Genetic Distinctiveness but Low Diversity Characterizes Rear-Edge *Thuja standishii* (Gordon) Carr. (Cupressaceae) Populations in Southwest Japan

James R. P. Worth [1,*], Ichiro Tamaki [2], Ikutaro Tsuyama [3], Peter A. Harrison [4], Kyoko Sugai [5], Hitoshi Sakio [6], Mineaki Aizawa [7] and Satoshi Kikuchi [3]

1   Department of Forest Molecular Genetics and Biotechnology, Forestry and Forest Products Research Institute, Tsukuba 305-8687, Japan
2   Gifu Academy of Forest Science and Culture, Sodai, Mino, Gifu 501-3714, Japan; garageit@gmail.com
3   Hokkaido Research Center, Forestry and Forest Products Research Institute, 7 Hitsujigaoka, Toyohira, Sapporo 062-8516, Japan; aquarianleo@gmail.com (I.T.); kikuchi.satoshi.ffpri@gmail.com (S.K.)
4   School of Natural Sciences, University of Tasmania, Private Bag 55, Hobart, TAS 7001, Australia; P.A.Harrison@utas.edu.au
5   Institute of Agricultural and Life Sciences, Academic Assembly, Shimane University, Matsue 690-8504, Japan; sugaikyoko@life.shimane-u.ac.jp
6   Sado Island Center for Ecological Sustainability, Niigata University, Sado, Niigata 952-2206, Japan; sakiohit@gmail.com
7   Department of Forest Science, School of Agriculture, Utsunomiya University, Utsunomiya 321-8505, Japan; aizawam@cc.utsunomiya-u.ac.jp
*   Correspondence: jrpw2326@affrc.go.jp

Abstract: Rear-edge populations are of significant scientific interest because they can contain allelic variation not found in core-range populations. However, such populations can differ in their level of genetic diversity and divergence reflecting variation in life-history traits, demographic histories and human impacts. Using 13 EST-microsatellites, we investigated the genetic diversity and differentiation of rear-edge populations of the Japanese endemic conifer *Thuja standishii* (Gordon) Carr. in southwest Japan from the core-range in northeast Japan. Range-wide genetic differentiation was moderate ($Fst = 0.087$), with northeast populations weakly differentiated ($Fst = 0.047$), but harboring high genetic diversity (average population-level $Ar = 4.76$ and $Ho = 0.59$). In contrast, rear-edge populations were genetically diverged ($Fst = 0.168$), but contained few unique alleles with lower genetic diversity ($Ar = 3.73$, $Ho = 0.49$). The divergence between rear-edge populations exceeding levels observed in the core-range and results from ABC analysis and species distribution modelling suggest that these populations are most likely relicts of the Last Glacial Maximum. However, despite long term persistence, low effective population size, low migration between populations and genetic drift have worked to promote the genetic differentiation of southwest Japan populations of *T. standishii* without the accumulation of unique alleles.

Keywords: approximate Bayesian computation; central-marginal hypothesis; gene-flow; genetic drift; Japanese endemic conifer; leading edge; relictual population

1. Introduction

Many plant species have distributions fitting a central-marginal model characterized by a core range surrounded by smaller edge populations. The core range consists of more or less contiguous large populations and is thought to reflect optimal conditions leading to high effective population size (Ne) and genetic diversity [1]. In contrast, at the margins of species' ranges, populations usually become smaller and increasingly geographically isolated confined to narrow areas of sub-optimal microclimates within or even outside the climatic envelope of the species core range [2]. These populations are usually characterized

by high genetic divergence and low effective population sizes [3]. Rear-edge populations, that is, those that are distributed at the low-latitude margin of species' ranges [4], are often considered to be disproportionately important in terms of their contribution to species overall genetic diversity and ecological and evolutionary processes [5,6]. This is because rear-edge populations have often experienced long-term isolation from the core range promoting genetic drift [7] and local adaptation [4]. For these reasons and the fact that rear-edge populations of keystone species like large trees can form crucial "islands" of habitat in an otherwise inhospitable climate for various other organisms [8], such populations are of immense scientific and conservation interest.

However, due to a range of factors including life history traits, distinct biogeographic and demographic histories and human impacts [9], rear-edge populations can differ widely in their levels of genetic diversity and divergence compared to the core range [10]. At one extreme are rear-edge populations that have persisted in geographically isolated stable climate refugia (i.e., regions that have provided suitable conditions for species persistence through both glacial and interglacial periods [11]) characterized by high genetic diversity and strong differentiation [4]. Even in plants with high gene flow, such refugia can harbor unique genetic diversity, thus making them important genetic reservoirs for species [12]. Some examples include populations of widespread temperate trees such as *Fraxinus excelsior* L. in Iran [12], *Cryptomoria japonica* (L.f) D.Don on Yaku Island, southern Japan [13] and *Pseudotsuga menziesii* (Mirbel) Franco in Mexico [14]. Long-term isolation of rear-edge populations can also result in ecological and trait differentiation from the core-range such as the occurrence of polyploidy [6], modification of dispersal traits [15] and ecophysiological changes such as increased drought tolerance [16]. In contrast, rear-edge populations that have narrow ranges, have been demographically unstable and/or have only recently been isolated from core range populations can display low genetic diversity. Some examples include populations of *Ulmus minor* Mill. in the Balearic Islands [17] and *Myrtus nivellei* Batt. and Trab., 1972 in the Sahara desert [18]. Understanding the biogeographic history and genetic diversity of rear-edge populations is therefore vital for evaluating their relationship to the core-range and whether they require specific conservation actions.

The cool temperate and subalpine flora of Japan is characterized by species with central-marginal like distributions. Their core range occurs in the northeast of Japan while small, isolated rear-edge populations are disjunctly distributed in the highest mountains of southwest Japan [19]. This study investigates the genetic diversity of rear-edge populations of the Japanese endemic conifer *Thuja standishii* (Gordon) Carr. (Cupressaceae) in southwest Japan and their relationships to large populations in its core range in northeast Japan (Figure 1). *Thuja standishii* is a widely distributed tree (40.67° N in northern Honshu to 33.49° N in Shikoku) that usually co-occurs with other conifers in cool temperate or subalpine forests, but sometimes dominates in steep, rocky or waterlogged habitats [20]. While a relatively common component of forests in northeast Japan, particularly on the Japan Sea side of Honshu [21], the species has a limited distribution in southwest Japan with only six known isolated occurrences. These populations are of great interest because of their geographic isolation from the core-range (varying from 285 km to 430 km) across the most important phylogeographic break in Japan located in the Kansai region [22–25] (Figure 1). In addition, their occurrence at the edge of the species climatic range encompasses the lowest elevation sites known in the species including growing as emergents in warm temperate evergreen forests [20].
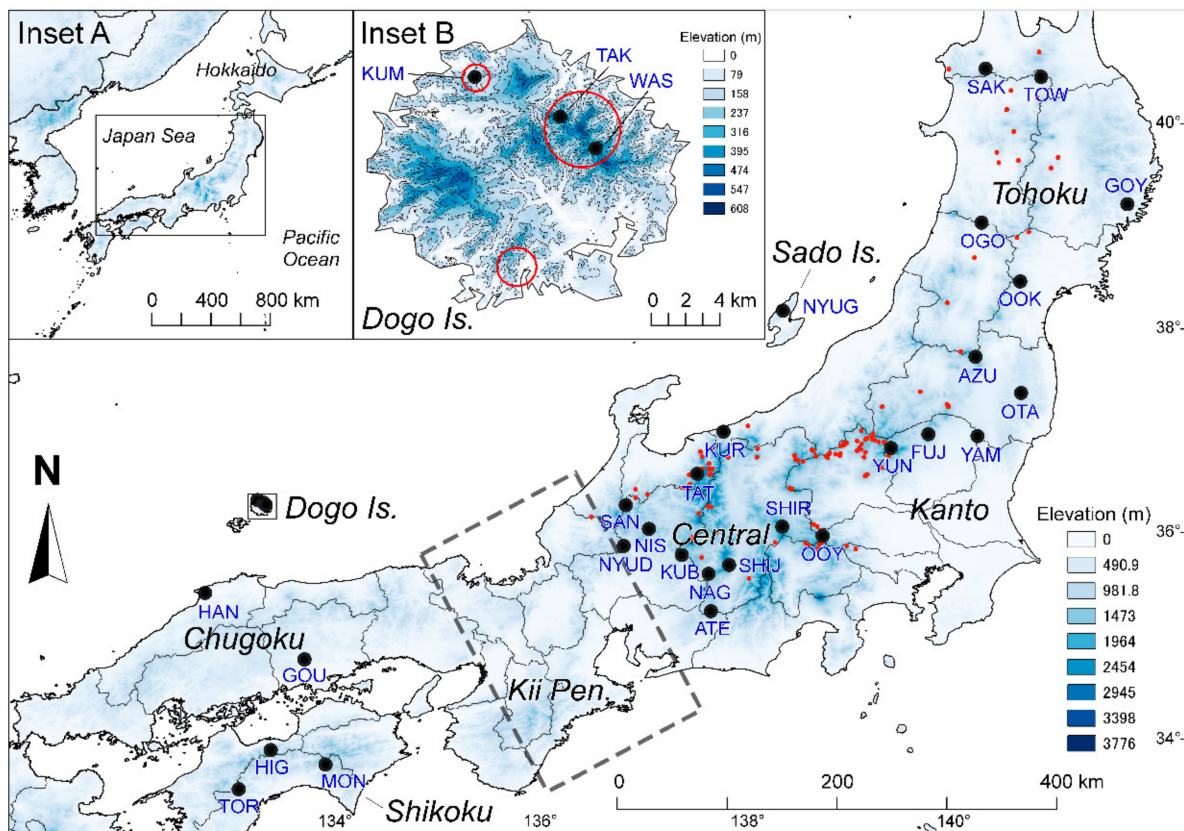
Figure 1. Location of the 30 sampled populations of *Thuja standishii* indicated as black circles. The current distribution of the species is shown as small red circles. For full names of populations see Table S1 in the Supplementary Materials. The area surrounded by dashed line indicates the broad area where the major phylogeographic break in the Japanese flora between northeast and southwest Japan is situated [22–25]. Inset A shows the location of the study area in relation to the rest of Japan and the Eurasian mainland while inset B shows the location of the three populations sampled on Dogo Island with 100 m contour intervals shown as grey lines. The parts of the island where *Thuja standishii* is distributed are enclosed in red circles.

The origin and genetic diversity of these rear-edge populations remains unknown. They may be the direct descendants of populations that persisted in stable climate refugia for multiple glacial–interglacial cycles in southwest Japan and contain unique genetic diversity. In fact, these populations are closest to large areas of suitable habitat for temperate forest tree species predicted to have occurred in southwest Japan during the Last Glacial Maximum [26,27]. Alternatively, southwest populations may have high genetic divergence but low genetic diversity as a result of genetic bottlenecking in small glacial refugia or even an origin via post-glacial dispersal events from the core-range. Lastly, these populations may be remnants of a more extensive past distribution of *T. standishii* in southwest Japan that has been mostly lost due to forest destruction by humans [20]. However, discerning between these scenarios is near impossible based on the fossil record because the pollen of *T. standishii* is very difficult to distinguish from other Cupressoideae species, which in Japan includes widespread species of the genera *Chamaecyparis*, *Juniperus* and *Thujopsis* [28]. In addition, macrofossils of *T. standishii* are extremely rare, with the youngest known from the Last Glacial Maximum of lowland northeast Japan from only two sites [29,30].

This study uses expressed sequence tagged (EST) microsatellites to investigate the range-wide genetic structure and diversity of *T. standishii*, including all known occurrences in southwest Japan. Specifically, we aim to (1) understand the genetic affinity of eight populations sampled from the six occurrences in southwest Japan to the core range in northeast Japan; (2) compare the genetic diversity of core range and rear-edge populations and (3) examine past demographic and distributional changes across the species range using approximate Bayesian computation (ABC) and species distribution modelling.

2. Materials and Methods

2.1. Population Sampling and Identification of Clones

A total of 30 populations (850 samples) of *T. standishii* were collected across the species range (Figure 1). In order to estimate population level genetic diversity, a minimum of 20–25 samples were collected where possible at least two plant heights apart. Twenty-two populations were sampled in northeast Japan, including eight from Tohoku (including the most northern populations), four from Kanto and ten from Japan Sea Side/central Honshu. In southwest Japan, all known populations were sampled in Shikoku (three populations) and mainland Chugoku (two populations)), while three were sampled from Dogo Island, the main island of the Oki island group, representing most of the species' distribution there (Figure 1). As populations differed in spatial extent, with most populations in southwest Japan and some in northeast Japan (e.g., NYUG, KUB and KUR) being small (Figure 1; for full names of populations see Table S1), we examined whether average spatial sampling distance between individuals in each population (calculated in Quantum GIS 3.01 using nearest neighbour analysis) may be correlated with population-level genetic diversity. All populations were included in this analysis except the small population on Sado Island (NYUG) where individual GPS points were not taken. Overall, there were no significant relationships between sampling distance and genetic diversity when examining all populations or within northeast or southwest Japan (Figure S1).

Individuals that shared the same genotype (i.e., almost certainly clones) were identified in GenAlex 6.5 [31] using the "find clones" function. Identical genotypes were identified in eight populations with six found at GOU, three at NYUG and KUR, two at SHIR and one each at the remaining four populations (Table S2). One individual from each clone was retained resulting in 33 samples being excluded from all analyses leaving 817 samples.

2.2. Laboratory Work and Genotyping

DNA extractions were undertaken using a modified CTAB protocol [32]. Genetic diversity and structure of T. standishii was assessed using 13 of the 15 Expressed Sequence Tagged (EST) microsatellite loci (Kurobe_18480, Kurobe_2969, Kurobe_23700, Kurobe_44557, Kurobe_15129, Kurobe_16758, Kurobe_6943, Kurobe_23263, Kurobe_38308, Kurobe_41636, Kurobe_42400, Kurobe_40825 and Kurobe_4219) developed by Worth et al. [33] and amplified following the same protocol. Genotyping of alleles was done in GeneMarker (SoftGenetics, LLC, PA, USA).

2.3. Data Integrity

For all 13 loci, linkage disequilibrium (i.e., allelic correlations among loci) were tested at the population level in Genepop v 4.2 [34] with sequential Bonferroni correction ($p < 0.05$). In addition, whether loci were affected by null alleles was assessed using the expectation-minimization (EM) algorithm [35] in FreeNA [36]. To examine how null alleles may have impacted estimates of population genetic differentiation, in the same program, *Fst* with and without null alleles, were also calculated.

EST microsatellite loci, which are located in genes, may be subject to directional (increasing allele fixation) or balancing (maintaining allele frequencies) selection [37]. However, the possibility that selection could influence genetic parameters of EST-SSR loci is low given that most genes do not experience positive selection [38] and EST-SSR loci have been found to display highly similar genetic differentiation to genomic SSRs [39–41]. Nonetheless, whether any loci were potentially under selection was tested using an *Fst* outlier method implemented in BayeScan v2.1 [42]. Default parameters were used except a more realistic prior odds of 1000 was used following Lotterhos and Whitlock [43]. Each run was repeated three times and assessed under Jeffreys' scale of evidence [44]. Existing genetic structuring can increase the number of false positives when detecting outlier loci [45]; therefore, the analysis was undertaken using genetic clusters identified using Bayesian analysis of population structure (BAPS) version 6 [46] run using a maximum *K* of

20. BAPS identified 11 clusters with probability of $K = 11$ being 0.968 versus the next best, $K = 10$, with a probability of 0.032.

2.4. Isolation by Distance and Environment

We investigated the importance of two major factors that may contribute to the genetic divergence of populations of *T. standishii*; namely isolation by distance (IBD: [47]), whereby geographic distance and landscape barriers cause restricted gene flow [48], and isolation by environment (IBE: [49]), in which gene flow is inhibited between populations occupying different environments [48]. Despite expectations that neutral markers unlinked to selected loci should undergo extensive gene flow between populations even in different environments [50], selection may work against the whole genome of migrants and hybrids reducing gene flow even at unliked neutral markers [50]. To examine the strength of evidence for IBD and IBE driving population differentiation, we fitted a generalized dissimilarity model (GDM), associating a matrix of genetic distances among populations defined by $Fst/(1—Fst)$ with geographic and climatic gradients using non-linear matrix regression [51,52]. Climatic gradients were expressed as principal components (PC) from a principal components analysis (PCA) fitted using *prcomp* in R version 4.0.3 [53] that summarised the climatic envelope of *T. standishii* into three orthogonal components, using the same climate variables utilized for the species distribution modelling (see Section 2.8 below). The GDM was fitted using the "gdm" package in R version 4.0.3 [53] and variable importance and its significance were estimated from 1000 permutations using the *gdm.varImp* function. The importance and magnitude of genetic change along the geographic and climatic gradients were visualized using the extracted I-splines from 1000 cross-validated GDMs, where each I-spline reflects the partial dependency of genetic change for a given gradient whilst keeping all other gradients constant. This was done over all 30 populations and separately for northeast and southwest Japan populations.

2.5. Range-Wide Genetic Structure

Genetic structure across the species range was investigated using both region, population and individual-based methods. For the purpose of comparison of genetic diversity, populations were grouped into regions based on a combination of geography and genetic relationships (see Results) and consisted of Shikoku, mainland Chugoku and Dogo Island in southwest Japan and Kanto, Japan Sea side/Central Honshu and Tohoku (including Sado Island) in northeast Japan. Some regions were supported (albeit weakly in some cases) by genetic data (e.g., Kanto, Tohoku, Japan Sea side/Central Honshu and Dogo Island) while others involved grouping of geographically close but genetically distinct populations (i.e., those in Shikoku and mainland Chugoku). Firstly, analysis of molecular variance (AMOVA) was used to assess the proportion of genetic variation distributed between northeast Japan and southwest Japan and all 30 populations in GenAlEx 6.5 with 999 permutations. Secondly, three genetic differentiation indices (Fst, Hedrick's $F'st$ and Josts' D) were calculated in Genodive 3.04 [54]. Thirdly, a neighbor-joining (NJ) tree based on a population-based matrix of the D_A genetic distance [55] was constructed in SplitsTree 4.14.4 [56]. The maximum possible value of D_A is one which occurs when two populations share no alleles at any loci [57]. D_A has been shown to be the most suitable genetic distance measure for obtaining correct tree topology under different microsatellite mutation models and demographic scenarios [58] while NJ trees are highly effective at displaying population structure under a range of evolutionary histories [59]. Lastly, BAPS v6 [46], which implements a Bayesian approach to determine which combination of predetermined sample groups is best supported by the data [60] was used implementing "clustering of groups of individuals" with a maximum K of 20 and ten independent runs. BAPS has been shown to perform well when population differentiation is relatively low (i.e., Fst is between 0.03–0.1) [61], which is usual for wind-pollinated conifers [62].

For individual-based analysis, the discriminant analysis of principal components (DAPC) multivariate method [63] was used. DAPC is designed to both identify and describe

clusters of genetically related individuals and maximizes the differentiation between groups, while minimizing variation within groups [63]. DAPC performs well under a range of dispersal scenarios including when genetic structuring is clinal [63] which can impact the results of other clustering programs via the detection of spurious clusters [64]. We used DAPC for two separate analyses to (1) assess the proportion of successful reassignment of individuals to pre-defined groups (i.e., an *a priori* analysis) using both the 30 sampled populations and BAPS clusters and to (2) identify clusters of genetically related individuals *a posteriori* using sequential *K*-means clustering and the Bayesian information criterion (BIC). DAPC analyses were implemented in R with *a posteriori* analysis performed using the "find.clusters" command with "max.n.clust" = 20 and 50 principal components retained. Cluster plots were drawn in Structure Plot v2 [65].

2.6. Genetic Diversity and Bottlenecks

The number of alleles (*Na*), number of effective alleles (*Ne*), observed (*Ho*) and expected heterozygosity (*He*) at the population and region level was assessed in GenAlEx 6.5. The inbreeding coefficient (*Fis*) was calculated using FSTAT V2.9.3.2 [66] with the significance level of deviation from zero calculated using 390,000 randomizations. Rarefied values of allelic richness (*Ar*) and private allelic richness (*PAr*) were calculated in HpRare 1.1 [67]. Correlations between population-level genetic diversity (*Ar* and *PAr*) with increasing latitude and with increasing "northeastness" (calculated by taking a logarithm of the sum of latitude and longitude) were assessed using the Pearson correlation coefficient. This analysis was done over all populations and separately for the northeast populations.

Recent genetic bottlenecks were tested for using two methods: Firstly, allele frequency distributions of each population were examined for mode shift distortion, which is characteristic of bottlenecked populations [68], in Bottleneck Version 1.2.02 [69]. Secondly, heterozygosity excess, a signature of past reduction in effective population size, was tested using the Wilcoxin signed-rank test under a two-phase model (TPM) of microsatellite evolution [70] in INEST 2.2 [71].

2.7. Past Demographic Change and Estimates of Migration

In order to investigate the population size change history of the species in each region (Shikoku, mainland Chugoku, Dogo Island, Japan Sea side/Central Honshu, Kanto and Tohoku) during and after the Last Glacial Period (11,700–115,000 years ago), we performed approximate Bayesian computation (ABC). ABC allows comparisons between different demographic models and to estimate the values of parameters of the models [72]. At first, alleles represented as fragment lengths of the 13 EST-microsatellite loci were converted into repeat number data by subtracting flanking region length and then dividing by the repeat motif length of each locus. To adjust for small indels, we added ± 1 to the corresponding alleles, by examining a histogram of allele size distribution. If there were putative large indels, such alleles were binned into the nearest high frequency allele. Six summary statistics, including the average and standard deviation of number of alleles, expected heterozygosity and allele size range, were calculated with arlsumstat version 3.5.2 [73].

We constructed three demographic models: a standard neutral model (SNM), population growth model (PGM) and size reduction model (SRM; Figure S2). The SNM assumes that population size has been constant over time and has one structural parameter, current effective population size (N_{CUR}), whose unit is the number of diploid individuals. The PGM assumes that population size has exponentially expanded from an ancestral effective population size (N_{ANC}) to N_{CUR} during T_2 to T_1. Population growth rate (*G*) was calculated as $G = 1/(T_2 - T_1) \log (N_{ANC}/N_{CUR})$. PGM has four structural parameters (N_{CUR}, N_{ANC}, T_1 and T_2). The SRM assumes that population size has declined from N_{ANC} to N_{CUR} at time T_1. SRM has three structural parameters (N_{CUR}, N_{ANC} and T_1). In PGM and SRM, we assumed that the ranges of T_1 and T_2 were from 1 to 10^3 generations ago and the generation time of *T. standishii* was 150 years per generation and, thus, the ranges of T_1 and T_2 in the

unit of years ago ranged from 150 to 150,000 years ago, which included the Last Glacial Period. All prior distributions of structural parameters are summarized in Table S3.

For the microsatellite mutation model, we used a generalized stepwise mutation model (GSM; [74]). GSM has two parameters: mutation rate (μ) and the geometric parameter (P_{GSM}). P_{GSM} ranges from 0 to 1 and represents the proportion of mutations that change allele sizes by more than one step, the value of zero means a strict stepwise mutation model. The value of 1×10^{-4} was used for the mean value of μ among the 13 EST-microsatellite loci [75]. The prior distribution for the value of μ for each locus was randomly drawn from gamma distribution with *shape* and *rate* parameters. The prior distribution of *shape* parameter was drawn from a uniform distribution from 0.5 to 5 and the *rate* parameter was calculated by *shape* / mean value of μ. The prior distribution of the mean value of P_{GSM} among 13 loci were drawn from a uniform distribution from 0 to 1 and each locus value was randomly drawn from a beta distribution with *a* and *b* parameters. The values of *a* and *b* were calculated by $0.5 + 199 \times$ mean value of P_{GSM} and $a \times (1$—mean value of P_{GSM})/mean value of P_{GSM}, respectively, according to Excoffier et al. [76].

Prior distributions were generated by R version 4.0.3 [53]. The three population size change models were each simulated 10,000 times using fastsimcoal2 version 2.6.0.3 [77] and summary statistics were calculated by arlsumstat. These models were compared using the ABC random forest (ABC-RF) approach implemented in the abcrf package version 1.8.1 of R [78]. The number of trees in random forest was set to 1000. The classification error and posterior probability of the best model were calculated. For the best model selected by ABC-RF, 2×10^5 simulations were repeated and summary statistics were calculated. With the 1000 simulations nearest the observed data set, posterior distributions of parameters were estimated by neural network regression implemented in the abc package version 2.1 of R [79,80]. The number of neural networks was set to 20. Logit transformation was used to keep the estimated value within the prior range. Using posterior distributions of the best model, we estimated an effective population size from the lower to the upper limits of T_2 and then calculated the mode and 95% highest posterior density (HPD). The posterior mode was estimated by the density function of R. The 95% HPD was estimated using the coda package version 0.19.4 [81]. By plotting the mode and 95% HPD over time, the trajectory of population size change in each region was obtained [82].

In order to investigate population divergence and migration patterns between the three regions in both northeastern and southwestern Japan, we constructed three population divergence models with different migration patterns (Figure S3). The northeast and southwest populations were analyzed separately because of evidence for their emergence from separate glacial refugia based on species distribution modelling, the fact that the northeast populations form a distinct genetic cluster from the southwest (see Results) and the fossil record indicating long term presence of the species in both regions [83,84]. Isolation with migration models (WMM) assume that the three regions have diverged at time T_3 and exchanged migrants at a rate M_{XY} between regions X and Y during the present to T_3. In WMM1, we assume all migrations between pairs of regions, i.e., an island model, while, in WMM2, we assume only migrations between pairs of adjacent regions, i.e., a stepping-stone model. Isolation without migration model (OMM) assumes no migration between regions. In all three models, population size change patterns in each region from the present to T_2 were determined from the results of population size change analyses (see results) and all population size and time parameters from the present to T_2 were fixed into posterior mode values estimated in the analyses to reduce computational costs. Prior distributions of T_3, ancestral population size before divergence (N_{ANC_ALL}) and M_{XY} were summarized in Table S3. We calculated seven additional summary statistics, overall number of alleles, overall expected heterozygosity, overall allele size range, overall *Fst* and pairwise *Fst* and, thus, used a total of 25 summary statistics for this analysis. Model comparison by ABC-RF were conducted in the same way as the population size change analyses. For the best model selected by ABC-RF, 4×10^5 simulations were repeated and summary statistics were calculated. With 1000 simulations nearest the observed data set,

posterior distributions of parameters were estimated in the same way as the population size change analyses. As migration rate only shows the proportion of migrants per generation in the population and does not directly reflect the intensity of migration, to evaluate the intensity of migration, we calculated number of migrants per generation (*NM*) by multiplying effective population size of source population backward-in-time and M_{XY}. In this analysis, all directions of migration were shown in time-forward, i.e., movements of individuals. As our divergence model assumes population size change after divergence at T_3, we calculated effective population size using posterior mode values obtained in population size change analyses, multiplied by M_{XY} and then averaged and, finally, the average *NM* from the present to T_3 was obtained.

In both population size change and population divergence analyses, to confirm the fit of the model to the observed data, posterior predictive simulation with 1000 randomly drawn posterior samples was conducted with R and fastsimcoal2 and summary statistics were calculated with arlsumstat and compared to the observed data [85].

2.8. Species Distribution Modelling

Species distribution modelling (SDM) was undertaken using the maximum entropy principle algorithm in MaxEnt [86] to investigate potential habitat for *T. standishii* under current climate and during the culmination of Last Glacial climate, the Last Glacial Maximum (LGM, ca. 21 ka). During the LGM, lower precipitation and mean summer temperature depression of between 5 to 9 °C in Japan had a profound impact on the abundance and distribution of plants including conifers [87]. There was a limited expansion of glaciers and permafrost conditions across Japan [88] and woody vegetation dominated by subalpine conifers covered large parts of the glacial landscape [89–92]. MaxEnt can provide robust predictive performance without absence data [93]. Analyses were undertaken using MaxEnt version 3.4.1 (https://biodiversityinformatics.amnh.org/open_source/maxent/ (accessed on 30 October 2020)) using the "dismo" package (http://rspatial.org/sdm/ (accessed on 30 October 2020)) in R version 3.6.2 [53]. Four bioclimatic variables [94] that are important for the distribution of conifers in cool temperate and subalpine zones in Japan [95] were used: mean temperature of warmest quarter (MeTWaQ), precipitation of warmest quarter (PrWaQ), minimum temperature of coldest month (MiTCM) and precipitation of the coldest quarter (PrCQ). The spatial resolution of the climatic variables was 30″ (ca. 1 km in latitude). Presence data for *T. standishii* (971 points; obtained from Worth [20] and new records based on personal observations) were used as response variable for the SDM. Duplicated presence records in each grid cell were removed to decrease the effect of spatial autocorrelation. We performed 10 runs using the cross-validation method. In order to calculate and summarize the accuracy of each model we used three methods: (1) the area under the curve (AUC, [96]); (2) the continuous Boyce index (CBI, [97,98]); (3) the Brier score [99]. We used the "ecospat" package (http://www.unil.ch/ecospat/home/menuguid/ecospat-resources/tools.html (accessed on 24 March 2021) in R to calculate CBI and the "DescTools" package (https://andrisignorell.github.io/DescTools/ (accessed on 15 April 2021) in R to calculate the Brier score.

Fossil data can be used to either calibrate models under past conditions or to assess the accuracy of projected models calibrated under current conditions [100]. Given that only two LGM fossil records of *T. standishii* are known, which cannot provide a complete picture of the species past LGM distribution [101] and is below the minimum sample size required for construction of accurate species distribution models [102,103], the latter approach was taken. Thus, models developed under current conditions were projected onto two global circulation models (GCMs) for the LGM: the Community Climate System Model version 4 (CCSM, [104]) and the Model for Interdisciplinary Research on Climate Earth System Model (MIROC,) [105]), both with a spatial resolution of 2.5′ (ca. 5 km in latitude). We applied two thresholds of occurrence probability based on cut-off sensitivity values [106] of 95% and 99% to define areas of potential habitat. These two thresholds had differing levels of success in predicting the two LGM fossil records (see Results).

3. Results

3.1. Data Integrity

A total of six out of the 2340 population by locus-pairs were found to have significant allelic associations. Of the nine loci within these significant population by locus-pairs, three (Kurobe_18480, Kurobe_4219 and Kurobe_42400) were observed twice while the other five were observed only once. Thus, as such, there was no strong evidence for linkage of any loci, all 13 loci were retained for analyses. Bayescan identified no loci under potential selection with posterior probability values near zero at all loci (Table S4). For each of the 13 loci, the average frequency of estimated null alleles per population averaged 2.3% and *Fst* values estimated with and without null alleles were highly similar (average *Fst* across all 13 loci of 0.085 versus 0.087, respectively) (Table S5). These results indicated that the problem of null alleles was low for the 13 loci and therefore, the unadjusted data set was used for all analyses.

3.2. Isolation by Distance and Environment

Over all 30 sampled populations, there was evidence for significant IBD ($p < 0.0001$) and IBE ($p = 0.004$) (Figure 2), with the model significantly ($p = 0.002$) explaining 71% of the deviance in the genetic distance among the populations. The greatest change in genetic distance was associated with climatic distance (Figure S4), particularly PC2 which reflected a longitudinal cline in the precipitation of the coldest quarter among the populations (Results not shown). Geographic distance showed a monotonic increase with genetic distance (after accounting for climatic distance), with the genetic differentiation increasing when populations were more than 200 km apart. When east and western Japan populations were analysed separately, evidence for IBD and IBE was only significant at the alpha = 0.1 level and explained 38% (northeastern Japan) and 65% (southwestern Japan) of the deviance in genetic differentiation (Figure S5).

3.3. Range-Wide Genetic Structure

Over all 30 populations, genetic differentiation based on *Fst* was 0.087 (with pairwise populations values ranging from 0.0033–0.327)), while Hedrick's *F'st* was 0.218 and Jost's *D* was 0.144. AMOVA showed that 89% of genetic variation was found within populations, 7.48% between populations and 3% between northeast and southwest Japan (Table 1). Genetic differentiation between populations was low in northeast Japan with an average *Fst* of 0.047 (varying between 0.0033–0.132) compared to 0.168 (0.024–0.327) in the southwest. The NJ tree indicated clear genetic divergence of northeast and southwest Japanese populations (Figure 3). Despite low overall genetic differentiation in northeast Japan, geographic based grouping of populations was evident with populations from Tohoku, Kanto and Japan Sea side/central Honshu forming separate clades with the exception of the Tohoku population OTA grouping with Kanto and the geographically isolated YAM (southern Fukushima Prefecture) population and a high elevation population from the Yatsugatake Mountains (SHIR) not grouping with any clade (Figure 3). In contrast, apart from the three closely related populations from Dogo Island, genetic relationships between populations in southwest Japan were not associated with geographic origin: MON (Shikoku) was closest to HAN (mainland Chugoku) and TOR (Shikoku), the most southern population, was closest to the small population GOU (mainland Chugoku). All populations in Shikoku and mainland Chugoku were characterized by high genetic divergence from northeast Japan populations, especially TOR and GOU. BAPS identified eleven clusters with all populations in northeast Japan forming one cluster except for four single population-based clusters: the most northern populations sampled (SAK and TOW), YAM and SHIR (Figure 4). In southwest Japan, all populations were identified as a separate cluster except for the three Dogo Island populations which formed a single cluster.
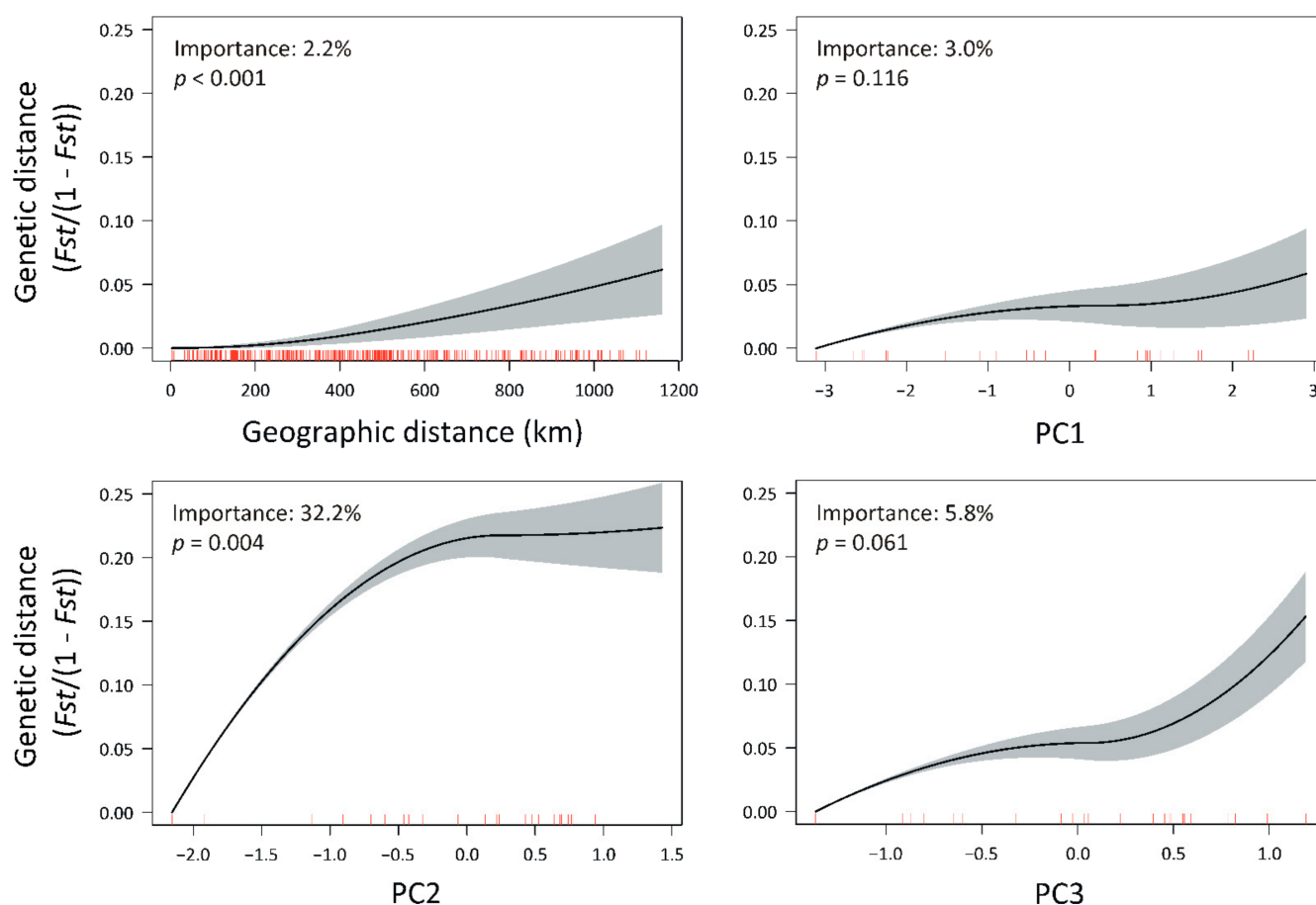
Figure 2. Visualization of the I-splines and their standard errors estimated from 1000 bootstraps of the Generalized Dissimilarity Model for the 30 populations of *T. standishii*. Shown are the I-splines for geographic distance and climatic distance represented by the first three principal components (PC) of a principal component analysis. The model explained 71% ($p = 0.002$) of the deviance in the genetic distance among the populations. The grey area shows one standard deviation of the bootstrapped I-splines. Red tick marks on the x-axis correspond to the spread of the data used in the analysis.

Table 1. Summary of analysis of molecular variance (AMOVA) results for northeast and southwest Japan and all 30 *Thuja standishii* populations.

Source	df	SS	MS	Est. Var.	%
Among northeast and southwest Japan	1	107.186	107.186	0.137	3.17
Among Pops	28	598.797	21.386	0.323	7.48
Within Pops	1604	6183.825	3.855	3.855	89.35
Total	1633	6889.808		4.314757	100

df = degrees of freedom, SS = sum of squares, MS = mean of the squares, Est. Var. = estimated variance of components, % = percentage of total variance contributed by each component.
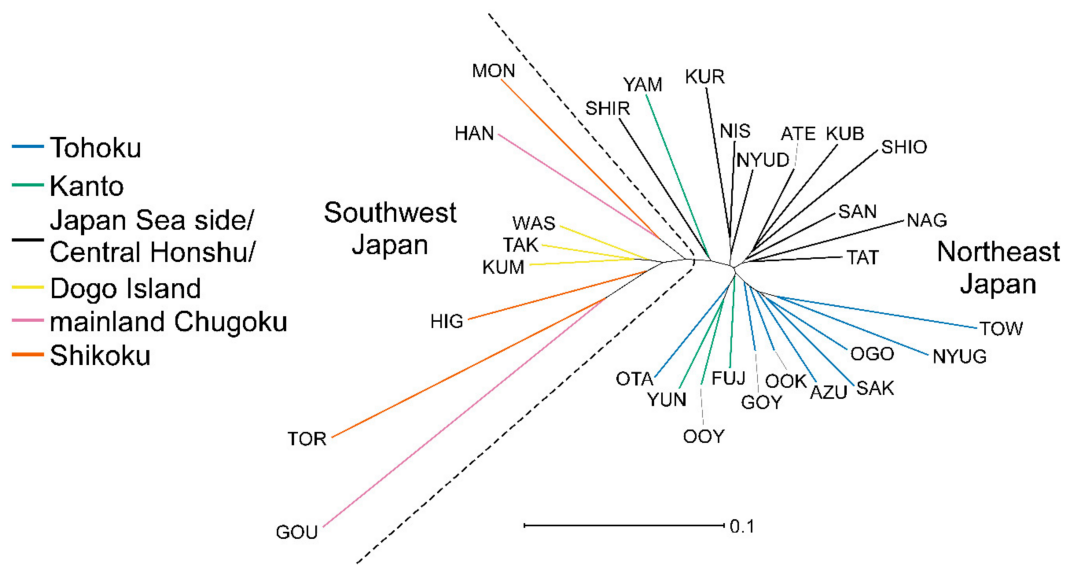
Figure 3. A neighbor-joining unrooted tree based on Nei's *Da* of all 30 populations of *Thuja standishii*. The geographic region of each population is indicated by the color of the branches while the division between southwest and northeast is indicated by a broken line. For full names of populations see Table S1 in the Supplementary Materials.
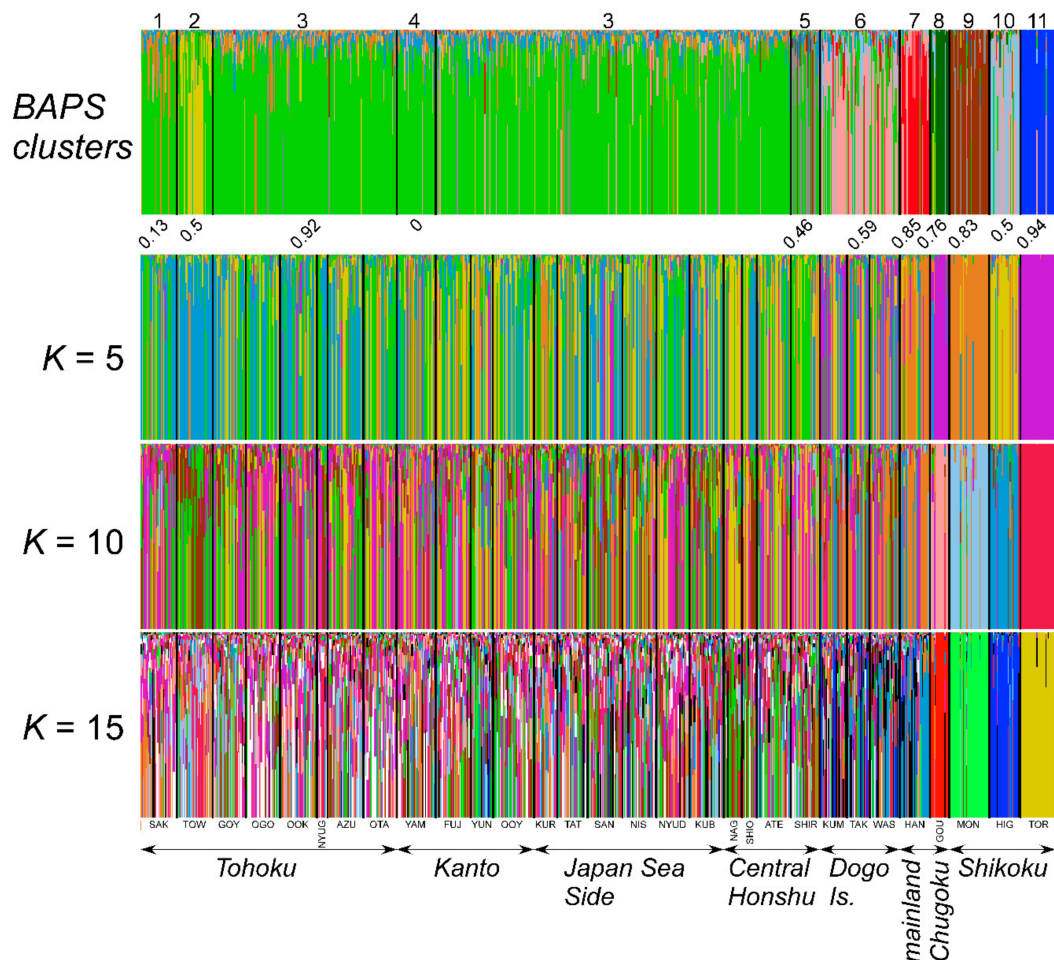

Figure 4. Membership probability to genetic clusters inferred by the discriminant analysis of principal components (DAPC) method. The uppermost plot shows the *a priori* analysis using eleven Bayesian analysis of population structure (BAPS) -based pre-defined populations, with the re-assignment probability indicated below each cluster. The bottom three plots are based on the *a posteriori* analysis with three levels of *K* (*K* = 5, 10 and 15) shown. Populations are ordered from northeast to southwest of Japan with populations and BAPS clusters delineated by black lines.

For the *a posteriori* DAPC analysis, the value of BIC decreased gradually from *K* = 10 to *K* = 15 (Figure S6). Clustering of individuals was mostly consistent with the genetic relationships based on the NJ tree with clear differentiation between northeast and southwest Japan populations, the grouping of the southwest populations GOU/ TOR and HAN/ MON at *K* = 5 and the strong divergence of southwest populations: all Shikoku and Chugoku populations were dominated by single unique clusters at *K* = 10 and *K* = 15 (Figure 4). In addition, DAPC and the NJ tree clearly shows that the three Dogo Island populations are closely genetically related. However, in contrast to the NJ tree, little discernible genetic differentiation was evident in the DAPC analysis within northeast Japan with the exception of the northern Tohoku TOW population and SHIR. The *a priori* DAPC analysis undertaken using the pre-defined 30 populations clearly supported the inferences made from the *a posteriori* analysis. Thus, the highest re-assignment values were in southwest Japan (Table S1, Figure 4), being over 0.82 for all Shikoku and mainland Chugoku populations indicating that these were clearly distinct populations. However, for the three Dogo Island populations, re-assignment values were low (average = 0.31). In northeast Japan, re-assignment values were low overall (average = 0.22). Tohoku had the highest values (average = 0.29) with the two most northern populations, SAK and TOW, forming moderately to clearly defined clusters with re-assignment values of 0.56 and 0.91, respectively. Lastly, Kanto (average = 0.17) and Japan Sea side/central Honshu (average = 0.19) had low values with the exception of SHIR (0.73). Re-assignment values based on the eleven pre-defined BAPS clusters were all equal to or above 0.5 except for SAK, SHIR and YAM populations (Figure 4).

3.4. Genetic Diversity and Bottlenecks

A total of 216 alleles were found at the 13 loci with 122 alleles shared between both northeast and southwest Japan. Ninety alleles were restricted to the northeast and only four to the southwest. Over all populations, average *Ar* was 4.49 and *Ho* was 0.57 (Table S1). Compared to the northeast (*Ar* = 4.76 and *Ho* = 0.59) lower genetic diversity was observed in the southwest (*Ar* = 3.73 and *Ho* = 0.49) (Table S1). *Fis* values at the population level varied between −0.069 and 0.301 with eleven populations displaying significant positive *Fis* values (Table S1). Four of the eight southwest populations had significant *Fis* values with the highest *Fis* value of 0.302 at the GOU population. Populations in northeast Japan with significant *Fis* values were mainly observed in the Japan Sea side/Central Honshu region including some small (KUB) and isolated (ATE, SHIO) populations. Only one population was observed with significant *Fis* values in both Kanto and Tohoku. At the regional level, genetic diversity levels differed markedly with highest *Ar*, *PAr* and *Ho* found in the Japan Sea side/Central Honshu and Tohoku regions and lowest in mainland Chugoku and Shikoku (Table 2: Figure 5). The lowest genetic diversity at the population level were observed in the isolated mainland Chugoku population, GOU, (*Ar* = 2.19 and *Ho* = 0.24)), while the highest was observed in the Japan Sea side population SAN (*Ar* = 5.42 and *Ho* = 0.66) (Table S1). *Ar* values > 5.0 were found on the Japan Sea side (SAN, TAT, NIS), Central Honshu (KUB, SHIO), Tohoku (GOY) and Kanto (YUN) (Figure 5) while lowest values for the northeast were observed at SHIR, TOW and YAM. In southwest Japan, population level *Ar* was lower than all northeast populations except for the Dogo Island populations (Figure 5). Twenty-one populations harboured private alleles with a maximum of four found in the Japan Sea side population TAT and the Tohoku population AZU and three each at GOY, OGO (both Tohoku) and YAM (Kanto) (Table S1). Twelve populations had one private allele of which three were from southwest Japan (TAK, HAN and MON). The Japan Sea side/Central Honshu and Tohoku regions had twenty-three and twenty-two private alleles, respectively, followed by eight in the Kanto region in comparison to two on Dogo Island and one each in mainland Chugoku and Shikoku regions (Table 2). Over all populations, significant increases in *Ar* with increasing latitude (*p* = 0.041) and "northeastness" (*p* = 0.0056) were observed while *PAr* significantly increased

with increasing "northeastness" ($p = 0.0035$) but not versus latitude ($p = 0.09$). No significant correlations were found for northeast populations (results not shown).

Table 2. Genetic statistics based on thirteen expressed sequence tags (EST) microsatellite loci for the six defined regions of *Thuja standishii* including the number of alleles (*Na*), number of effective alleles (*Ne*), observed (*Ho*) and expected (*He*) heterozygosity, fixation index (*Fis*), rarefied allelic richness (*Ar*) and rarefied private alleles (*PAr*) (both rarefied using a minimum sample size of 43 individuals) and the actual number of observed private alleles. In addition, the genetic statistics for northeast and southwest Japan groupings are also shown.

| Region	*N*	*Na*	*Ne*	*Ho*	*He*	*Fis*	*Ar* (43 ind.)	*PAr* (43 ind.)	No. of Private Alleles
Tohoku	230	12.54	3.86	0.58	0.63	0.07	9.07	1.06	22
Kanto	122	10.77	3.40	0.58	0.61	0.05	8.82	0.8	8
Japan Sea side/Central Honshu	255	13.46	4.18	0.60	0.65	0.07	9.52	1.15	23
Dogo Island	71	8.15	3.63	0.57	0.62	0.07	7.36	0.2	2
Mainland Chugoku	44	5.00	2.65	0.40	0.54	0.25	4.98	0.09	1
Shikoku	95	7.69	3.26	0.49	0.62	0.21	6.91	0.24	1
northeast Japan	607	16.31	4.02	0.59	0.64	0.08	9.73	4.13	90
southwest Japan	210	9.69	3.85	0.50	0.63	0.20	7.7	2.17	4
Total	27.23	6.22	3.14	0.57	0.58	0.02			
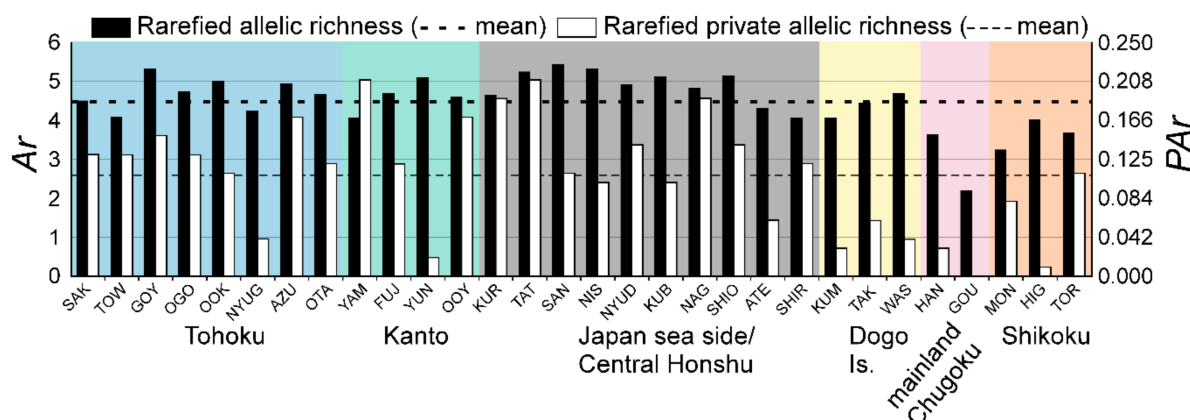
Figure 5. Rarefied allelic richness (*Ar*) and rarefied private allelic richness (*PAr*) for the 30 populations of *Thuja standishii* based on the thirteen nuclear EST microsatellites. Populations are ordered by decreasing latitude from the northeast to southwest Japan. For full names of populations see Table S1 in the Supplementary Materials.

Only one population, GOU from mainland Chugoku, was found to have evidence for a past bottleneck using the mode-shift test (Table S6). On the other hand, no population showed evidence for a bottleneck based on heterozygosity excess (Table S6).

3.5. Past Demographic Change and Estimates of Migration

The average classification error rate values of ABC-RF model choice in population size change and population divergence analyses for the six and two regions were 0.335 and 0.176, respectively (Table S7). Therefore, our model choices were considered to be conducted with relatively moderate precision. We also checked the goodness-of-fit of the best models for each region with predictive simulation and confirmed that in most summary statistics, predictive values of the best model were located near the observed value (results not shown).

In the population size change analysis of the six regions, PGM was selected in the three northeast Japan regions (Tohoku, Kanto and Japan Sea side/Central Honshu) and two southwest Japan regions (Dogo Island and Shikoku) with posterior probability between 0.533–0.914 (Table S7). SNM was selected in only the mainland Chugoku region. However, in Shikoku, as the proportion of votes of the best and the second-best models (PGM and SNM) were very near (0.516 and 0.451) and all the other results based on genetic diversity and species distribution modeling were inconsistent with PGM, we selected SNM for Shikoku. Trajectories of population size change of the three northeast Japan regions and Dogo Island were quite contrasting (Figure 6). The three northeast Japan regions started population expansion before 100,000 years ago and reached near the current size ca. 10,000 years ago while Dogo Island slowly grew in population size and suddenly expanded its size around 1000 years ago and recently reached near the current size. Mainland Chugoku and Shikoku showed smaller current effective population size compared with the other regions (Figure 7).
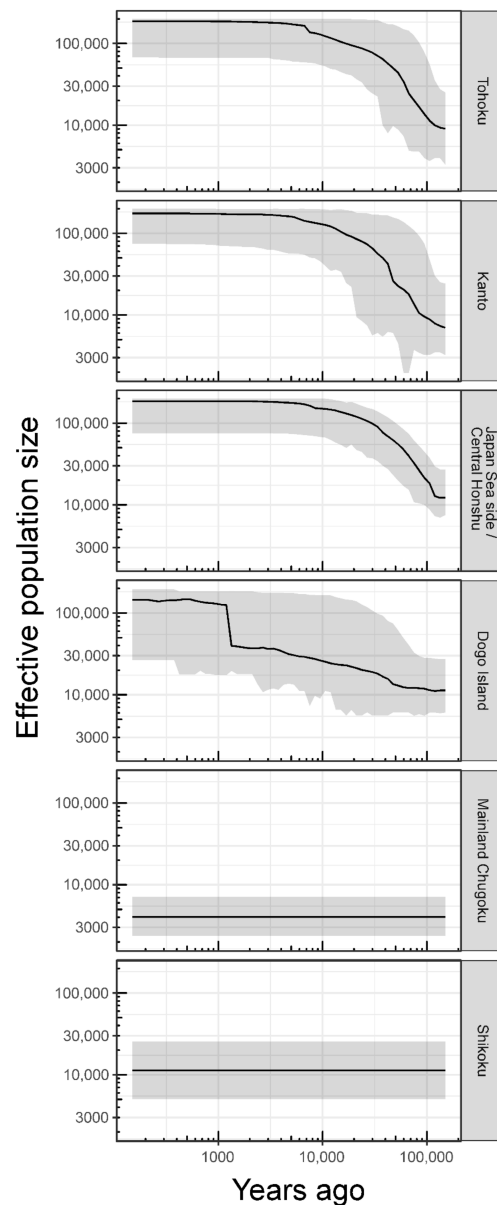
Figure 6. Trajectories of effective population size in the six regions of *Thuja standishii* as calculated by approximate Bayesian computation (ABC). Black lines and gray areas indicate the posterior mode and 95% highest posterior density, respectively.
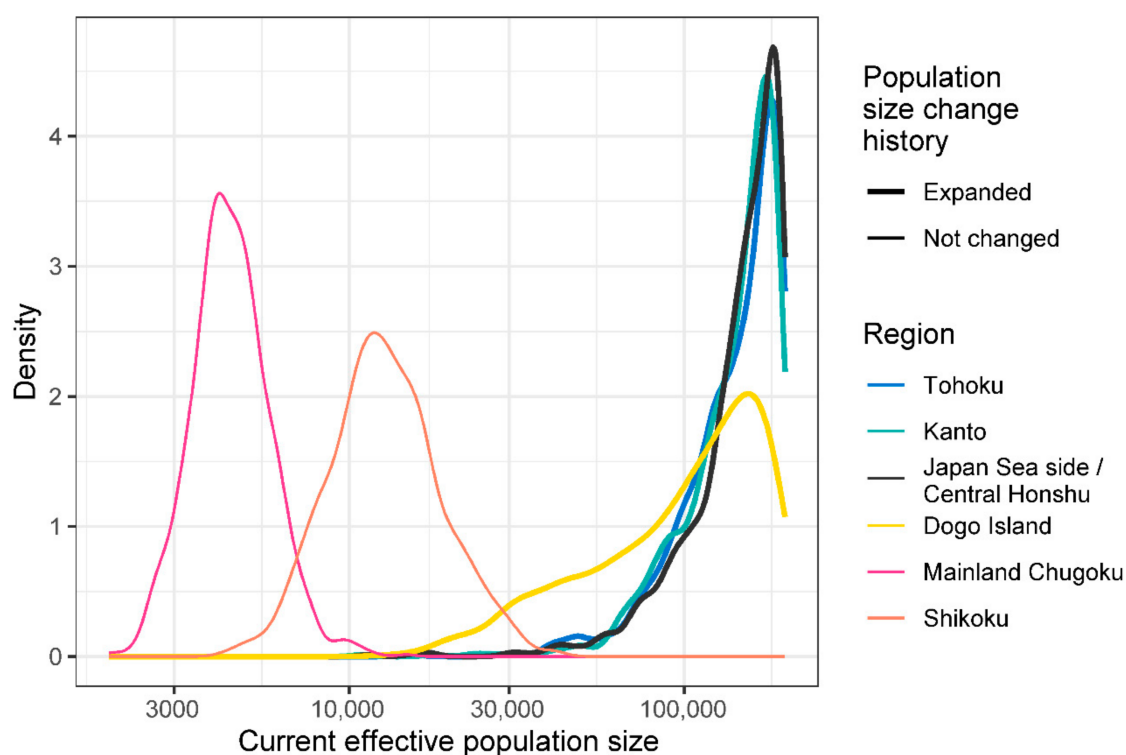
Figure 7. Distributions of current effective population size in the six regions of *Thuja standishii* and their inferred population size change history calculated by approximate Bayesian computation (ABC).

In the population divergence analysis of within northeast and southwest Japan regions, WMM1 was selected in both regions with posterior probability 0.699 and 0.612, respectively (Table S7). Posterior modes (95% HPDs) of divergence time (T_3) for northeast and southwest Japan regions were 176,000 (150,000–2,730,000) and 250,000 (153,000–11,412,000) years ago, respectively (Table S8). Posterior modes (95% HPDs) of ancestral effective population size before divergence (N_{ANC_ALL}) were 1050 (17–16,100) and 15,400 (40–159,000), respectively. The average numbers of migrants per generation (*NM*) ranged between 2.91–29.84 and 0.47–3.73 in northeast and southwest Japan, respectively (Figure S7).

3.6. Species Distribution Modelling

The mean and standard deviation of the AUC, CBI and Brier score values for the MaxEnt models were 0.92 ± 0.02, 0.92 ± 0.04 and 0.055 ± 0.001 which indicates excellent prediction accuracy [98,107,108]. MeTWaQ was the most important climatic variable, followed by PrWaQ and MiTCM while PrCQ has the lowest contribution (i.e., importance) to the model (Table S9). The response of occurrence probability to the four climatic variables (Figure S8) showed that cooler summer mean temperature (lower than 14 °C), moderate precipitation during summer (ca. 800 mm) and moderate minimum temperature of coldest month (ca. −8 °C) were suitable climatic conditions for *T. standishii* distribution.

The predicted occurrence for the current climate showed that modelled potential habitat for the species was consistent with its actual distribution, with highest suitability values in Japan Sea side/central Honshu and mountain ranges of Tohoku and lower probability in southwest Japan (Figure 8). Most occurrences were within areas with occurrence probabilities of values over 0.1 and well within the area considered suitable habitat under both thresholds. The major exceptions were the isolated GOU and HAN populations in mainland Chugoku which were predicted only by the 99% threshold (occurrence probability values of between 0.001–0.02) (Figure 8). Southern Hokkaido, the highest mountains of the Kii Pensinsula and most of the mountains in Chugoku were predicted as potential habitat where the species is absent (i.e., empty habitat, [95]). During the LGM, both MIROC and CCSM predicted the highest habitat suitability to be in southwest Japan (Chugoku

and Shikoku), the Kii Peninsula and the southern portion of the species range in the Japan Sea side/central Honshu region (Figure 9). The northern range limit during the LGM was predicted to be coastal areas of Tohoku either on the Japan Sea side (under CCSM) or on the Pacific side (under MIROC). In contrast, northern parts of central Honshu, inland and northern Tohoku and Hokkaido were predicted to be non-habitat during the period. Some areas, particularly in southwest Japan were predicted to have had suitable climates from the LGM to the present. The two LGM fossil records from northeast Japan were only predicted as suitable habitat using the 99% threshold with occurrence probability values of between 0.003–0.006, except for the fossil record from the Japan Sea side which had relatively high habitat suitability of 0.12 under CCSM (Figure 9).
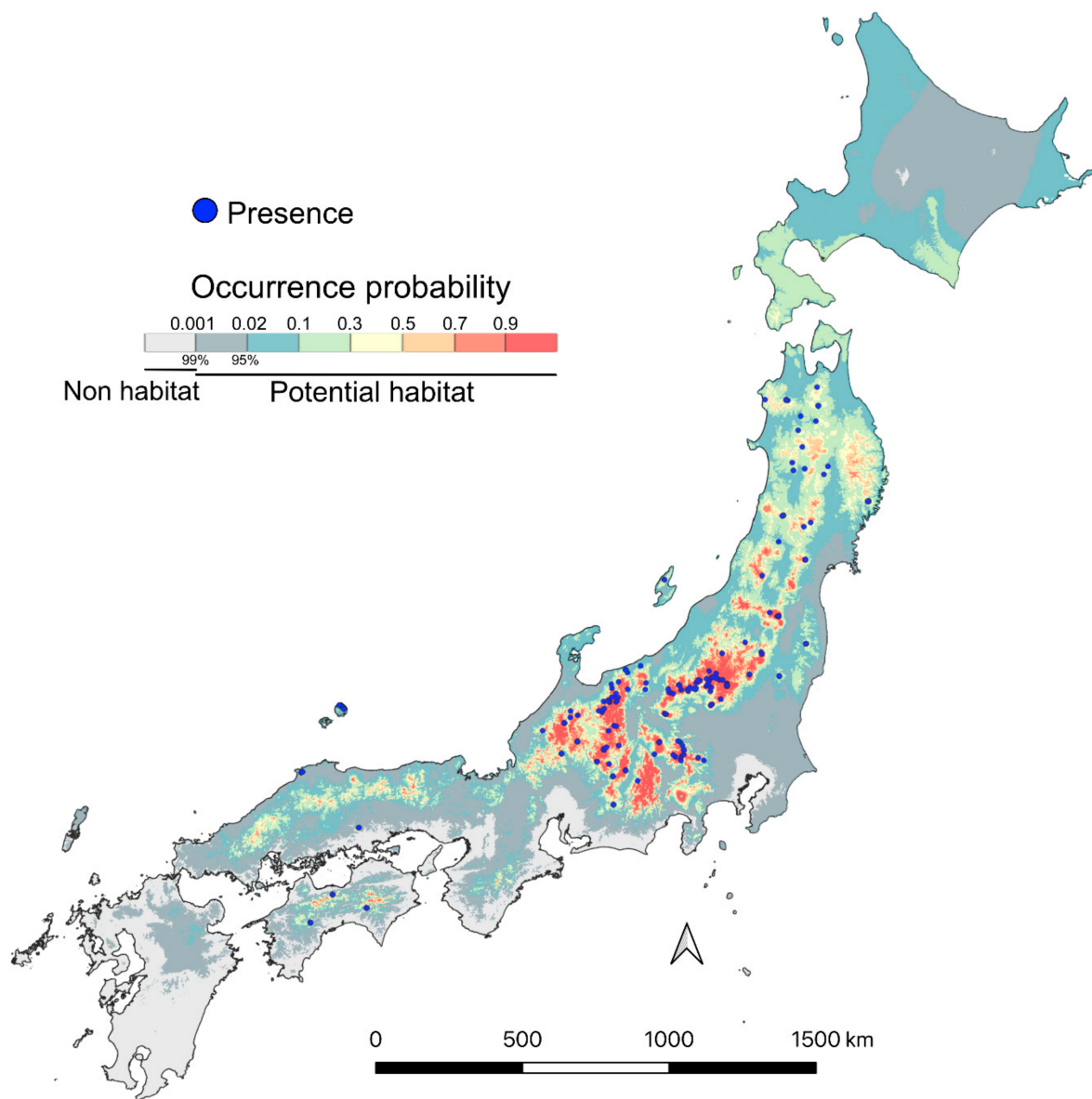
Figure 8. The modelled occurrence probability under current climate for *Thuja standishii* based on four climatic variables. Presence records used for the model are shown as blue dots. Both thresholds of occurrence probability based on cut-off sensitivity values of 95% and 99% indicating areas of potential habitat are shown.
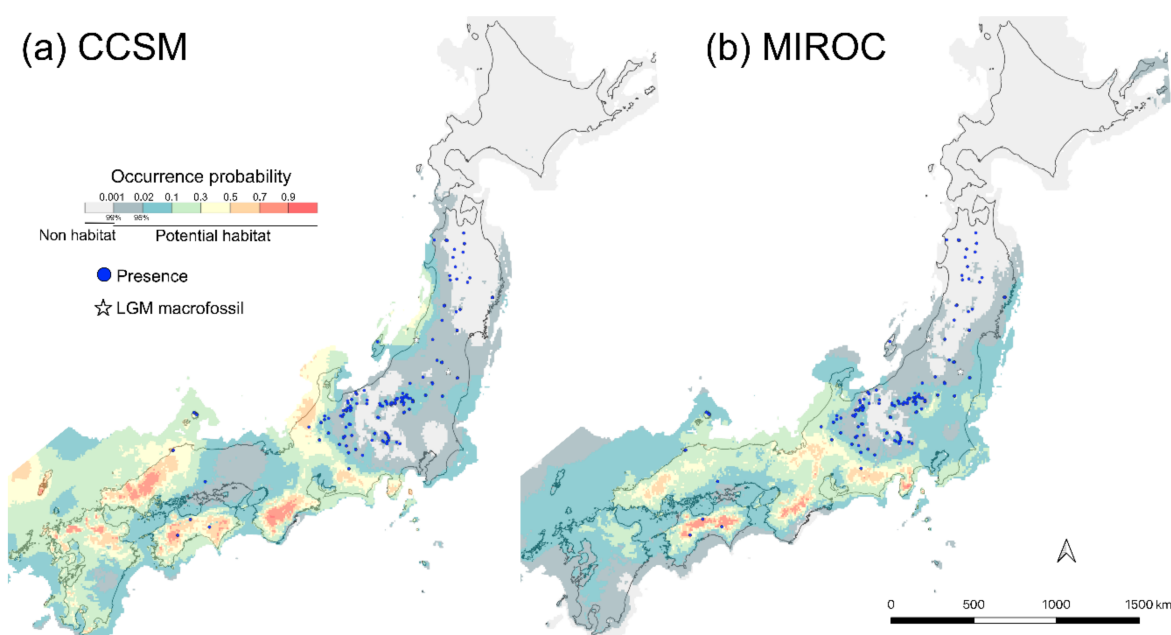
Figure 9. The modelled occurrence probability of *Thuja standishii* under the (**a**) CCSM and (**b**) MIROC Last Glacial Maximum global circulation model. Both thresholds of occurrence probability based on cut-off sensitivity values of 95% and 99 % indicating areas of potential habitat are shown. Model occurrence probabilities are projected on to land down to the LGM coastline with the present coastline outlined in black. The locations of the two known LGM fossil records [29,30] are indicated by white stars and current present records are shown as blue dots.

4. Discussion

Assessment of the range-wide population genetic structure and diversity of *T. standishii* shows that the species fits a central-marginal genetic pattern with highest genetic diversity and effective population size in the core range in northeast Japan. Despite low differentiation between populations in northeast Japan, most likely due to the thorough sampling, some genetic relationships (although weak) were evident. Thus, Tohoku and Japan Sea side/Central populations formed separate clusters in the neighbour joining tree (Figure 3). On other hand, the most diverged populations were at the rear-edge in southwest Japan and, to a lesser extent, some leading-edge populations in northern Tohoku. The isolated southwest populations in Shikoku and mainland Chugoku were the most diverged and had the smallest effective population sized with evidence for limited gene flow between them. The factors underlying the observed genetic diversity and differentiation in both southwest and northeast Japan and conservation implications of the genetic findings are discussed below.

4.1. High Diversity and Range Dynamics in Northeast Japan

The high genetic diversity and low differentiation of populations in northeast Japan show that this region has supported meta-populations of *T. standishii* over time. ABC results support an overall net expansion of the species in northeast Japan beginning approximately 100,000 years ago (Figure 6) during the Last Glacial. A number of lines of evidence suggest that the core range in northeast Japan has undergone expansion and contraction repeatedly in response to the glacial interglacial cycles particularly in the most northern part of the range. Firstly, during the LGM areas of high habitat suitability are modelled to have been restricted to the southern portion of the species present northeast range and along the coastline of Tohoku. Secondly, a submerged forest of *T. standishii* from the Middle Pleistocene (marine isotope stage 8 starting 300–243 kya) from Aomori Prefecture beyond the present northern limit of the species at 41.38° latitude [84], shows that the species was present in Tohoku during previous glacial periods. However, the extent of contraction and expansion latitudinally is uncertain. Indeed, there is evidence

for some populations persisting in areas of northeast Japan with no or extremely low modelled occurrence probability during the LGM including: (1) high levels of unique alleles in Tohoku populations, including GOY, OGO and AZU; (2) no significant decline in genetic diversity with increasing distance northwards in northeast Japan, as may be expected if Tohoku populations were established via postglacial migration [109]; and (3) no trend for an increase in genetic bottlenecks and inbreeding in Tohoku populations which were signatures of recent expansion into northern Japan in *Pinus densiflora* Siebold and Zucc. [110]. Similar evidence for small refugia in areas outside areas of modelled suitable habitat during the LGM in Tohoku has been found in other temperate trees including *Cryptomeria japonica* [111]) and *Kalopanax septemlobus* (Thunb. ex A.Murr.) Koidz. [112]. This discrepancy between modelled LGM distribution and genetic evidence could be due to the global circulation models used for LGM modelling simply not accurately reflecting LGM climates at the regional level or potentially the realized niches of species have changed since the LGM [113].

4.2. Genetic Distinctiveness and Biogeographic History of Rear-Edge Populations in Shikoku and Chugoku

The combined evidence from genetic data and SDM suggest that southwest populations of *T. standishii* are the relicts of populations that have persisted in the region since at least the Last Glacial. Firstly, a recent origin via postglacial spread from the northeast is unlikely because all populations in the southwest (excluding between populations on Dogo Island) are highly differentiated from one another and their divergence exceeds that between nearly all populations in the northeast range [114]. Secondly, ABC indicates that divergence of southwest regions occurred earlier than between regions in northeast Japan. Lastly, the SDM results show that suitable habitat has been available in southwest Japan from the LGM to present, although habitat suitability was higher in the LGM. This is also supported by fossil evidence that indicates the widespread occurrence of cool climate forests, including Cupressaceae conifers, in lowland areas of southwest Japan during the LGM (Gotanda et al. 2008).

The low genetic diversity of southwest rear-edge populations versus northeast populations, with southwest populations only harboring a subset of the variation found in the northeast with very few unique alleles, was unexpected given the evidence for the long-term existence of the species in the southwest. However, this finding is not unique, having been observed in a range of other plant species [115,116] including in the widespread endemic Japanese conifers *Sciadopitys verticillata* (Thunb.) Siebold and Zucc. [117] and *Thujopsis dolabrata* (Thunb. ex. L.f.) Siebold and Zucc. [118]. In *T. standishii*, the low diversity is likely a result of the forces of strong genetic drift eroding allelic diversity probably related to the declining habitat suitability since the LGM. Genetic drift also reduces the probability that new mutations survive [119,120], which may explain the almost complete absence of private alleles. Genetic drift has likely been stronger than in northeast Japan with for example even the extremely small Sado Island population (NYUG) having higher *Ar* than any population in mainland Chugoku or Shikoku (Figure 5). Another factor may be low levels of ongoing gene flow (*NM*) as evidenced by the migration analysis. A similar case has been observed in the southeast Australian cool temperate rainforest tree *Nothofagus cunninghamii* (Hook.f.) Heenan and Smissen [121], whereby rear-edge mainland Australian populations harbor only a subset of the genetic variation of the core range ~250 km to the south on the island of Tasmania. This was explained as a result of low effective population size, including during the LGM, and low levels of past and ongoing gene flow from core populations in Tasmania being insufficient to overcome genetic drift [121].

Apart from the most extreme case of population extinction, the impact of humans on *T. standishii* in the southwest is uncertain given the inherent difficulty in evaluating disturbance from humans on genetic diversity of forest trees [122]. In southwest Japan, fragmentation by humans has been implicated in genetic bottlenecks observed in the warm temperate evergreen tree *Castanopsis cuspidata* (Thunb.) Schottky [123] and it is likely that during the long history of human habitation in southwest Japan that *T. standishii* forests

have also been impacted. Indeed, population extinction of *T. standishii* due to logging has occurred in southwest Japan in the 20th century [20] which, based on the results of this study, likely involved the loss of a unique genetic lineages (as is found in all extant populations in the southwest), but little decline in the species overall neutral allelic diversity. In extant southwest populations evidence for recent genetic bottlenecks is lacking, with the only exception being the small GOU population of mainland Chugoku, where a mode shift in allele frequency was detected.

4.3. Distinct Genetic Lineage of T. standishii on Dogo Island

Fossil evidence shows that Dogo Island was an important Last Glacial Maximum refugia for temperate forest species in Japan [124]. However, to date the genetic status of Dogo Island species has seldom been investigated including for an enigmatic group of usually subalpine/alpine species that occur down to sea level on Dogo Island consisting of *Thuja standishii*, *Quercus crispula* Blume, *Schizocodon soldanelloides* (Siebold and Zucc.) Makino and *Allium schoenoprasum* L. var. *orientale* Regel [125]. In the one previous range-wide population genetic study including a population from Dogo Island, in *Cryptomeria japonica*, although low genetic differentiation of Dogo Island from mainland Honshu populations was found, the number of rare and private alleles was one of the highest in the species [13] consistent with fossil pollen evidence of the species LGM survival [124]. In this study, we found that *T. standishii* on Dogo Island forms a distinct genetic lineage with genetic diversity exceeding any other populations in southwest Japan. ABC results indicate a unique demographic history of the species on the island involving a postglacial increase in effective population size. However, despite this there is evidence for low levels of gene flow as having occurred and possibly still occurring, with other southwest populations. Within the island, low genetic differences between the populations indicates sufficient gene-flow between populations.

4.4. Conservation Implications

Despite a scattered distribution and being rarer than many other Japanese conifers, *T. standishii* in northeast Japan was found to consist of a large meta-population with high genetic diversity, high gene flow (apart from the most northern populations in Tohoku) and little evidence for widespread inbreeding or genetic bottlenecks. This suggests that the species in the northeast of its range is resilient to factors that could degrade genetic diversity. In contrast, a long history of genetic drift and low gene flow in southwest populations, may result in further loss of genetic diversity in the future. However, decline in these rear-edge populations is not at all certain due to local adaptation and other local mechanisms [2]. In fact, their occurrence at relatively low elevations at the margins of the species ecological range make it possible that they harbor unique adaptive genetic variation [8]. Further research is needed to understand the extent of divergence of these populations at adaptive regions of the genome including using both genomic and common garden-based methods.

for each population; Table S7: Summary of ABC-RF model comparison; Table S8: Posterior mode and 95% highest posterior density of structural parameters in the best demographic models; Table S9: The percent contribution and permutation importance of each of the four bioclimatic variables used for species distribution modelling.

References

1. Provan, J.; Maggs, C.A. Unique genetic variation at a species' rear edge is under threat from global climate change. *Proc. R. Soc. B Boil. Sci.* **2011**, *279*, 39–47. [CrossRef] [PubMed]
2. Vilà-Cabrera, A.; Premoli, A.C.; Jump, A.S. Refining predictions of population decline at species' rear edges. *Glob. Chang. Biol.* **2019**, *25*, 1549–1560. [CrossRef] [PubMed]
3. Arnaud-Haond, S.; Teixeira, S.; Massa, S.I.; Billot, C.; Saenger, P.; Coupland, G.; Duarte, C.M.; Serrao, E. Genetic structure at range edge: Low diversity and high inbreeding in Southeast Asian mangrove (*Avicennia marina*) populations. *Mol. Ecol.* **2006**, *15*, 3515–3525. [CrossRef] [PubMed]
4. Hampe, A.; Petit, R.J. Conserving biodiversity under climate change: The rear edge matters. *Ecol. Lett.* **2005**, *8*, 461–467. [CrossRef]
5. Morales-Molino, C.; Tinner, W.; Perea, R.; Carrión, J.S.; Colombaroli, D.; Valbuena-Carabaña, M.; Valbuena-Carabaña, M.; Zafra, E.; Gil, L. Unprecedented herbivory threatens rear-edge populations of *Betula* in southwestern Eurasia. *Ecology* **2019**, *100*, e02833. [CrossRef]
6. Lepais, O.; Muller, S.D.; Saad-Limam, S.B.; Benslama, M.; Rhazi, L.; Belouahem-Abed, D.; Gammar, A.M.; Ghrabi-Gammar, Z.; Valbuena-Carabaña, M.; Zafra, E.; et al. High genetic diversity and distinctiveness of rear-edge climate relicts maintained by ancient tetraploidisation for *Alnus glutinosa*. *PLoS ONE* **2013**, *8*, e75029. [CrossRef]
7. Zardi, G.I.; Nicastro, K.R.; Serrão, E.A.; Jacinto, R.; Monteiro, C.A.; Pearson, G.A. Closer to the rear edge: Ecology and genetic diversity down the core-edge gradient of a marine macroalga. *Ecosphere* **2015**, *6*, 1–25. [CrossRef]
8. Woolbright, S.A.; Whitham, T.G.; Gehring, C.A.; Allan, G.J.; Bailey, J.K. Climate relicts and their associated communities as natural ecology and evolution laboratories. *Trends Ecol. Evol.* **2014**, *29*, 406–416. [CrossRef]
9. Pironon, S.; Papuga, G.; Villellas, J.; Angert, A.L.; García, M.B.; Thompson, J.D. Geographic variation in genetic and demographic performance: New insights from an old biogeographical paradigm. *Biol. Rev.* **2016**, *92*, 1877–1909. [CrossRef]
10. Diekmann, O.E.; Serrão, E.A. Range-edge genetic diversity: Locally poor extant southern patches maintain a regionally diverse hotspot in the seagrass *Zostera marina*. *Mol. Ecol.* **2012**, *21*, 1647–1657. [CrossRef]
11. Tzedakis, P.C.; Lawson, I.T.; Frogley, M.R.; Hewitt, G.M.; Preece, R.C. Buffered tree population changes in a Quaternary refugium: Evolutionary implications. *Science* **2002**, *297*, 2044–2047. [CrossRef]
12. Erichsen, E.O.; Budde, K.B.; Sagheb-Talebi, K.; Bagnoli, F.; Vendramin, G.G.; Hansen, O.K. Hyrcanian forests-Stable rear-edge populations harbouring high genetic diversity of *Fraxinus excelsior*, a common European tree species. *Divers. Distrib.* **2018**, *24*, 1521–1533. [CrossRef]
13. Takahashi, T.; Tani, N.; Taira, H.; Tsumura, Y. Microsatellite markers reveal high allelic variation in natural populations of *Cryptomeria japonica* near refugial areas of the last glacial period. *J. Plant Res.* **2005**, *118*, 83–90. [CrossRef]
14. Gugger, P.F.; González-Rodríguez, A.; Rodríguez-Correa, H.; Sugita, S.; Cavender-Bares, J. Southward Pleistocene migration of Douglas-fir into Mexico: Phylogeography, ecological niche modeling, and conservation of 'rear edge' populations. *New Phytol.* **2011**, *189*, 1185–1199. [CrossRef]

15. Hampe, A.; Bairlein, F. Modified dispersal-related traits in disjunct populations of bird-dispersed Frangula alnus (Rhamnaceae): A result of its Quaternary distribution shifts? *Ecography Cop.* **2000**, *23*, 603–613. [CrossRef]

16. Stojnić, S.; Suchocka, M.; Benito-Garzón, M.; Torres-Ruiz, J.M.; Cochard, H.; Bolte, A.; Cocozza, C.; Cvjetković, B.; De Luis, M.; Martínez-Vilalta, J.; et al. Variation in xylem vulnerability to embolism in European beech from geographically marginal populations. *Tree Physiol.* **2018**, *38*, 173–185. [CrossRef]

17. Fuentes-Utrilla, P.; Valbuena-Carabaña, M.; Ennos, R.; Gil, L. Population clustering and clonal structure evidence the relict state of Ulmus minor Mill. in the Balearic Islands. *Heredity* **2014**, *113*, 21–31. [CrossRef]

18. Migliore, J.; Baumel, A.; Juin, M.; Fady, B.; Roig, A.; Duong, N.; Médail, F. Surviving in mountain climate refugia: New insights from the genetic diversity and structure of the relict shrub *Myrtus nivellei* (Myrtaceae) in the Sahara Desert. *PLoS ONE* **2013**, *8*, e73795. [CrossRef]

19. Takahara, H.; Sugita, S.; Harrison, S.P.; Miyoshi, N.; Morita, Y.; Uchiyama, T. Pollen-based reconstructions of Japanese biomes at 0, 6000 and 18,000 14C yr bp. *J. Biogeogr.* **2000**, *27*, 665–683. [CrossRef]

20. Worth, J.R.P. Current distribution and climatic range of the Japanese endemic conifer *Thuja standishii* (Cupressaceae). *Bull. FFPRI* **2019**, *18*, 275–288.

21. Gansert, D. Treelines of the Japanese Alps – altitudinal distribution and species composition under contrasting winter climates. *Flora-Morphol. Distrib. Funct. Ecol. Plants* **2004**, *199*, 143–156. [CrossRef]

22. Tsumura, Y.; Suyama, Y. *Seedling Transfer Guideline of Japanese Tree Spcies*; Bun-ichi Co. Ltd.: Tokyo, Japan, 2015. (In Japanese)

23. Kikuchi, R.; Jae-Hong, P.; Takahashi, H.; Maki, M. Disjunct distribution of chloroplast DNA haplotypes in the understory perennial *Veratrum album* ssp. *oxysepalum* (Melanthiaceae) in Japan as a result of ancient introgression. *New Phytol.* **2010**, *188*, 879–891. [CrossRef] [PubMed]

24. Sugahara, K.; Kaneko, Y.; Ito, S.; Yamanaka, K.; Sakio, H.; Hoshizaki, K.; Suzuki, W.; Yamanaka, N.; Setoguchi, H. Phylogeography of Japanese horse chestnut (*Aesculus turbinata*) in the Japanese Archipelago based on chloroplast DNA haplotypes. *J. Plant Res.* **2010**, *124*, 75–83. [CrossRef] [PubMed]

25. Higashi, H.; Sakaguchi, S.; Ikeda, H.; Isagi, Y.; Setoguchi, H. Multiple introgression events and range shifts in *Schizocodon* (Diapensiaceae) during the Pleistocene. *Bot. J. Linn. Soc.* **2013**, *173*, 46–63. [CrossRef]

26. Iwasaki, T.; Aoki, K.; Seo, A.; Murakami, N. Comparative phylogeography of four component species of deciduous broad-leaved forests in Japan based on chloroplast DNA variation. *J. Plant Res.* **2011**, *125*, 207–221. [CrossRef]

27. Tsukada, M. *Cryptomeria Japonica*: Glacial refugia and Late-Glacial and postglacial migration. *Ecology* **1982**, *63*, 1091–1105. [CrossRef]

28. Takahara, H.; Kitagawa, H. Vegetation and climate history since the last interglacial in Kurota Lowland, western Japan. *Palaeogeogr. Palaeoclim. Palaeoecol.* **2000**, *155*, 123–134. [CrossRef]

29. Suzuki, K. *Picea* cone-fossils from Pleistocene strata of northeast Japan. *Saito Ho-on Kai Mus. Nat. Hist. Res. Bull.* **1991**, *59*, 1–41.

30. Kamoi, Y.; Saito, M.; Fujita, H.; Kobayashi, I. Plant fossil assemblage of the Last Glacial age in the northern part of Niigata Prefecture, Central Japan. *Quat. Res. (Daiyonki-Kenkyu)* **1988**, *27*, 21–29. [CrossRef]

31. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research–an update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef]

32. Murray, M.; Thompson, W. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **1980**, *8*, 4321–4326. [CrossRef]

33. Worth, J.R.P.; Chang, K.S.; Ha, Y.-H.; Qin, A. Development of microsatellite markers for the Japanese endemic conifer *Thuja standishii* and transfer to other East Asian species. *BMC Res. Notes* **2019**, *12*, 1–5. [CrossRef]

34. Raymond, M.; Rousset, F. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. *J. Hered.* **1995**, *86*, 248–249. [CrossRef]

35. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–22.

36. Chapuis, M.-P.; Estoup, A. Microsatellite Null Alleles and Estimation of Population Differentiation. *Mol. Biol. Evol.* **2006**, *24*, 621–631. [CrossRef]

37. Ellis, J.R.; Burke, J.M. EST-SSRs as a resource for population genetic analyses. *Heredity* **2007**, *99*, 125–132. [CrossRef]

38. Tiffin, P.; Hahn, M.W. Coding sequence divergence between two closely related plant species: *Arabidopsis thaliana* and *Brassica rapa* ssp. *pekinensis. J. Mol. Evol.* **2002**, *54*, 746–753. [CrossRef]

39. Pettenkofer, T.; Finkeldey, R.; Müller, M.; Krutovsky, K.V.; Vornam, B.; Leinemann, L.; Gailing, O. Genetic variation of introduced red oak (*Quercus rubra*) stands in Germany compared to North American populations. *Eur. J. For. Res.* **2020**, *139*, 321–331. [CrossRef]

40. Lind, J.F.; Gailing, O. Genetic structure of *Quercus rubra* L. and *Quercus ellipsoidalis* EJ Hill populations at gene-based EST-SSR and nuclear SSR markers. *Tree Genet. Genomes* **2013**, *9*, 707–722. [CrossRef]

41. Woodhead, M.; Russell, J.; Squirrell, J.; Hollingsworth, P.M.; MacKenzie, K.; Gibby, M.; Powell, W. Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Mol. Ecol.* **2005**, *14*, 1681–1695. [CrossRef]

42. Foll, M.; Gaggiotti, O.E. A genome-scan method to Identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* **2008**, *180*, 977–993. [CrossRef] [PubMed]

43. Lotterhos, K.E.; Whitlock, M.C. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol. Ecol.* **2014**, *23*, 2178–2192. [CrossRef] [PubMed]

44. Jeffreys, H. The theory of probability. *Nat. Cell Biol.* **1922**, *109*, 132–133. [CrossRef]

45. Excoffier, L.; Hofer, T.; Foll, M. Detecting loci under selection in a hierarchically structured population. *Heredity* **2009**, *103*, 285–298. [CrossRef] [PubMed]
46. Corander, J.; Waldmann, P.; Sillanpää, M.J. Bayesian analysis of genetic differentiation between populations. *Genetics* **2003**, *163*, 367–374. [CrossRef]
47. Wright, S. Isolation by distance. *Genetics* **1943**, *28*, 114. [CrossRef]
48. Wang, I.J.; Glor, R.E.; Losos, J.B. Quantifying the roles of ecology and geography in spatial genetic divergence. *Ecol. Lett.* **2012**, *16*, 175–182. [CrossRef]
49. Wang, I.J.; Summers, K. Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Mol. Ecol.* **2010**, *19*, 447–458. [CrossRef]
50. Thibert-Plante, X.; Hendry, A.P. When can ecological speciation be detected with neutral loci? *Mol. Ecol.* **2010**, *19*, 2301–2314. [CrossRef]
51. Fitzpatrick, M.C.; Keller, S.R. Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* **2015**, *18*, 1–16. [CrossRef]
52. Ferrier, S.; Manion, G.; Elith, J.; Richardson, K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* **2007**, *13*, 252–264. [CrossRef]
53. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020.
54. Meirmans, P.G.; Van Tienderen, P.H. genotype and genodive: Two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* **2004**, *4*, 792–794. [CrossRef]
55. Nei, M.; Tajima, F.; Tateno, Y. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J. Mol. Evol.* **1983**, *19*, 153–170. [CrossRef]
56. Huson, D.H.; Bryant, D. Application of Phylogenetic Networks in Evolutionary Studies. *Mol. Biol. Evol.* **2005**, *23*, 254–267. [CrossRef]
57. Kalinowski, S.T. Evolutionary and statistical properties of three genetic distances. *Mol. Ecol.* **2002**, *11*, 1263–1273. [CrossRef]
58. Takezaki, N.; Nei, M. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **1996**, *144*, 389–399. [CrossRef]
59. Kalinowski, S.T. How well do evolutionary trees describe genetic relationships among populations? *Heredity* **2009**, *102*, 506–513. [CrossRef]
60. Mondin, L.A.D.C.; Machado, C.B.; De Resende, E.K.; Marques, D.K.S.; Galetti, P.M.J. Genetic pattern and demographic history of *Salminus brasiliensis*: Population expansion in the Pantanal Region during the Pleistocene. *Front. Genet.* **2018**, *9*, 1. [CrossRef]
61. Latch, E.K.; Dharmarajan, G.; Glaubitz, J.C.; Rhodes, O.E. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conserv. Genet.* **2006**, *7*, 295–302. [CrossRef]
62. Provan, J.; Beatty, G.E.; Hunter, A.M.; McDonald, R.A.; McLaughlin, E.; Preston, S.J.; Wilson, S. Restricted gene flow in fragmented populations of a wind-pollinated tree. *Conserv. Genet.* **2007**, *9*, 1521–1532. [CrossRef]
63. Jombart, T.; Devillard, S.; Balloux, F. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet.* **2010**, *11*, 94. [CrossRef] [PubMed]
64. Perez, M.F.; Franco, F.F.; Bombonato, J.R.; Bonatelli, I.A.S.; Khan, G.; Romeiro-Brito, M.; Fegies, A.C.; Ribeiro, P.M.; Silva, G.A.R.; Moraes, E.M. Assessing population structure in the face of isolation by distance: Are we neglecting the problem? *Divers. Distrib.* **2018**, *24*, 1883–1889. [CrossRef]
65. Ramasamy, R.K.; Ramasamy, S.; Bindroo, B.B.; Naik, V.G. STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus* **2014**, *3*, 1–3. [CrossRef] [PubMed]
66. Goudet, J. FSTAT (Version 1.2): A computer program to calculate F-statistics. *J. Hered.* **1995**, *86*, 485–486. [CrossRef]
67. Kalinowski, S.T. hp-rare 1.0: A computer program for performing rarefaction on measures of allelic diversity. *Mol. Ecol. Notes* **2005**, *5*, 187–189. [CrossRef]
68. Luikart, G.; Allendorf, F.W.; Cornuet, J.-M.; Sherwin, W.B. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J. Hered.* **1998**, *89*, 238–247. [CrossRef]
69. Piry, S.; Luikart, G.; Cornuet, J.M. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* **1999**, *90*, 502–503. [CrossRef]
70. Di Rienzo, A.; Peterson, A.C.; Garza, J.C.; Valdes, A.M.; Slatkin, M.; Freimer, N.B. Mutational processes of simple-sequence repeat loci in human populations. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3166–3170. [CrossRef]
71. Chybicki, I.J.; Burczyk, J. Simultaneous estimation of null alleles and inbreeding coefficients. *J. Hered.* **2008**, *100*, 106–113. [CrossRef]
72. Beaumont, M.A. Approximate Bayesian computation in evolution and ecology. *Annu. Rev. Ecol. Evol. Syst.* **2010**, *41*, 379–406. [CrossRef]
73. Excoffier, L.; Lischer, H.E.L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **2010**, *10*, 564–567. [CrossRef]
74. Estoup, A.; Jarne, P.; Cornuet, J.-M. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol. Ecol.* **2002**, *11*, 1591–1604. [CrossRef]
75. Setsuko, S.; Sugai, K.; Tamaki, I.; Takayama, K.; Kato, H.; Yoshimaru, H. Genetic diversity, structure, and demography of *Pandanus boninensis* (Pandanaceae) with sea drifted seeds, endemic to the Ogasawara Islands of Japan: Comparison between young and old islands. *Mol. Ecol.* **2020**, *29*, 1050–1068. [CrossRef]
76. Excoffier, L.; Estoup, A.; Cornuet, J.-M. Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genet.* **2005**, *169*, 1727–1738. [CrossRef]

77. Excoffier, L.; Foll, M. Fastsimcoal: A continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics* **2011**, *27*, 1332–1334. [CrossRef]

78. Pudlo, P.; Marin, J.-M.; Estoup, A.; Cornuet, J.-M.; Gauthier, M.; Robert, C.P. Reliable ABC model choice via random forests. *Bioinformatics* **2015**, *32*, 859–866. [CrossRef]

79. Csilléry, K.; François, O.; Blum, M.G.B. abc: An R package for approximate Bayesian computation (ABC). *Methods Ecol. Evol.* **2012**, *3*, 475–479. [CrossRef]

80. Blum, M.G.B.; François, O. Non-linear regression models for approximate Bayesian computation. *Stat. Comput.* **2010**, *20*, 63–73. [CrossRef]

81. Plummer, M.; Best, N.; Cowles, K.; Vines, K. CODA: Convergence diagnosis and output analysis for MCMC. *R News* **2006**, *6*, 7–11.

82. Lu, R.-S.; Chen, Y.; Tamaki, I.; Sakaguchi, S.; Ding, Y.-Q.; Takahashi, D.; Li, P.; Isaji, Y.; Chen, J.; Qiu, Y.-X. Pre-quaternary diversification and glacial demographic expansions of *Cardiocrinum* (Liliaceae) in temperate forest biomes of Sino-Japanese Floristic Region. *Mol. Phylogenetics Evol.* **2020**, *143*, 106693. [CrossRef]

83. Iwauchi, A.; Hase, Y. Late Cenozoic vegetation and paleoenvironment of northern and central Kyushu, Japan - Part 5 Yoshino area (Middle Pleistocene). *J. Geol. Soc. Jpn.* **1992**, *98*, 205–221. [CrossRef]

84. Fujii, S.; Nara, M.; Hatanaka, S.; Yamaguchi, G.; Suzuki, M.; Takeuti, S.; Yoshida, A.; Noshiro, S.; Muramoto, J.; Horiuchi, K.; et al. Submarine forests around Shimokita Peninsula, north end of Honshu, Japan. *Earth Sci. (Chikyu Kagaku)* **2006**, *60*, 375–387. (In Japanese)

85. Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, B.D.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*; CRC Press: Boca Raton, FL, USA, 2013.

86. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **2006**, *190*, 231–259. [CrossRef]

87. Tsukada, M. Vegetation and climate during the Last Glacial Maximum in Japan. *Quat. Res.* **1983**, *19*, 212–235. [CrossRef]

88. Ono, Y. Last Glacial paleoclimate reconstructed from glacial and periglacial landforms in Japan. *Geogr. Rev. Jpn. Ser. B.* **1984**, *57*, 87–100. [CrossRef]

89. Yoshida, A.; Akihiko, S.; Motonari, O.; Masataka, H.; Akifumi, I. Vegetation reconstruction during the Last Termination on Mt. Chokai, northeastern Japan and habitat of *Larix gemelinii*. *Jpn. J. Histor. Bot.* **2014**, *23*, 21–26. (In Japanese)

90. Minaki, M.; Matsuba, C. Plant macrofossil assemblage from about 18,000 years ago in Tado-cho, Mie Prefecture, central Japan. *Quat. Res. (Daiyonki-Kenkyu)* **1985**, *24*, 51–55. [CrossRef]

91. Momohara, A.; Yoshida, A.; Kudo, Y.; Nishiuchi, R.; Okitsu, S. Paleovegetation and climatic conditions in a refugium of temperate plants in central Japan in the Last Glacial Maximum. *Quat. Int.* **2016**, *425*, 38–48. [CrossRef]

92. Noshiro, S.; Terada, K.; Tsuji, S.-I.; Suzuki, M. Larix-Picea forests of the Last Glacial Age on the eastern slope of Towada Volcano in northern Japan. *Rev. Palaeobot. Palynol.* **1997**, *98*, 207–222. [CrossRef]

93. Elith, J.; Graham, C.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann, F.; Leathwick, J.R.; Lehmann, A.; et al. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **2006**, *29*, 129–151. [CrossRef]

94. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **2005**, *25*, 1965–1978. [CrossRef]

95. Tsuyama, I.; Nakao, K.; Higa, M.; Matsui, T.; Shichi, K.; Tanaka, N. What controls the distribution of the Japanese endemic hemlock, *Tsuga diversifolia*? Footprint of climate in the glacial period on current habitat occupancy. *J. For. Res.* **2014**, *19*, 154–165. [CrossRef]

96. Metz, C.E. Basic principles of ROC analysis. *Semin. Nucl. Med.* **1978**, *8*, 283–298. [CrossRef]

97. Boyce, M.S.; Vernier, P.R.; Nielsen, S.E.; Schmiegelow, F.K. Evaluating resource selection functions. *Ecol. Model.* **2002**, *157*, 281–300. [CrossRef]

98. Hirzel, A.H.; Le Lay, G.; Helfer, V.; Randin, C.; Guisan, A. Evaluating the ability of habitat suitability models to predict species presences. *Ecol. Model.* **2006**, *199*, 142–152. [CrossRef]

99. Brier, G.W. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* **1950**, *78*, 1–3. [CrossRef]

100. Moreno-Amat, E.; Rubiales, J.M.; Morales-Molino, C.; García-Amorena, I. Incorporating plant fossil data into species distribution models is not straightforward: Pitfalls and possible solutions. *Quat. Sci. Rev.* **2017**, *170*, 56–68. [CrossRef]

101. Svenning, J.-C.; Fløjgaard, C.; Marske, K.A.; Nógues-Bravo, D.; Normand, S. Applications of species distribution modeling to paleobiology. *Quat. Sci. Rev.* **2011**, *30*, 2930–2947. [CrossRef]

102. Williams, J.N.; Seo, C.; Thorne, J.; Nelson, J.K.; Erwin, S.; O'Brien, J.M.; Schwartz, M.W. Using species distribution models to predict new occurrences for rare plants. *Divers. Distrib.* **2009**, *15*, 565–576. [CrossRef]

103. Van Proosdij, A.S.J.; Sosef, M.S.M.; Wieringa, J.J.; Raes, N. Minimum required number of specimen records to develop accurate species distribution models. *Ecography* **2016**, *39*, 542–552. [CrossRef]

104. Gent, P.R.; Danabasoglu, G.; Donner, L.J.; Holland, M.M.; Hunke, E.C.; Jayne, S.R.; Lawrence, D.M.; Neale, R.B.; Rasch, P.J.; Vertenstein, M.; et al. The Community Climate System Model Version. *J. Clim.* **2011**, *24*, 4973–4991. [CrossRef]

105. Watanabe, S.; Hajima, T.; Sudo, K.; Nagashima, T.; Takemura, T.; Okajima, H.; Nozawa, T.; Kawase, H.; Abe, M.; Yokohata, T.; et al. MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geosci. Model Dev.* **2011**, *4*, 845–872. [CrossRef]

106. Pearson, R.G.; Dawson, T.P.; Liu, C. Modelling species distributions in Britain: A hierarchical integration of climate and land-cover data. *Ecography* **2004**, *27*, 285–298. [CrossRef]

107. Hosner, S.; Lemeshow, D.W. *Applied Logistic Regression*; John Wiley & Son: New York, NY, USA, 1989.
108. Maguire, K.-C.; Nieto-Lugilde, D.; Blois, J.L.; Fitzpatrick, M.C.; Williams, J.W.; Ferrier, S.; Lorenz, D.J. Controlled comparison of species- and community-level models across novel climates and communities. *Proc. R. Soc. B* **2016**, *283*, 20152817. [CrossRef]
109. Comps, B.; Gömöry, D.; Letouzey, J.; Thiébaut, B.; Petit, R.J. Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* **2001**, *157*, 389–397. [CrossRef]
110. Iwaizumi, M.G.; Tsuda, Y.; Ohtani, M.; Tsumura, Y.; Takahashi, M. Recent distribution changes affect geographic clines in genetic diversity and structure of *Pinus densiflora* natural populations in Japan. *For. Ecol. Manag.* **2013**, *304*, 407–416. [CrossRef]
111. Kimura, M.K.; Uchiyama, K.; Nakao, K.; Moriguchi, Y.; Jose-Maldia, L.S.; Tsumura, Y. Evidence for cryptic northern refugia in the last glacial period in *Cryptomeria japonica*. *Ann. Bot.* **2014**, *114*, 1687–1700. [CrossRef]
112. Sakaguchi, S.; Qiu, Y.-X.; Liu, Y.-H.; Qi, X.-S.; Kim, S.-H.; Han, J.; Takeuchi, Y.; Worth, J.R.P.; Yamasaki, M.; Sakurai, S.; et al. Climate oscillation during the Quaternary associated with landscape heterogeneity promoted allopatric lineage divergence of a temperate tree *Kalopanax septemlobus* (Araliaceae) in East Asia. *Mol. Ecol.* **2012**, *21*, 3823–3838. [CrossRef]
113. Worth, J.R.P.; Williamson, G.J.; Sakaguchi, S.; Nevill, P.G.; Jordan, G.J. Environmental niche modelling fails to predict Last Glacial Maximum refugia: Niche shifts, microrefugia or incorrect palaeoclimate estimates? *Glob. Ecol. Biogeogr.* **2014**, *23*, 1186–1197. [CrossRef]
114. Callahan, C.M.; Rowe, C.A.; Ryel, R.J.; Shaw, J.D.; Madritch, M.D.; Mock, K.E. Continental-scale assessment of genetic diversity and population structure in quaking aspen (*Populus tremuloides*). *J. Biogeogr.* **2013**, *40*, 1780–1791. [CrossRef]
115. Cho, W.; So, S.; Han, E.K.; Myeong, H.H.; Park, J.S.; Hwang, S.H.; Hwang, S.-H.; Kim, J.-H.; Lee, J.-H. Rear-edge, low-diversity, and haplotypic uniformity in cold-adapted *Bupleurum euphorbioides* interglacial refugia populations. *Ecol. Evol.* **2020**, *10*, 10449–10462. [CrossRef]
116. Dering, M.; Kosiński, P.; Wyka, T.P.; Pers-Kamczyc, E.; Boratyński, A.; Boratyńska, K.; Reich, P.B.; Romo, A.; Zadworny, M.; Żytkowiak, R.; et al. Tertiary remnants and Holocene colonizers: Genetic structure and phylogeography of Scots pine reveal higher genetic diversity in young boreal than in relict Mediterranean populations and a dual colonization of Fennoscandia. *Divers. Distrib.* **2017**, *23*, 540–555. [CrossRef]
117. Worth, J.R.P.; Yokogawa, M.; Pérez-Figueroa, A.; Tsumura, Y.; Tomaru, N.; Janes, J.K.; Isagi, Y. Conflict in outcomes for conservation based on population genetic diversity and genetic divergence approaches: A case study in the Japanese relictual conifer *Sciadopitys verticillata* (Sciadopityaceae). *Conserv. Genet.* **2014**, *15*, 1243–1257. [CrossRef]
118. Inanaga, M.; Hasegawa, Y.; Mishima, K.; Takata, K. Genetic diversity and structure of Japanese endemic genus *Thujopsis* (Cupressaceae) Using EST-SSR Markers. *Forests* **2020**, *11*, 935. [CrossRef]
119. Burton, O.J.; Travis, J.M.J. Landscape structure and boundary effects determine the fate of mutations occurring during range expansions. *Heredity* **2008**, *101*, 329–340. [CrossRef] [PubMed]
120. McInerny, G.; Turner, J.; Wong, H.; Travis, J.; Benton, T. How range shifts induced by climate change affect neutral evolution. *Proc. R. Soc. B Boil. Sci.* **2009**, *276*, 1527–1534. [CrossRef]
121. Duncan, C.J.; Worth, J.R.P.; Jordan, G.J.; Jones, R.C.; Vaillancourt, R.E. Genetic differentiation in spite of high gene flow in the dominant rainforest tree of southeastern Australia, *Nothofagus cunninghamii*. *Heredity* **2016**, *116*, 99–106. [CrossRef]
122. Ledig, F.T. Human impacts on genetic diversity in forest ecosystems. *Oikos* **1992**, *63*, 87. [CrossRef]
123. Aoki, K.; Ueno, S.; Kamijo, T.; Setoguchi, H.; Murakami, N.; Kato, M.; Tsumura, Y. Genetic differentiation and genetic diversity of *Castanopsis* (Fagaceae), the dominant tree species in Japanese broadleaved evergreen forests, revealed by analysis of EST-associated microsatellites. *PLoS ONE* **2014**, *9*, e87429. [CrossRef]
124. Takahara, H.; Tanida, K.; Miyoshi, N. The full-Glacial refugium of *Cryptomeria japonica* in the Oki Islands, Western Japan. *Jpn. J. Palynol.* **2001**, *47*, 21–33.
125. Inoue, K.; Mishima, M.; Fukaya, H.; Yahata, H.; Nobe, K. Distributions of several species of northern plants in Oki Islands, Shimane Prefecture. *Bull. Shimane Nat. Mus. Mt Sanbe* **2019**, *17*, 37–43.