* D. Don, Zasshunbo, with a heterozygous male-sterility gene in a heterozygous state, selected from Toyama Prefecture, Japan. Forest Genetics and Tree Breeding. 2015; 4: 45–51. (in Japanese with English abstract)
31. Le Provost G, Herrera R, Paiva JA, Chaumeil P, Salin F, Plomion C. A micromethod for high throughput RNA extraction in forest trees. Biol Res. 2007; 40: 291–297. /S0716-97602007000400003 PMID: 18449457
32. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
33. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv. 2013; 1303.3997v2.
34. Murray M, Thompson WF. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 1980; 8: 4321–4325. PMID: 7433111
35. Tani N, Takahashi T, Iwata H, Mukai Y, Ujino-Ihara T, Matsumoto A et al. A consensus linkage map for sugi (*Cryptomeria japonica*) from two pedigrees, based on microsatellites and expressed sequence tags. Genetics. 2003; 165: 1551–1568. PMID: 14668402
36. Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 2002; 93: 77–78. PMID: 12011185
37. Neves LG, Davis JM, Barbazuk WB, Kirst M. A high-density gene map of loblolly pine (*Pinus taeda* L.) based on exome sequence capture genotyping. G3-Genes Genomes Genet. 2014; 4: 29–37.
38. Westbrook JW, Chhatre VE, Wu LS, Chamala S, Neves LG, Muñoz P et al. A consensus genetic map for *Pinus taeda* and *Pinus elliottii* and extent of linkage disequilibrium in two genotype-phenotype discovery populations of *Pinus taeda*. G3-Genes Genomes Genet. 2015; 5: 1685–1694.
39. Plomion C, Chancerel E, Endelman J, Lamy JB, Mandrou E, Lesur I et al. Genome-wide distribution of genetic diversity and linkage disequilibrium in a mass-selected population of maritime pine. BMC Genomics. 2014; 15: 171. <https://doi.org/10.1186/1471-2164-15-171> PMID: 24581176
40. Pavy N, Pelgas B, Laroche J, Rigault P, Isabel N, Bousquet J. A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. BMC Biol. 2012; 10: 84. <https://doi.org/10.1186/1741-7007-10-84> PMID: 23102090

LG1

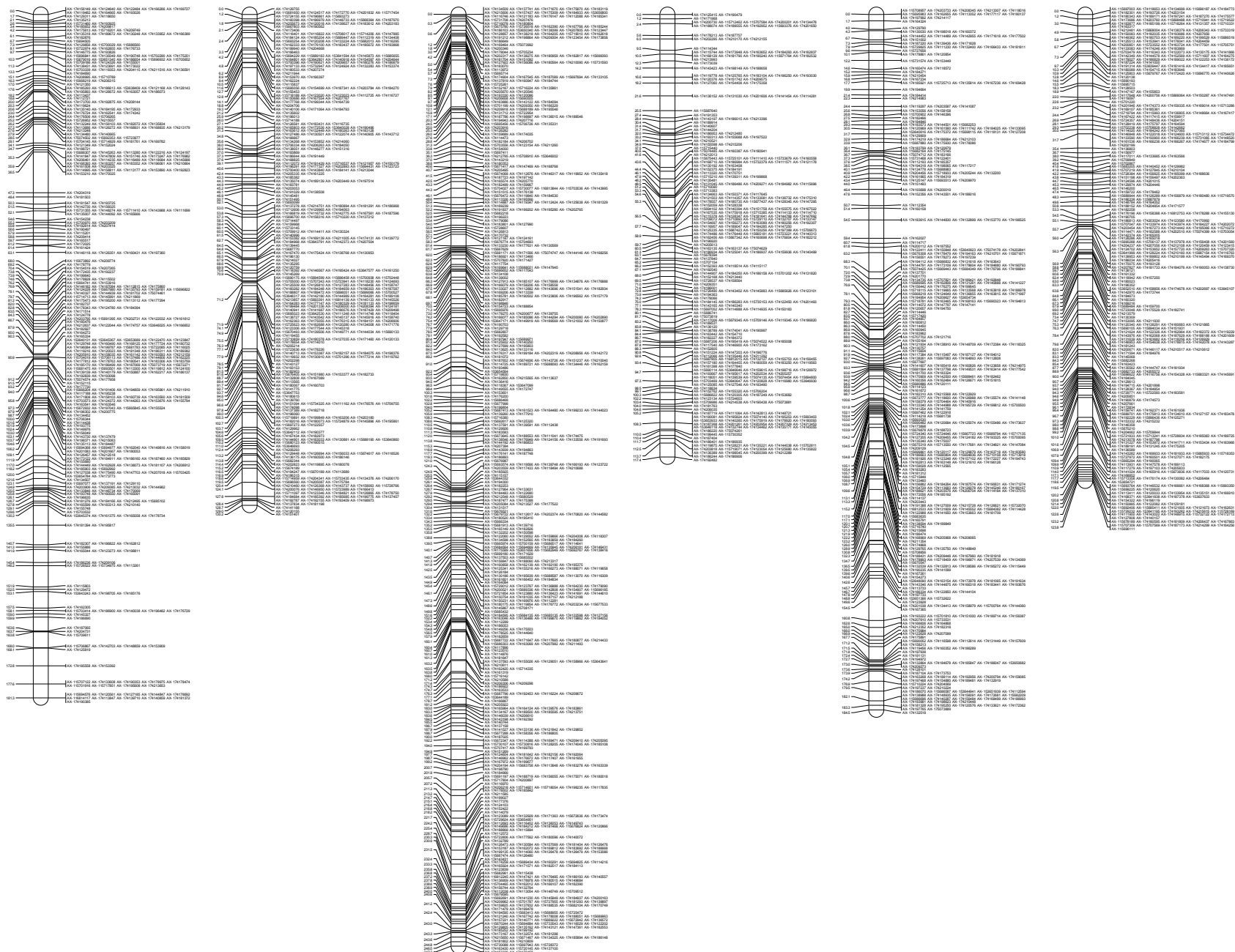
LG2

LG3

LG4

LG5

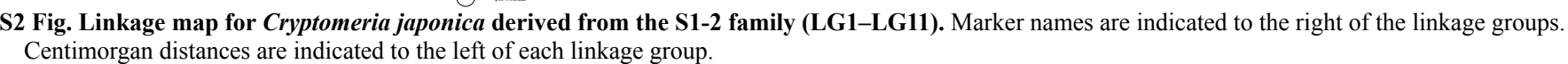
LG6



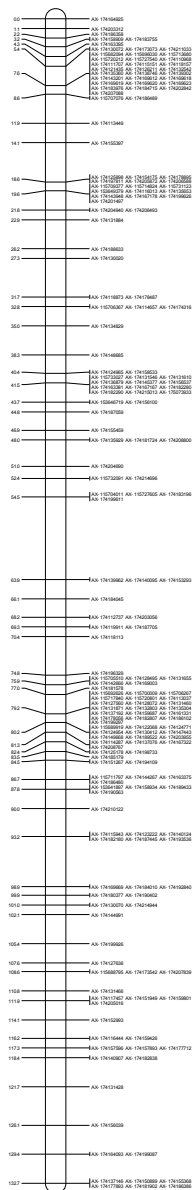
S1 Fig. Linkage map for *Cryptomeria japonica* derived from the F107 family (LG1–LG11). Marker names are indicated to the right of the linkage groups. Centimorgan distances are indicated to the left of each linkage group.

LG11

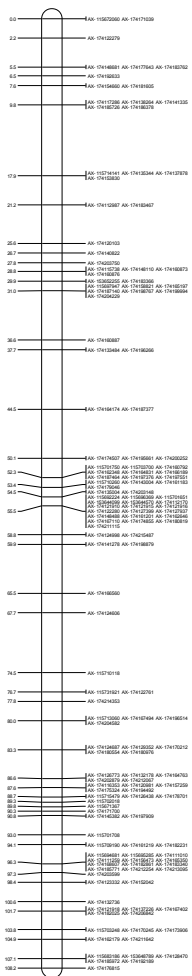


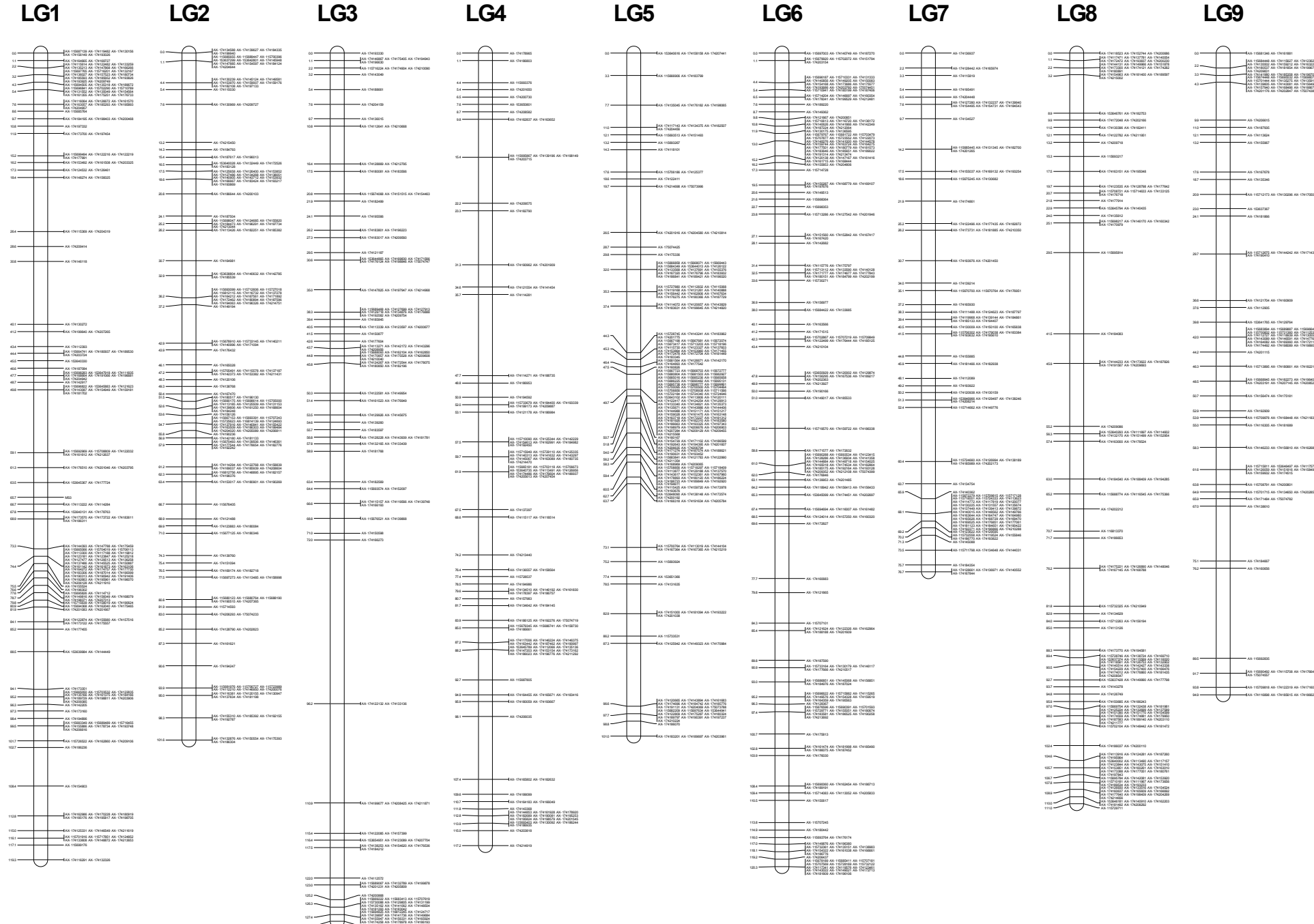


LG10



LG11

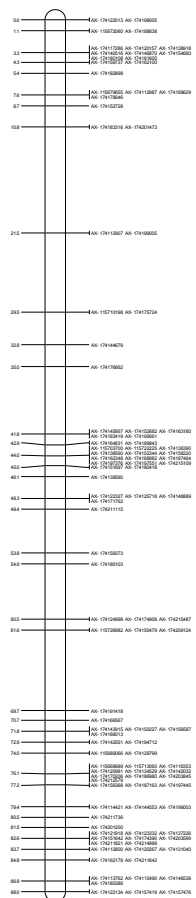
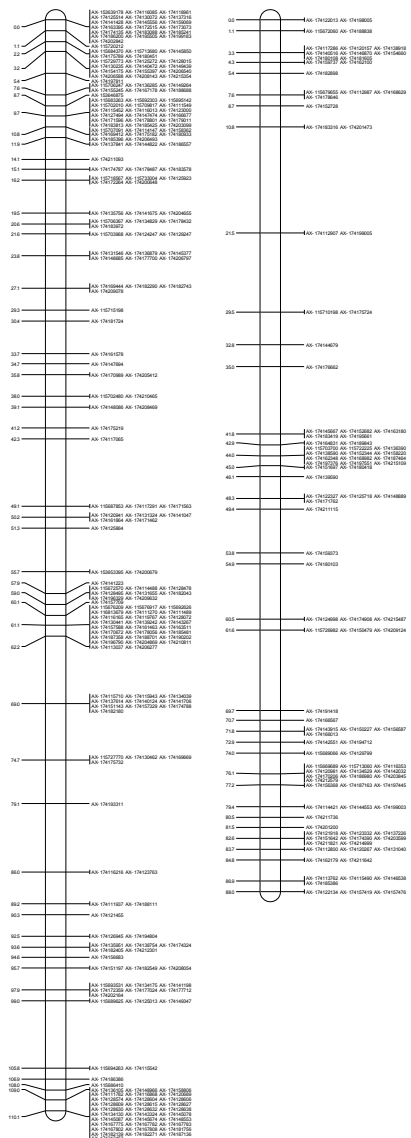


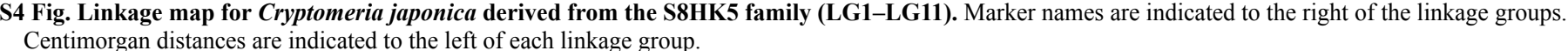


S3 Fig. Linkage map for *Cryptomeria japonica* derived from the S5HK7 family (LG1–LG11). Marker names are indicated to the right of the linkage groups. Centimorgan distances are indicated to the left of each linkage group.

LG10

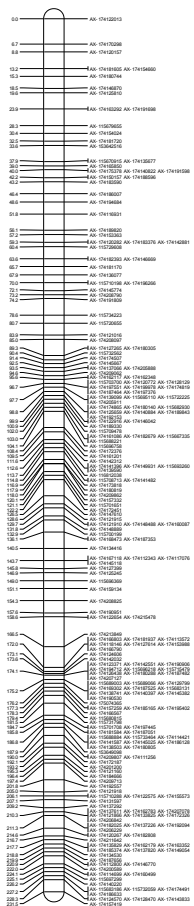
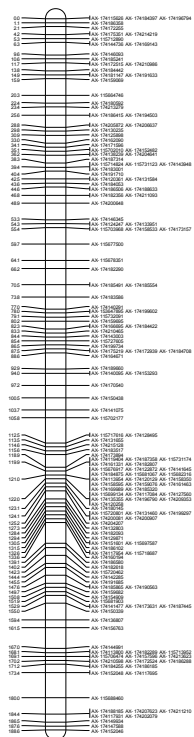
LG11





LG10

LG11



MSI

S5 Fig. Genotypes of the male-sterile gene (*MS1*) and markers obtained using Affymetrix microarray. Genotypes with red backgrounds showed inconsistencies between the marker genotypes and genotype for the *MS1* locus.

S6 Fig. Genotypes of the male-sterile gene (*MS4*) and markers obtained using Affymetrix microarray. Genotypes with red backgrounds showed inconsistencies between the marker genotypes and genotype for the *MS4* locus.