BMC Plant Biology

RESEARCH

Open Access

# Next-generation sequencing-based population genetics unravels the evolutionary history of *Rhodomyrtus tomentosa* in China

Xing-ming Xu[1†], Bo-yong Liao[1,4*†], Su-jiao Liao[1†], Qiao-mei Qin[2], Chun-yan He[2], Xin Ding[2], Wei Wu[1], Long-yuan Wang[1], Fang-qiu Zhang[2], Li-xia Peng[2], Bryan T. Drew[3] and Yong-quan Li[1,4*]

## Abstract

**Background** *Rhodomyrtus tomentosa* (Ait.) Hassk. is useful for its ornamental, medicinal, and ecological characteristics, and is considered a "Neglected and Underutilized Crop Species". However, our understanding of the geographic structure and evolutionary history of its wild populations is limited. To address this gap, we investigated genomic data from 284 samples of *R. tomentosa* from 28 wild populations in southern China.

**Results** The genetic diversity of populations in different regions revealed the similar trends using whole-genome and RAD-seq data, and Hainan Island having a higher genetic diversity than other regions. The 28 populations clustered into three distinct groups: (a) GROUP1 on the eastern mainland within Guangdong, Fujian, and Hunan Provinces; (b) GROUP2 on the western mainland within Guangxi and Yunnan Provinces; and (c) GROUP3 on Hainan Island. Mantel tests and redundancy analyses revealed population differentiation was affected by distance and environmental factors such as annual average radiation. Demographic history and gene flow analyses indicated the mainland populations and the Hainan Island populations diverged around 0.93 MYA, with gene flow primarily occurring from Hainan Island and the coastal regions (such as Zhanjiang in Guangdong and Fangchenggang in Guangxi) towards the mainland, reflecting an expansion trend within the species. PSMC' analyses indicated that the populations of the three groups underwent a bottleneck during the Pleistocene due to glacial-interglacial cycles and geological events. Niche analysis revealed that the ice ages caused habitat contraction for the species, and populations with higher genetic diversity are generally distributed in areas with more suitable habitats.

**Conclusions** This study elucidates the current genetic distribution of the species within China and suggests that drastic Pleistocene climate change and geographical events caused population divergence and fluctuations in effective population size, shaping the current genetic distribution of *R. tomentosa*. These findings provide a theoretical basis for the genetic conservation and improvement of *R. tomentosa*.

†Xing-ming Xu, Bo-yong Liao and Su-jiao Liao contributed equally to this work.

*Correspondence:
Bo-yong Liao
liaoby05@126.com
Yong-quan Li
yongquanli@zhku.edu.cn

Full list of author information is available at the end of the article

**Keywords** Population genetics, Divergence time, Evolutionary history, Gene flow, Pleistocene, Ecological niche

## Background

*Rhodomyrtus tomentosa* (Ait.) Hassk. (Myrtaceae) is a robust monoecious evergreen shrub distributed in East and Southeast Asia, including China, Japan, Thailand, the Philippines, Vietnam, and Malaysia [1]. In China, *R. tomentosa* predominantly inhabits regions between latitudes of 25°18'N and 26°18'N [2]. *Rhodomyrtus tomentosa* requires full sunlight and typically occurs along forest fringes [3]. For monoecious insect pollinated plants such as *R. tomentosa*, the reproduction mode varies between facultative outcrosses, cross-pollination, and same plant outcrosses [4].

*Rhodomyrtus tomentosa* is a versatile plant with several valuable attributes. Its aesthetic appeal makes it a popular choice for potted plants and bonsai owing to its exquisite color-changing flowers and vibrant red fruits [5]. In addition to its ornamental charm, *R. tomentosa* has a wide variety of medicinal benefits, including antimicrobial efficacy [6, 7], anti-tumor potential [8], and anti-inflammatory [9] and antioxidant properties [10]. *Rhodomyrtus tomentosa* also plays a major role in ecological landscapes, improving microhabitats in moderately degraded grassland ecosystems and creating a favorable environment for sandalwood seedlings [11, 12]. Recently, *R. tomentosa* has gained attention and was identified as one of the 240 "Neglected and Underutilized Crop Species" by the scientific project "Agrofolio" [10]. To further advance our understanding and utilization of *R. tomentosa*, it is essential to explore the genetic foundations of its distribution patterns and adaptive traits in wild populations.

Population genetics is an effective approach to investigate the genetic basis of distribution patterns and adaptive traits of species. Recent studies using various molecular markers (e.g., ISSR and SSR) have investigated genetic diversity and distribution patterns among populations in different regions [13–15]. However, due to the limitations of ISSR and SSR markers, past studies have not detailed the population history and current distribution patterns of *R. tomentosa*. Species distribution patterns result from the interaction of both external (e.g., historical climate, topography) and internal factors (e.g., biology, population history). External factors typically shape species geographical distributions [16–18], whereas internal factors often contribute to geographic structural diversity [19, 20]. Understanding how external factors, such as climate history, and internal factors, such as population history, affect species distribution patterns is crucial for predicting range shifts and guiding conservation efforts [21]. Therefore, conducting population genetics research on *R. tomentosa* using informative molecular markers over a broad geographical range is essential for understanding both the historical distribution and present genetic variation within the species.

Restriction site associated DNA sequencing (RAD-seq) is a simplified genome sequencing method developed using next-generation sequencing (NGS) technology [22–24]. Its high throughput, ease of use, and relatively low cost have made it a popular choice for genetic diversity analyses [25–28]. Despite these advantages, RAD-seq has limitations due to its reliance on restriction enzymes, which can pose challenges in applications. For example, in studies of evolutionary history over broader time scales, using chloroplast or low copy nuclear gene data can expand the range of outgroups and available fossil evidence, and can improve the accuracy of analyses. Furthermore, the low genome coverage of RAD-seq data renders its results unreliable for analyzing historical effective population size [29], whereas whole-genome sequencing (WGS) data meet the coverage and depth requirements necessary for this analysis. Therefore, combining RAD-seq datasets with other sequencing data, such as whole chloroplast genome and WGS data is a more efficient approach.

This study addresses three main questions: (a) What are the genetic diversity and structural distribution patterns of current *R. tomentosa* populations in southern China; (b) What are the impacts of geographical isolation and environmental factors on *R. tomentosa* population differentiation; and (c) How are the current genetic distribution and population differentiation of *R. tomentosa* related to climate oscillations, the formation of Hainan Island, and marine transgressions?

## Materials and methods

### DNA extraction and RAD libraries

Following field investigations in southern China, we selected 28 wild populations and sampled 10–20 plants (at distances exceeding 30 m) at each population location (Additional file 1: Table S1; Fig. 1C). The sample collection work was approved by the Forestry Bureau of Guangdong Province, China, and identified by Dr. Longyuan Wang from Zhongkai University of Agricultural Engineering. Two voucher specimens for each plant were retained and deposited separately in the Zhongkai University of Agricultural Engineering and Guangdong Eco-engineering Polytechnic herbaria (accession number: ZKU-LBY). The sampled leaves were dried directly with silica gel in the wild, then taken back to the lab and frozen in − 80 °C until DNA extraction.

Genomic DNA was extracted from leaf samples using a modified CTAB method [30]. Subsequently, DNA quality

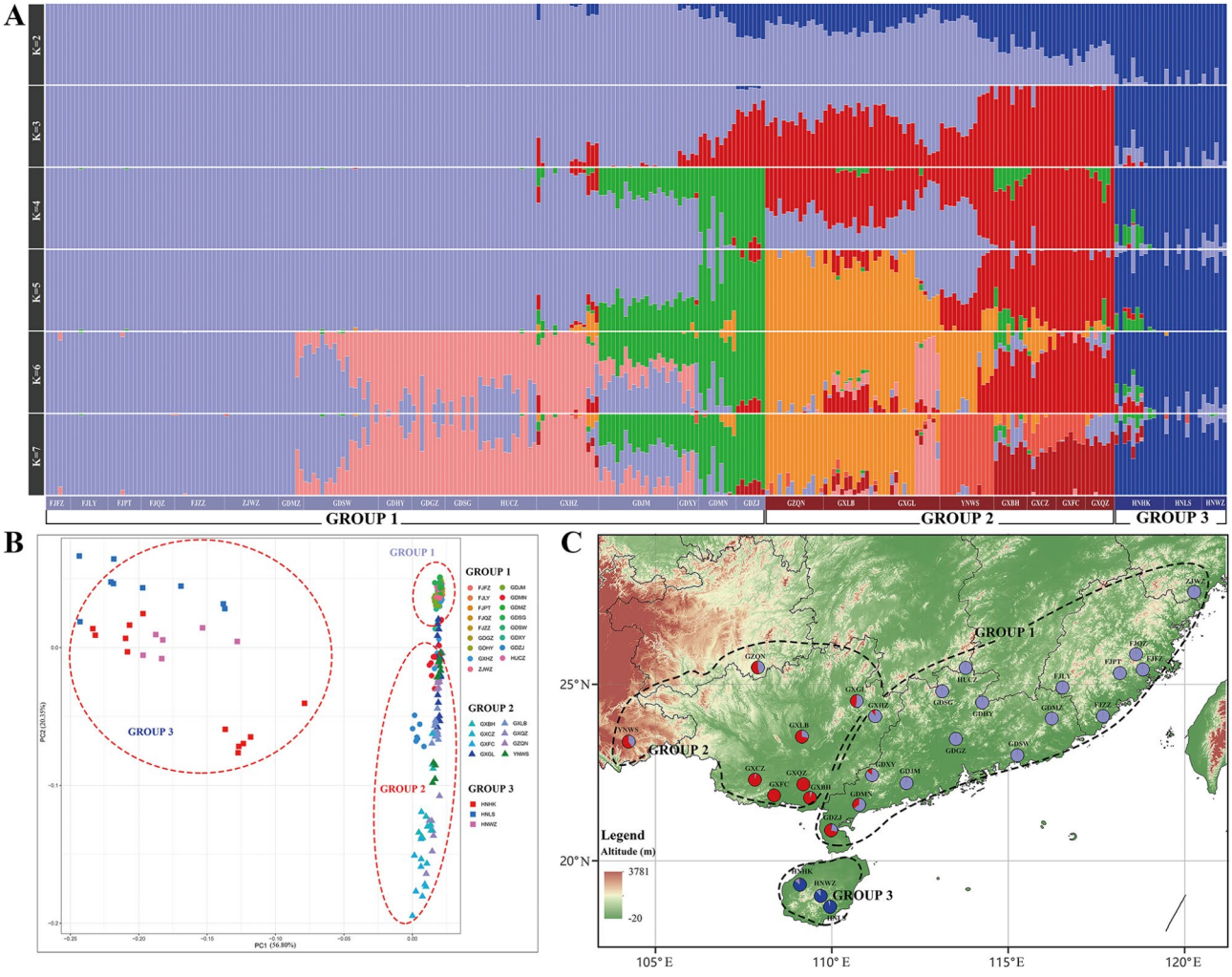


**Fig. 1** Patterns of the genetic structure of *Rhodomyrtus tomentosa*. **A** ADMIXTURE assignments for 28 populations. Model-based population assignments at K from 2 to 7. Each vertical bar represents a sample, with its assignment probability to genetic clusters represented by different colors. **B** Principal components analysis (PCA). Percent variation explained by each component is shown in parentheses. The color of the dots represents the population of the samples, while the shape of the dots indicates the group of populations. **C** The geographic distribution of the 28 populations of *R. tomentosa*. The pie charts depict the ancestry composition of each population for K = 3 as inferred by ADMIXTURE

and concentration were assessed using 1% agarose gel electrophoresis. DNA concentration was measured using a ThermoFish Qubit 4.0 fluorometer (Invitrogen, USA), and sample concentration was diluted to 150 ng/μl. Purified DNA was digested using *Mse I* and *EcoR I* enzymes (New England Biolabs, Beverly, USA). The fragmented DNA was ligated to unique paired-end adapters (P1) and a common adapter (P2). Unique P1 adapters were used to differentiate samples during downstream sequence analysis. Samples were run on 2% gels, and DNA in the range of 300–500 bp was excised and recovered using a GEL Extraction Kit (Omega, Norcross, USA). Purified products were amplified using PCR with forward and reverse amplification primers (IGEBIO, Guangzhou). The products were sequenced on an Illumina HiSeq 2500 platform (Jierui Biotech, Guangzhou), generating 135 bp paired-end reads. Full uncropped Gels and Blots images are provided in Additional file 3.

### SNP calling

After evaluation and quality control of the raw sequencing data using FastQC, we used the process_radtags tool from Stacks v.2.55 [31] to demultiplex the data into individuals. The demultiplexed data were aligned to the *R. tomentosa* reference genome (GenBank accession number: GCA_028455895.1) [32] using BWA-MEM v.0.7.17 [33]. The alignment results from the BWA-MEM were processed using SAMtools v.1.10 [34] and converted to BAM files. Quality assessment was performed using SAMtools flagstat (Additional file 1: Table S2).

To address various analytical requirements, different datasets were generated from the initial total single

nucleotide polymorphism (SNP) data using three different filtering strategies. For Stacks, the gstacks program assembled double-digested RAD loci for each locus from the alignment results, and the population program extracted and filtered SNPs. For the total SNP dataset, the Stacks population program was configured with the following settings: -r 0.7, to remove samples with missing data greater than 30%; --min-mac 10 to exclude SNPs with minor allele counts < 10; and --max-obs-het 0.7, to eliminate SNPs with observed heterozygosity > 0.7 across all populations; resulting in dataset (A) Dataset A was further filtered using VCFtools v.0.1.16 [35]. First, samples with > 20% missing data were removed using --max-missing 0.8. Next, SNPs with Hardy-Weinberg equilibrium (HWE) p-values < $10^{-7}$ within populations were filtered out using --hwe 0.0000001, and records with more than two alleles were excluded using --max-alleles 2, resulting in dataset (B) To minimize the impact of linkage disequilibrium (LD) on population structure inference, additional LD filtering was performed on dataset B using PLINK v.1.9 [36] with parameters set as --indep-pairwise 50 10 0.2, resulting in dataset C.

#### Shallow whole-genome sequencing, alignment, chloroplast genome assembly

From each population, 1–2 representative samples were selected (Additional file 1: Table S3). WGS was performed on the Illumina NovaSeq platform (provided by Guangzhou Jierui Biotechnology) with paired-end, 2 × 150 bp reads. Each sample generated approximately 6 Gb of clean data. Quality-controlled data were aligned to the *R. tomentosa* reference genome (GenBank accession number: GCA_028455895.1) using BWA-MEM and sorted into BAM files using SAMtools [34]. The alignment results from BWA-MEM were assessed separately for quality using SAMtools flagstat and SAMtools coverage (Additional file 1: Table S3). Chloroplast genome assembly was conducted with the GetOrganelle toolkit [37] using k-mer values of 21, 55, 85, and 115, using the published *R. tomentosa* chloroplast genome (NC_043848.1) as a random seed input. The filtered assembly results were manually inspected for accuracy and completeness using Bandage [38] before generating the final sequence.

#### Population genetic diversity and structure analysis

Population genetic diversity was calculated using the Stacks population pipeline for dataset A, although this may lead to biased estimates of population genetic parameters [39]. Dataset A was used to calculate the occurrence rate of private alleles (PA), average observed heterozygosity ($H_O$), average expected heterozygosity ($H_E$), nucleotide diversity (Π), and Wright's inbreeding coefficient ($F_{IS}$). Pairwise mean population genetic differentiation coefficient ($F_{ST}$) was computed using dataset B.

To assess population genetic structure of *R. tomentosa*, we used ADMIXTURE v.1.3.0 [40] in conjunction with principal component analysis (PCA) [41]. ADMIXTURE uses a maximum likelihood-based approach to assign ancestral allele frequencies across the entire genome. It calculates the CV-error for K = 1–10 using a cross-validation procedure and visualizes the genetic structure of various populations. PCA of the samples was conducted using PLINK. ADMIXTURE uses a likelihood model to estimate ancestral allele frequencies based on SNPs whereas PCA does not rely on biological assumptions, provides complementary insights with ADMIXTURE.

#### Validating population structure based on whole-genome sequencing data

First, the alignment results of the RAD-seq data for 30 samples were input into Stacks. The output results were filtered using VCFtools with the parameter "--maf 0.01" to select SNPs with a minor allele frequency greater than 0.01. Subsequently, the parameter "--max-missing 0.9" was applied to retain only those variants for which genotype information was available in 90% or more of the samples. Next, alignment results from the BWA-MEM of the WGS data were input into GATK v.4.3.0 [42] to extract SNPs. HaplotypeCaller was used to extract variant information for each sample individually, followed by merging the alignment results for the 30 samples. SelectVariants was used to extract SNPs from the variant information. In VariantFiltration, the criteria "QD < 2.0|| MQ < 40.0|| FS > 60.0|| SOR > 3.0|| MQRankSum < -12.5|| ReadPosRankSum < -8.0" were used to filter low-quality SNPs. The output results were filtered using VCFtools with filtering strategies consistent with the RAD-seq data.

Genetic diversity for each group was calculated using PLINK and the Hardy-Weinberg equilibrium was assessed. Finally, both SNP datasets, after LD filtering, were subjected to genetic structure analysis using ADMIXTURE.

#### Testing the impact of geographical distance and environmental differences on population differentiation

Isolation by Distance (IBD), Isolation by Environment (IBE), and Redundancy Analysis (RDA) can be used to elucidate the influence of geographical distance and environmental differences on population genetic differentiation. A genetic differentiation matrix (linearized pairwise genetic differentiation, I.E., $F_{ST}$ / (1 – $F_{ST}$)) was generated from the $F_{ST}$ matrix. Subsequently, the *distm* function from the *geosphere* package in R [43] was used to convert the latitude and longitude coordinates of each population (Additional file 1: Table S1) into a geographic distance matrix. To obtain the environmental distance matrix,

data for 19 bioclimatic factors (BIO1–19), three soil factors (soil organic matter, soil moisture content, soil pH), and four solar ultraviolet radiation factors (Uvb1–4) were retrieved from the WORLDCLIM (https://www.worldclim.org/), Center for Sustainability and the Global Environment (http://www.sage.wisc.edu/atlas/index.php/), and gIUV (http://www.ufz.de/gluv) database, respectively. Environmental data for each population were extracted using ArcGIS v.10.7 based on sampling locations (Additional file 1: Table S1). To mitigate the impact of multicollinearity, variables with a Variance Inflation Factor (VIF) > 10 were excluded as bioclimatic factors, resulting in the retention of four bioclimatic factors: BIO1, BIO2, BIO12, and BIO14. In total, 11 environmental factors were selected for further analysis (Table 1). In R, the *dist* function from the *vegan* package [44] was used to calculate the Euclidean distance of standardized environmental factors, which served as the environmental distance matrix for 28 *R. tomentosa* populations. After obtaining these three matrices, the IBD and IBE Mantel tests were conducted using GENALEX v.6.5 [45], with 999 permutations.

RDA was used to estimate the percentage of genetic variation attributed to the collinear portions of each environmental factor. The *pcnm* r function was used to convert the pairwise genetic matrix into the principal coordinates of neighboring matrices [46]. The *rda* function in the R *vegan* package was used to perform redundancy analysis of the 11 environmental factors and genetic distance, and the results were visualized using *ggplot2* package.

#### Demographic history analysis

The population history of *R. tomentosa* consists of two key components, phylogenetic history and gene flow events. To reconstruct this history, we employed: (1) BEAST molecular clock analyses using chloroplast whole-genome sequences; (2) Pairwise Sequentially Markovian Coalescent (PSMC') modeling using shallow WGS data to estimate effective population sizes; (3) DIYABC analysis to verify the evolutionary history; and (4) TreeMix and (5) Dsuite analyses, to elucidate the gene flow events among the 28 distinct *R. tomentosa* natural populations.

**Table 1** The environmental factors used to construct the environmental distance matrix and conducting ecological niche predictions

| NO | Code | Variable types | Environment variable |
|---|---|---|---|
| 1 | BIO1 | Bioclimatic factors | Annual Mean Temperature |
| 2 | BIO2 | Bioclimatic factors | Mean Diurnal Range |
| 3 | BIO12 | Bioclimatic factors | Annual Precipitation |
| 4 | BIO14 | Bioclimatic factors | Precipitation of Driest Month |
| 5 | Soil1 | Soil factors | Soil pH |
| 6 | Soil2 | Soil factors | Soil organic matter |
| 7 | Soil3 | Soil factors | Soil moisture content |
| 8 | Uvb1 | Radiation factors | Annual Mean UV-B |
| 9 | Uvb2 | Radiation factors | UV-B Seasonality |
| 10 | Uvb3 | Radiation factors | Mean UV-B of Highest Month |
| 11 | Uvb4 | Radiation factors | Mean UV-B of Lowest Month |

First, we combined the 30 assembled chloroplast sequences with 34 additional chloroplast whole-genome sequences from Myrtaceae taxa accessible on NCBI (Additional file 1: Table S4). These sequences were aligned using MAFFT v.7.520 [47], examined using MEGA X [48], and converted to FASTA format. Subsequently, using the ModeFinder tool within IQ-TREE v.1.4.4 [49] the GTR + F + R5 model was determined to be the best-fitting DNA substitution model, with a Bayesian Information Criterion (BIC) score of 903836.022. Divergence time analysis was performed with BEAST v.2.4 [50] using the Markov chain Monte Carlo (MCMC) method. The BEAUti settings included a site model with the nucleotide substitution model set to GTR + F + R5, determined by ModeFinder. The Clock model was set to Strict Clock, and the speciation model was set to the Yule Model. For Myrtoideae stem age, we used a normal distribution with a mean of 85.0, Sigma of 1.5, and Offset of 5.0, based on fossil records for *Sterculia* (http://www.malvaceae.info/Classification/ Sterculioideae.html). Additionally, based on fossil records from the Timetree website (https://timetree.org/home), we set the time mean for *Corymbia* and *Angophora* to 2.0, with a Sigma of 1.0 and an Offset of 0.5. In BEAST, the MCMC chain length was set to 10,000,000 and samples were taken every 10,000 generations. The chain stability and convergence were assessed using Tracer 1.5 (beast.community/tracer). A consensus tree was generated using TreeAnnotator after discarding the initial 10% of the tree.

The PSMC' model [51] was used to infer changes in the historical effective population size from genomic sequences. Initially, the mpileup command from BCFtools v.1.17 [52] and bamCaller.py (https://github.com/stschiff/msmc-tools) were used to generate VCF and mask files from BAM files (Additional file 1: Table S3) with default parameters. This process masked the gaps, repetitive sequences, ad insertions in the genome. Subsequently, generate_multihetsep.py (https://github.com/stschiff/msmc-tools) generated input files for MSMC2, which analyzed the effective population size for each pair of population genomes. The grouping was as follows: GDMZ, GDSG, and FJZZ represented GROUP1; GXFC, GXLB, and GZQN represented GROUP2; HNWZ, HNWS, and HNHK represented GROUP3 (Additional file 1: Table S3) with default parameters. The final output results were plotted using R. The parameters were scaled assuming a generation time (g) of five years and a mutation rate (μ) of $1 \times 10^{-8}$ per generation per site [53]. The

effective population size and time were represented by θ/(4μ) and d/(2μ/g), respectively.

Finally, the population history was investigated using approximate Bayesian computations in DIYABC v2.1.0 [54] based on SNP dataset C. Scenarios 1–3 were assumed to verify the results of the divergence history analysis. The prior information for the assumed population effective size (Ne) and time (t) was set according to the results from BEAST and PSMC'. The generation time was assumed to be g = 5, with t2 ranging from 150,000 to 200,000 and t1 ranging from 1,500 to 2,000. All Ne values ranged from 10,000 to 100,000, with t2 being greater than or equal to t1. The admixture rate (r) ranged from 0.001 to 0.999. According to the software's recommendation, the analysis was performed with 400,000 simulations.

TreeMix v.1.13 [55] was used to test the gene flow in a phylogenetic context. The VCF file was converted to a frequency file that could be transformed into a TreeMix file using the plink2treemix script (https://github.com/barbatom/plink2treemix). TreeMix simultaneously estimated an ML species tree and the direction and weight (w) of gene flow among populations and groups based on allele frequencies. Initially, an ML species tree without migration was built, and migration events were sequentially added until the model's explanatory power no longer increased significantly.

Dsuite [56] software, based on SNP dataset C, infers Patterson's D statistic, also known as the ABBA-BABA statistic, often referred to as the *f*4 ratio statistic. In this test, an outgroup was included for comparison. The analysis involved four populations, distinguishing between ancestral ("A") and derived ("B") alleles, and quantified the patterns known as "ABBA" and "BABA." Under the assumption of a phylogenetic tree with no gene flow, SNPs counts displaying "ABBA" and "BABA" patterns should be equally frequent. An excess of either pattern in the genome suggests a potential gene flow between different populations, which can be calculated using Patterson's D statistic. For this analysis, the maximum likelihood tree inferred with zero gene flow (m = 0) events using TreeMix was used as the input tree.

**Table 2** Usage of different sequencing data sets

| Data sets | | Usage |
|---|---|---|
| RAD-seq | A | Genetic diversity |
| | B | IBD, IBE, RDA |
| | C | Population genetic structure, Gene flow |
| Whole-genome sequencing | | Genetic diversity, Chloroplast Genome Assembly, PSMC' |
| chloroplast whole-genome sequence | | BEAST2 |

### Niche model construction

After merging the geographic coordinates of *R. tomentosa* obtained from the Chinese Virtual Herbarium (CVH, https://www.cvh.ac.cn/) and the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/) with the geographic coordinates of 28 populations (Additional file 1: Table S1), they were imported into ArcGIS for distance calculation. Points that were too close were manually removed, leaving 174 specimen coordinates for analysis. The four bioclimatic variables used to construct the environmental difference matrix (Table 1) were also used as predictive factors for the potential habitat of *R. tomentosa*. The geographic resolution of the environmental raster dataset was standardized to 2.5' (approximately 4.5 $km^2$).

For habitat prediction, geographic coordinates and environmental raster data were imported into MaxEnt v.3.4.4 [57]. Each model was run ten times, using 75% of available records for training and 25% for validation. All other parameters were set to their default values. Predictive accuracy was assessed using Receiver Operating Characteristic (ROC) analysis. The Area Under the Curve (AUC) was used as an indicator of model prediction accuracy, with models showing AUC values close to 1 considered more accurate [58]. These were classified (0.5–0.6), poor (0.6–0.7), fair (0.7–0.8), good (0.8–0.9), and excellent (0.9–1.0) [59, 60]. The average of ten output asc files was used for subsequent analyses. Additionally, we collected four bioclimatic variables from the Mid-Holocene (MH, approximately 6,000 years BP), the Last Glacial Maximum (LGM, approximately 22,000 years BP), and Last Interglacial (LIG, approximately 140,000 years BP) periods in the WORLDCLIM bioclimatic database (https://www.worldclim.org/data/v1.4/paleo1.4.html) to simulate the distribution of *R. tomentosa* during different historical periods. The output of the suitability distribution raster data was imported into ArcGIS for mapping, yielding four levels of potential habitat suitability: excellent (> 0.6), good (0.4–0.6), fair (0.2–0.4), and poor (< 0.2) suitability [60, 61].

## Results

### Sequencing data preprocessing and snp calling

After initial filtering, RAD-seq data from 284 samples representing 28 *R. tomentosa* populations were retained for subsequent analysis (Additional file 1: Table S2). The alignment results showed an average of 3,285,192 clean reads with a 96.47% average alignment rate and 4.86% average genome coverage. Three SNP datasets were generated based on the alignment results (Table 2): dataset A contained 220,734 SNPs for genetic diversity analysis; dataset B contained 13,229 SNPs for calculating FST, Mantel tests, and RDA; and dataset C contained 3,021 SNPs for genetic structure and gene flow analysis. The

WGS alignment results showed a total of 47,660,548 average clean reads, with a 93.98% average alignment rate, 90.5% average genome coverage, and an average depth of 13.95× (Additional file 1: Table S3). The nucleotide composition analysis of the 30 successfully assembled complete chloroplast genome sequences (Additional file 1: Table S5) revealed an average sequence length of 156,153 base pairs, with an A + T content of 62.8% and a C + G content of 37.1%.

#### Population genetic diversity and structure analysis

The genetic diversity results from Stacks (using Variant positions) indicated that the number of PA across all populations ranged from 45 (GDMZ) to 5,323 (GDZJ) (Table 3). $H_O$ varied from 0.0222 (ZJWZ) to 0.1924 (GDZJ), while $H_E$ ranged from 0.0705 (FJQZ) to 0.2535 (HNHK). Π ranged from 0.0759 (FJQZ) to 0.2665 (HNHK), and the $F_{IS}$ varied from 0.0282 (FJLY) to 0.3155 (GXGL). The GDZJ population in the Guangdong Leizhou Peninsula and the HNHK population in northern Hainan Island had the highest number of PA, $H_O$, $H_E$, and Π. Populations with the highest $F_{IS}$ included highland populations GXGL, YNWS, and GZQN, and island populations HNLS and HNWZ. Populations with higher genetic diversity indices were mainly found near Beibu Gulf in northern Guangxi Province, Leizhou Peninsula in Guangdong Province, and Hainan Island. Other regions had relatively lower genetic diversity indices, with Fujian and Guangdong having lower indices than those of Guangxi. Populations in mountainous and island areas had higher $F_{IS}$ values. The pairwise genetic differentiation coefficient $F_{ST}$ (Additional file 2: Fig. S1), and the highest value was observed between ZJWZ and HNWZ, reaching 0.23. Most $F_{ST}$ values between the populations ranged from 0.05 to 0.15.

The cross-validation error rates for the number of ancestral populations (K = 1–10), indicated no significant differences between K = 3 and K = 9 (Additional file 2: Fig. S2). PCA further supported the grouping strategy at K = 3 (Fig. 1B). PC1 (56.80%) differentiated the Hainan Island populations from the mainland populations, whereas PC2 (20.35%) separated the Guangxi Province population from those in Guangdong and Fujian Provinces. For K = 3–9, the PCA results consistently supported K = 3 (Fig. 1A and B). Based on ADMIXTURE, PCA, and population geographic distributions, the 28 populations were

**Table 3** Population genetic diversity measures for *Rhodomyrtus tomentosa* populations based on 220,734 polymorphic sites

| Pop ID | Variant positions | | | | | All positions (variant and fixed) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | PA | $H_O$ | $H_E$ | Π | $F_{IS}$ | $H_O$ | $H_E$ | Π | $F_{IS}$ |
| FJFZ | 116 | 0.0607 | 0.0753 | 0.0828 | 0.0500 | 0.0004 | 0.0004 | 0.0005 | 0.0003 |
| FJLY | 316 | 0.0778 | 0.0836 | 0.0889 | 0.0282 | 0.0005 | 0.0006 | 0.0006 | 0.0002 |
| FJPT | 862 | 0.0742 | 0.0855 | 0.0924 | 0.0455 | 0.0003 | 0.0004 | 0.0004 | 0.0002 |
| FJQZ | 434 | 0.0544 | 0.0705 | 0.0759 | 0.0528 | 0.0003 | 0.0004 | 0.0004 | 0.0003 |
| FJZZ | 92 | 0.0739 | 0.0983 | 0.1033 | 0.0752 | 0.0004 | 0.0005 | 0.0005 | 0.0004 |
| GDGZ | 711 | 0.1310 | 0.1556 | 0.1677 | 0.0852 | 0.0007 | 0.0008 | 0.0009 | 0.0004 |
| GDHY | 1218 | 0.1174 | 0.1287 | 0.1383 | 0.0516 | 0.0007 | 0.0007 | 0.0008 | 0.0003 |
| GDJM | 181 | 0.1031 | 0.1397 | 0.1442 | 0.1222 | 0.0004 | 0.0006 | 0.0006 | 0.0005 |
| GDMN | 235 | 0.1065 | 0.1470 | 0.1573 | 0.1315 | 0.0005 | 0.0007 | 0.0008 | 0.0007 |
| GDMZ | 45 | 0.0727 | 0.1125 | 0.1242 | 0.1149 | 0.0004 | 0.0007 | 0.0007 | 0.0007 |
| GDSG | 931 | 0.0901 | 0.1246 | 0.1346 | 0.1032 | 0.0005 | 0.0007 | 0.0007 | 0.0006 |
| GDSW | 75 | 0.1039 | 0.1541 | 0.1596 | 0.1505 | 0.0003 | 0.0005 | 0.0005 | 0.0005 |
| GDXY | 130 | 0.0569 | 0.1544 | 0.1748 | 0.2509 | 0.0002 | 0.0004 | 0.0005 | 0.0007 |
| GDZJ | 5323 | 0.1924 | 0.2011 | 0.2184 | 0.0625 | 0.0011 | 0.0012 | 0.0013 | 0.0004 |
| GXBH | 97 | 0.1415 | 0.2048 | 0.2218 | 0.1857 | 0.0004 | 0.0006 | 0.0007 | 0.0006 |
| GXCZ | 1337 | 0.1562 | 0.1942 | 0.2121 | 0.1304 | 0.0009 | 0.0011 | 0.0012 | 0.0007 |
| GXFC | 664 | 0.1712 | 0.2044 | 0.2230 | 0.1164 | 0.0009 | 0.0011 | 0.0012 | 0.0006 |
| GXGL | 342 | 0.0968 | 0.2103 | 0.2179 | 0.3155 | 0.0006 | 0.0012 | 0.0013 | 0.0019 |
| GXHZ | 369 | 0.0973 | 0.1697 | 0.1766 | 0.2218 | 0.0005 | 0.0009 | 0.0009 | 0.0011 |
| GXLB | 1236 | 0.1692 | 0.2289 | 0.2422 | 0.1858 | 0.0010 | 0.0013 | 0.0014 | 0.0011 |
| GXQZ | 276 | 0.1298 | 0.1906 | 0.2093 | 0.1805 | 0.0006 | 0.0008 | 0.0009 | 0.0008 |
| GZQN | 363 | 0.0787 | 0.1612 | 0.1682 | 0.2408 | 0.0004 | 0.0008 | 0.0008 | 0.0012 |
| HNHK | 4389 | 0.1744 | 0.2535 | 0.2665 | 0.2256 | 0.0010 | 0.0014 | 0.0015 | 0.0013 |
| HNLS | 264 | 0.1138 | 0.2145 | 0.2299 | 0.2770 | 0.0003 | 0.0005 | 0.0005 | 0.0006 |
| HNWZ | 2591 | 0.1547 | 0.2130 | 0.2348 | 0.1765 | 0.0007 | 0.0010 | 0.0011 | 0.0008 |
| HUCZ | 138 | 0.0959 | 0.1521 | 0.1590 | 0.1606 | 0.0005 | 0.0008 | 0.0008 | 0.0008 |
| YNWS | 1048 | 0.0504 | 0.1149 | 0.1200 | 0.2185 | 0.0003 | 0.0006 | 0.0007 | 0.0012 |
| ZJWZ | 134 | 0.0222 | 0.0921 | 0.0963 | 0.1652 | 0.0001 | 0.0004 | 0.0004 | 0.0007 |

classified into three groups (Fig. 1A and C). GROUP1, located in the eastern mainland, included 17 populations distributed in the Guangdong, Fujian, Zhejiang, and Hunan provinces. GROUP2, in the western mainland, consisted of eight populations distributed in the Guangxi, Yunnan, and Guizhou provinces. GROUP3 comprised three populations on Hainan Island. The populations at the distribution boundaries of the three groups exhibited mixed genetic components. Substructures appeared at K = 4–7 (Fig. 1A). At K = 4, the GDJM, GDXY, GDZJ, and GDMN populations, located at the boundary between GROUP1 and GROUP2, were separated from GROUP1. At K = 5, the populations in the southern and northern parts of the GROUP2 distribution area were also differentiated. Finally, at K = 7, the YNWS population was separated from GROUP2.

We further validated the results of this analysis using WGS data. In genetic diversity analyses, WGS data retained 10,521,071 SNPs, while the RAD-seq data retained 47,524 SNPs. In genetic structure analyses WGS data retained 1,661,223 SNPs, and RAD-seq data retained 13,476 SNPs. The genetic diversity analysis of WGS data (Additional file 2: Fig. S3) showed that $H_O$ ranged from 0.071 (GROUP1) to 0.188 (GROUP3) and $H_E$ ranged from 0.145 (GROUP1) to 0.200 (GROUP3). In RAD-seq data, $H_O$ ranged from 0.068 (GROUP1) to 0.102 (GROUP3) and $H_E$ from 0.130 (GROUP1) to 0.186 (GROUP2) (Additional file 2: Fig. S3). The genetic structure analysis (Additional file 2: Fig. S4) showed that the cross-validation error rates reached a minimum at K = 2, differing from the results obtained from 284 samples (Additional file 2: Fig. S2). This discrepancy may be due to the filtering strategy for the 30 samples, which resulted in the deletion of population-specific SNPs. The genetic structure results indicate that at K = 2–5, both the sequencing data types yielded similar results for the same samples. These results indicate that although RAD-seq data only capture a subset of genetic variation, they demonstrate results nearly identical to the WGS data, further supporting the reliability of our results.

#### Environmental and geographical effect on genetic variance

Genetic distance, geographic distance ($R^2 = 0.1808$, $P = 0.001$), and environmental distance ($R^2 = 0.0514$, $P = 0.035$) were significantly correlated across all populations (Fig. 2A and B). In the RDA (Fig. 2C), RDA1 (19.667%) and RDA2 (18.86%) accounted for the majority of genetic distance variation related to the predictor variables. Among the 11 environmental factors (Table 1), BIO1 (Annual Mean Temperature), BIO14 (Precipitation of Driest Month), Uvb1 (Annual Mean UV-B), Uvb4 (Mean UV-B of Lowest Month), and Soil1 (Soil pH) most strongly explained the genetic distance effects. Notably, Uvb1 appears to explain the differentiation between GROUP3 and the other populations quite well.

#### Demographic history

The consensus tree generated using BEAST (Fig. 3A) revealed that the 28 *R. tomentosa* populations can be categorized into three groups. Clade 3 was diverged 0.93 (0.78–1.10, 95% HPD) million years ago (MYA), while Clades 1 and 2 began to diversify around 0.09 (0.06–0.13, 95% HPD) MYA. With a few exceptions, grouping in the chloroplast tree aligned closely with the population divisions inferred from the ADMIXTURE analysis (Fig. 2A).

The PSMC' method successfully inferred the historical effective population sizes over the past 5 MYA (Fig. 3B and C, and 3D). All three groups experienced bottlenecks during the Pleistocene. Around 1.0–0.9 MYA, the effective size of all populations expanded to their maximum, with GROUP3 expanding to $9 \times 10^4$, while GROUP1 and GROUP2 expanded to $11 \times 10^4$. However, from 0.9 to 0.05 MYA, there was a population contraction phase, with all three groups shrinking to $1 \times 10^4$, followed by expansion again. Comparing the historical effective populations of GROUP1 and GROUP2 (Fig. 3B), a divergence in effective population sizes occurred around 0.2 to 0.1 MYA. Comparing GROUP3 with GROUP2 (Fig. 3D) and GROUP1 (Fig. 3C), it was found that GROUP3 experienced a contraction earlier, around 1.0 MYA, compared to the other groups.

The DIYABC results (Fig. 4D) indicate that Scenario 1 (Fig. 4A) has the best performance. Under this scenario, the population history includes two divergence events: GROUP3 emerged first, followed by the divergence of GROUP1 and GROUP2. The divergence timeline for Scenario 1 shows (Additional file 1: Table S6) that the first divergence (t2) occurred 152,000 generations ago, approximately 760,000 years ago, with t1 occurring 10,000 years ago.

In the TreeMix analysis, as the number of gene flow events increased from 0 to 8, the explanatory power of the model increased from 92.10 to 97.44%. Therefore, the model with m = 8 was selected for presentation. Among the eight gene flow events (Fig. 3E and F), events ①, ②, ③, ④, ⑤, and ⑧ were directed from coastal regions toward inland areas, with the strongest gene flow originating from GROUP3 toward GDZJ, with a migration weight of 0.54. Gene flow was also observed from Guangxi toward Hainan Island (⑥) and from eastern to western mainland regions (⑦).

The Dsuite analysis results (Fig. 3G) indicated that within GROUP1, the gene flow was relatively lower, being the strongest between GDJM and GDMN, similar to the results from TreeMix. Some Hybridization events were detected among the GDZJ, GZQN, and YNWS populations in GROUP1. And the HNHK population

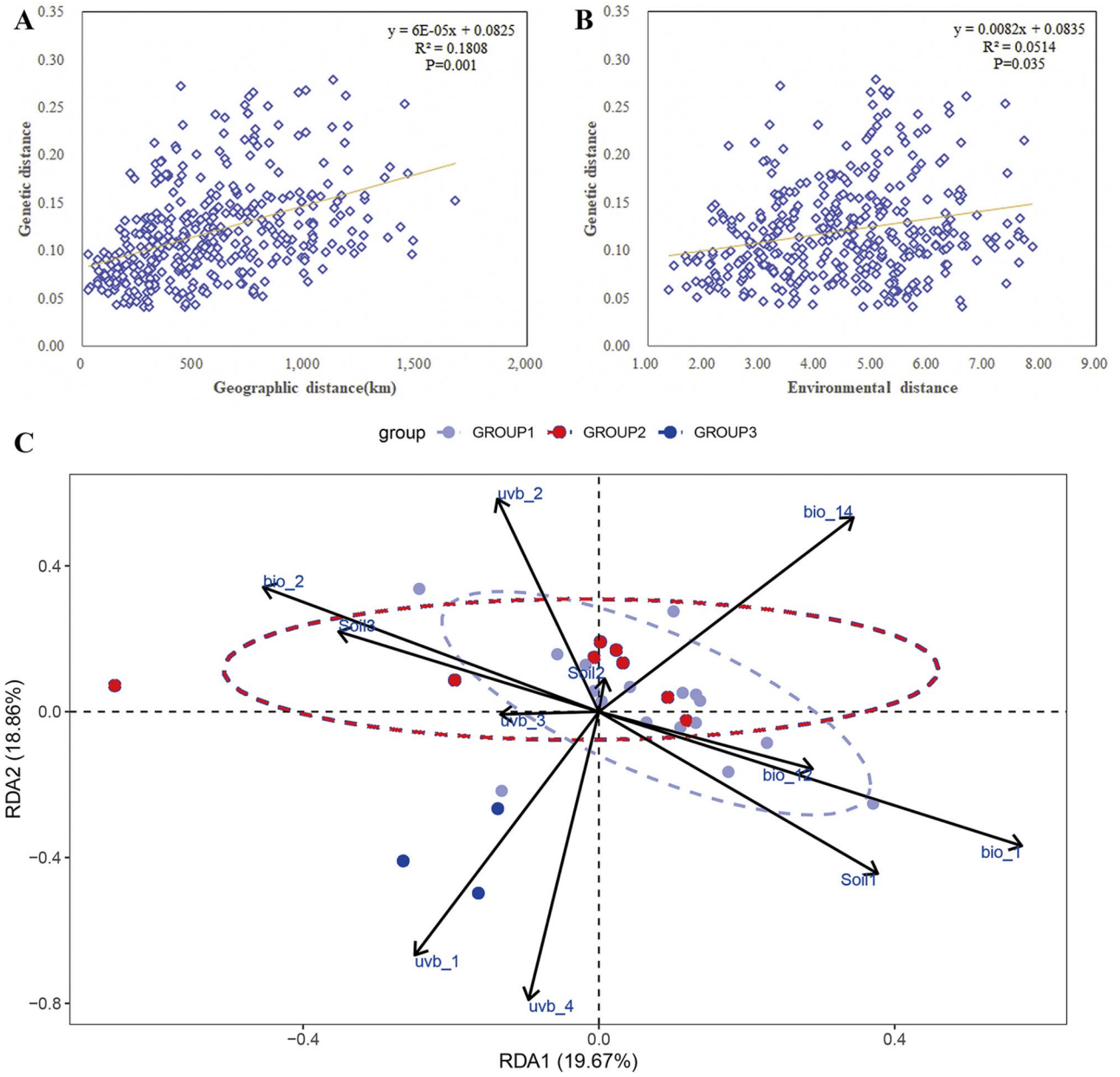


**Fig. 2** Environmental and geographical effect on the genetic variance of *Rhodomyrtus tomentosa*. **A** Correlation between geographical distances (km) and genetic distances computed as $F_{ST}$ / (1 − $F_{ST}$) between site pairs. **B** Correlation between environmental distances and genetic distances computed as $F_{ST}$ / (1 − $F_{ST}$) between site pairs. **C** Redundancy analysis (RDA) results show the distribution of the 28 populations of *R. tomentosa* on RDA axes 1 and 2. The vectors are the environmental variables. The relative arrangement of individuals and variables in the ordination space reflects their relationship with the ordination axes, which are linear combinations of the predictor variables

has hybridized with almost all populations, with a stronger effect of gene flow on GROUP2 compared to that on GROUP1.

### Ecological niche modeling

Figures S5A (Additional file 2) displays the test omission rate and predicted area, averaged over 10 runs, against the cumulative threshold. The black line in the graph represents the predicted false negative rate within the average range. The light blue line indicates the false negative rate for model training samples, showing consistency with predicted rates for the test set. Results suggest that the model adapts well to both training and unique test data. The average training AUC for 10 repetitions of ecological niche modeling under current climatic conditions (Additional file 2: Fig. S5B) is 0.966 with a standard deviation of 0.004, indicating excellent predictive accuracy of the model, making it suitable for identifying

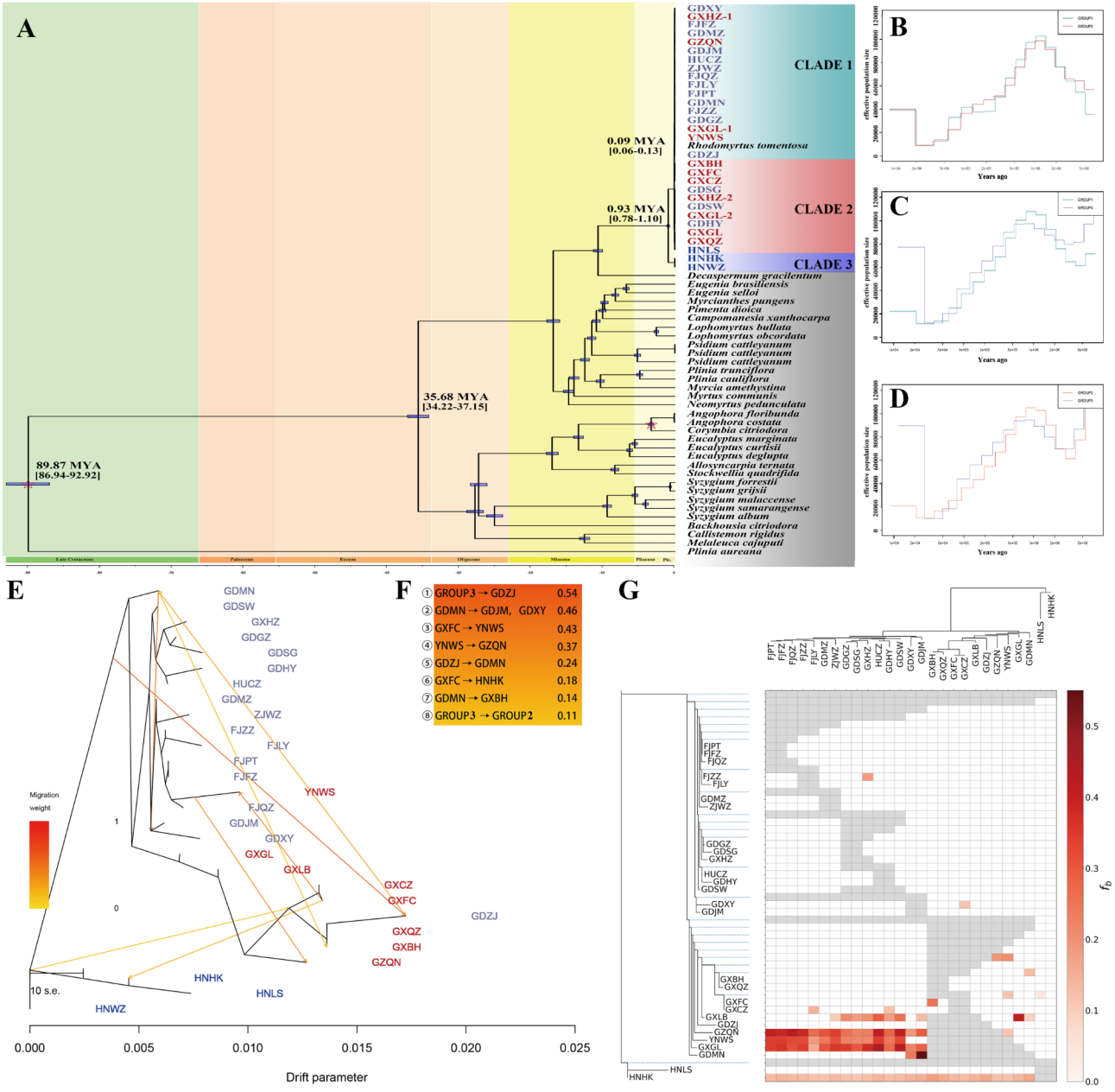


**Fig. 3** Evolutionary history and gene flow of *Rhodomyrtus tomentosa* populations. **A** BEAST time tree with node heights scaled to median divergence time estimates. Black numbers indicate the posterior mean age for each node. Blue bars represent the 95% highest posterior density (HPD) intervals. Sample names are colored according to the results of the genetic structure analysis (Fig. 1A). **B–D** Population size history inferred by PSMC′ using the three whole-genome sequenced individuals from GROUPs 1, 2, and 3. **E** TreeMix diagram illustrates the relationships and eight putative migration events among 28 populations of *R. tomentosa*. In the phylogenetic tree, HNWZ located at the root is designated as an outgroup (Fig. 3A). **F** Summarize the direction and migration weight of gene flow across populations (Fig. 3E). **G** The $f_b$ statistic (summary of $f_4$ admixture ratios). Gray color corresponds to tests that are not possible because of constraints on the phylogeny

potential habitat for *R. tomentosa* in China under current conditions.

Under the current climatic conditions (Fig. 5A), the potential distribution of *R. tomentosa* covers many southern regions of mainland China, including the Guangdong, Guangxi, Hainan, and Fujian provinces, as well as parts of Zhejiang, Jiangxi, Hunan, Guizhou, and Yunnan (Fig. 5A). The excellent suitable area covers 8,782 grid cells, while the good suitable area covers 10,997 grid cells (Additional file 2: Fig. S6). In the identified three time slices, the area with excellents and good suitability followed the order: MH > LIG > LGM, which is consistent with the temperature order of these periods. This indicates that during historical cold periods, the suitable

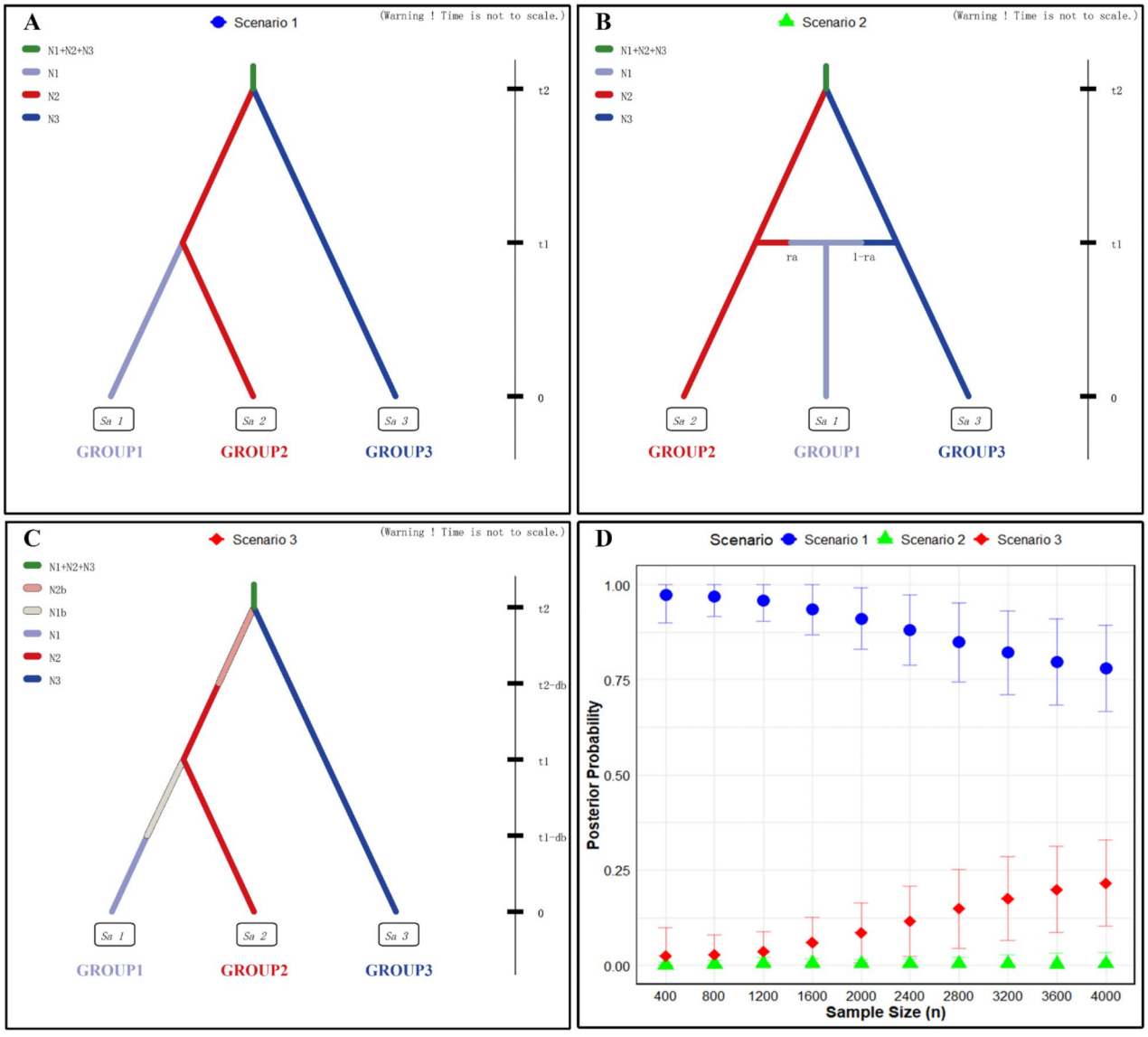


**Fig. 4** Demographic scenarios designed for DIYABC analysis. **A**–**C** represent five possible demographic scenarios, and **D** shows the logistic regression results of posterior probabilities, indicating that the third scenario is the most suitable

habitat of the species retreated southwards to areas near Hainan Island and the Leizhou Peninsula (Fig. 5B, C, and D).

## Discussion

### Genetic diversity

Discerning genetic diversity within populations is crucial for evaluating the evolutionary potential and environmental adaptability of species and is regarded as one of the most important parameters for prioritizing conservation genetics research [62]. Unlike traditional molecular markers such as SSRs, SNP markers offer higher density and a more uniform distribution, enabling a more comprehensive and accurate understanding of species genetic diversity [63]. For the first time, we assessed the genetic diversity of natural *R. tomentosa* populations distributed in China using a large-scale genomic SNP dataset. Genetic diversity varied significantly among populations in different distribution areas, with $H_O$ ranging from 0.0222 to 0.1924, $H_E$ ranging from 0.0705 to 0.2535, and Π ranging from 0.0759 to 0.2665. Populations in regions such as the Leizhou Peninsula (Guangdong), Beibu Gulf (Guangxi), and northern Hainan Island exhibited the highest levels of genetic diversity ($H_O$ > 0.1500, $H_E$ > 0.2000, Π > 0.2200). Conversely, populations in eastern Guangdong, Fujian, Zhejiang, and Hunan displayed lower genetic diversity ($H_O$ < 0.1000, $H_E$ < 0.1000, and Π < 0.1000). Compared with other studies that used

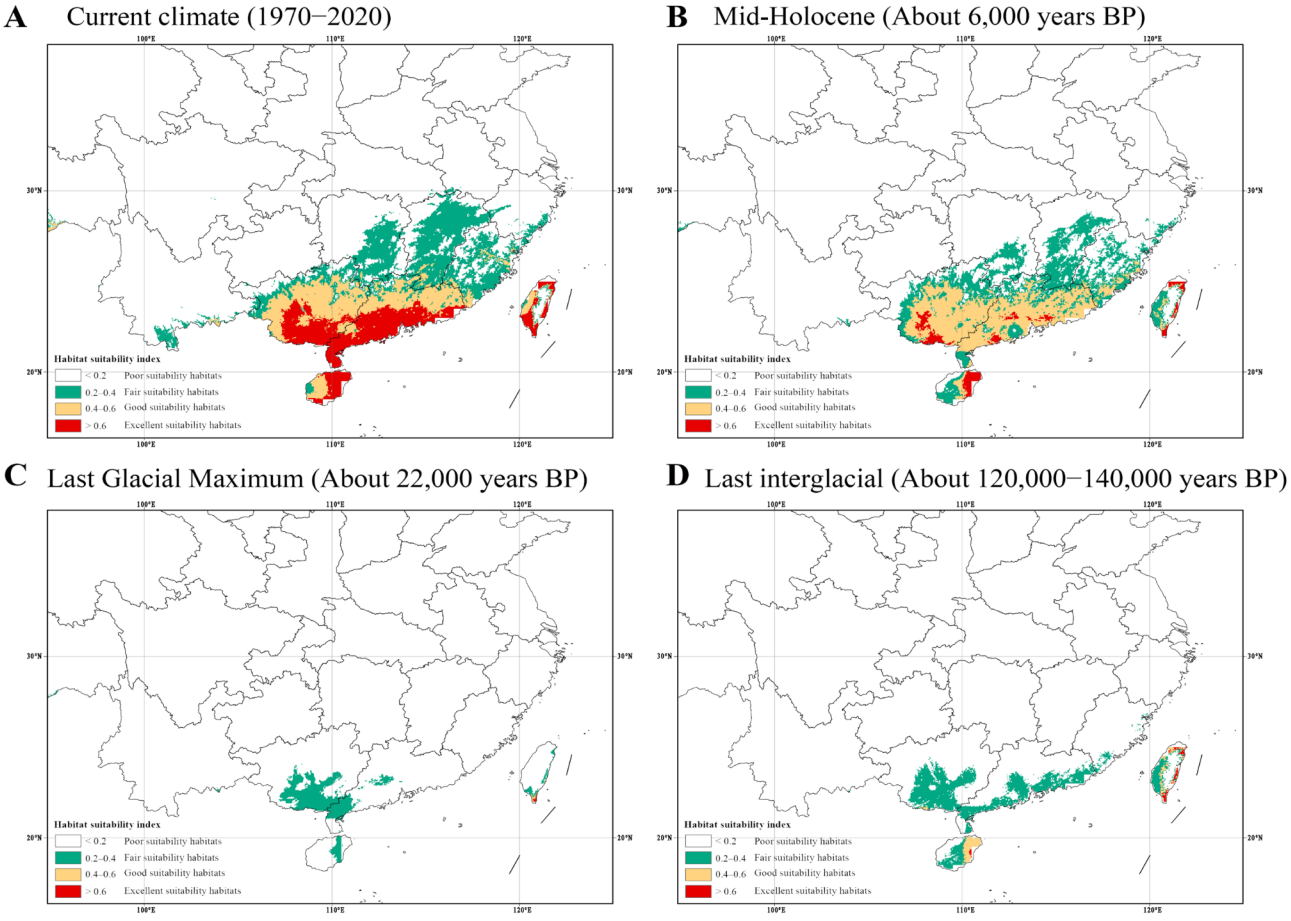


**Fig. 5** Species distribution models for *Rhodomyrtus tomentosa* based on ecological niche modeling using MaxEnt. **A** Current Distribution of Suitable Habitats. **B** Mid-Holocene (About 6000 years BP) Distribution of Suitable Habitats. **C** Last Glacial Maximum (About 22,000 years BP) Distribution of Suitable Habitats. **D** Last interglacial (About 120,000–140,000 years BP) Distribution of Suitable Habitats. 0.000–1.000 marked different colors: predicated distribution probabilities as logistic values. The map was created by ArcGIS and MaxEnt, and the base map is provided by Standard Map Service System (http://bzdt.ch.mnr.gov.cn; No. GS(2020)4619)

Stacks to calculate genetic diversity, the genetic diversity of *R. tomentosa* is lower than that of other common shrub species, such as *Camellia tunghinensis* ($H_O$ = 0.253–0.297, $H_E$ = 0.275–0.277, Π = 0.227–0.280) [64], and ($H_O$ = 0.058–0.111, Π = 0.204–0.234) [65], as well as *Corylus avellana* ($H_O$ = 0.217–0.326, $H_E$ = 0.113–0.290, Π = 0.195–0.316) [66]. However, it is higher than that of endangered species, such as *Rhododendron meddianum* ($H_O$ = 0.0440–0.0578, $H_E$ = 0.0646–0.0781, Π = 0.0671–0.0813) [67] and *Glyptostrobus pensilis* ($H_O$ = 0.0754–0.1008, $H_E$ = 0.0263–0.0466, Π = 0.0634–0.1010) [68]. Geographically, the center of genetic diversity for the species in China is located in coastal regions of Hainan Island and Guangxi (around Beibu Gulf), while lower genetic diversity, comparable to that of endangered species, is found in the eastern coastal regions of Zhejiang and Fujian. These results provide a foundation for the genetic conservation of *R. tomentosa*.

Previously, genetic diversity calculations based on SSR markers revealed expected heterozygosity ranging from 0.63 to 0.85 in 12 wild *R. tomentosa* populations in Thailand [13]. In contrast, 16 wild *R. tomentosa* populations in Guangxi, China, exhibited $H_O$ and $H_E$ ranging from 0.000 to 0.563 and 0.121–0.667, with average values of 0.240 and 0.414, respectively [15]. Furthermore, based on the ISSR markers, Malaysian populations displayed low levels of genetic diversity ($H_S$ = 0.0073; I = 0.1085; PPB = 20.14%) [14]. Overall, *R. tomentosa* populations from different geographical regions differed significantly, aligning with our findings. Although genetic diversity results calculated using different molecular markers are generally challenging to compare directly, previous research indicates that SSR markers estimate heterozygosity at a rate two to three times higher than SNP markers [69, 70]. Therefore, upon further comparison, we found that the Leizhou Peninsula, Beibu Gulf, and northern Hainan Island *R. tomentosa* populations had genetic diversity similar to those in Thailand [13], whereas the remaining populations were more similar to those in Guangxi Province [15]. Additionally, *R. tomentosa* populations located

in mountainous regions and on islands had higher $F_{IS}$ values, but still maintained high genetic diversity ($H_O$, $H_E$), indicating the influence of geographic barriers on island populations, as documented in various plant species [4, 71]. An increase in the probability of selfing can lead to the loss of certain alleles within *R. tomentosa* populations, thereby reducing the population genetic diversity [72]. This suggests that although environmental barriers may increase selfing in *R. tomentosa* populations, these regional populations have retained relatively high genetic diversity, likely due to the high levels of genetic diversity accumulated during the expansion following a bottleneck (Fig. 3C and D).

### Environmental factor impacts on *R. tomentosa* population differentiation

Natural populations differentiating into groups is often caused by neutral processes such as random genetic drift, admixtures from different population sources, bottlenecks, founder effects, and local differentiation driven by local environmental conditions during range expansion [73–76]. The magnitude of geographic distance is related to the interplay between genetic drift and migration, and both geographic distance and environmental differences are the most important driving factors influencing population genetic differentiation [18, 77]. IBD is typically closely related to the migration patterns of populations and the limitations of gene flow, and it applies to most geographically isolated species. In the IBE model, genetic differences between populations are mainly driven by environmental factors, rather than solely by geographic distance [77].

Investigating and quantifying the impact of environmental factors on this differentiation is crucial for species conservation. Our IBD and IBE analyses indicated a significant correlation between *R. tomentosa* genetic distance and both geographic and environmental distances. Similar results were found in studies of *Cephalotaxus oliveri* populations in the same distribution areas, where both IBD and IBE were found to jointly influence species differentiation [78]. However, previous research in Thailand did not detect significant IBD [13], possibly owing to insufficient geographic isolation in the sampling areas. Through RDA of 11 environmental factors, we found the environmental factors Uvb1, BIO1, BIO14, and Soil1 explained a significant portion of the genetic distance variation. Notably, Uvb1 significantly influenced the differentiation between other populations and GROUP3 (Hainan Island populations). We used RDA to provide initial insights into the contributions of environmental differences to population genetic differentiation by comparing the explanatory power of 11 environmental factors (Table 1) on the differentiation of 28 populations.

### Historical formation of the distribution pattern of *R. tomentosa*

Our study provides insights into the origin and diffusion pathway of *R. tomentosa* in China. In the BEAST tree (Fig. 3A), GROUP3 is located at the base of all populations, representing an ancestral population. Evidence of gene flow also suggests migration from Hainan Island to the mainland (Fig. 3F ① and ⑧) and from coastal populations inland (Fig. 3F ②, ③, ④, and ⑤), indicating that mainland populations of *R. tomentosa* originated from the expansion of populations from Hainan Island or coastal regions. To further determine the origin of Hainan Island populations, we considered the geological history of Hainan Island. During the Eocene, Hainan Island separated from the Beibu Gulf region of northern Guangxi and began a clockwise rotation of approximately 150°, with most of this rotation occurring between 40 and 24 MYA [79]. Genetic distance analysis shows that, despite Hainan Island being geographically closer to the Leizhou Peninsula in Guangdong, the genetic distance between the populations of Hainan Island and those in Guangxi is significantly smaller. This evidence supports the hypothesis that the species originated in Hainan Island or the coastal areas of Guangxi. Although gene flow from GROUP3 to GROUP2 (Fig. 3F ⑧) could suggest migration from Hainan Island to Guangxi, gene flow from GXFC to HNHK (Fig. 3F ⑥) indicates another possibility, making it difficult to determine whether *R. tomentosa* first appeared on Hainan Island or the coastal areas of Guangxi. Considering the migration patterns of other subtropical species in China [80–82], we believe that the model of migration from Southeast Asia into Guangxi, spreading to Hainan Island and further expanding inland, is more plausible.

Our study also provides insights into the two major diversification events of *R. tomentosa* in China. The first divergence occurred around 0.93 MYA, when populations on Hainan Island (GROUP3) and mainland populations (GROUP1 and GROUP2) diverged. This divergence may have been driven by three important factors: First, historical effective population size analysis showed a continuous decline in effective population size from 1 to 0.05 MYA, coinciding with the abrupt cooling period of MIS 20 to MIS 21 (around 0.9 MYA), which caused habitat contraction, reduced species interaction, and promoted species differentiation. Land bridges between Hainan Island and the mainland formed three times during the Middle Pleistocene, approximately 0.6–0.8 MYA, 0.42–0.48 MYA, and 0.13–0.3 MYA [83]. However, since the divergence occurred around 0.93 MYA, the absence of timely land bridges may have contributed to species differentiation. Third, changes in bird activity due to extreme climate conditions could have altered gene flow, further promoting species isolation. Evidence

suggests that the shrub-feeding bird *Alcippe morrisonia* on Hainan Island diverged into two lineages around 1.15 MYA [84]. The second divergence event occurred around 0.09 MYA between populations of the eastern mainland (GROUP1) and the western mainland (GROUP2). Historical effective population size analysis revealed a slight contraction-expansion cycle in GROUP1 populations around 0.2–0.1 MYA (Fig. 3B). This period coincides with marine transgression events in eastern China [85]. We infer that the second divergence was caused by isolation due to species exchange barriers during marine transgression events.

The distribution pattern of a species is the result of the interaction between external factors (such as historical climate and topography) and internal factors (such as biological characteristics, population history, and gene flow). Unveiling the distribution pattern of a species not only aids in its understanding and conservation but also provides insights into historical organismal evolution within East Asia.

## Conclusions

In the present study, genetic data from 284 samples belonging to 28 populations were collected to assess population genetic structure, distribution, and genetic diversity characteristics of *R. tomentosa* in southern China. The populations were divided into three groups: the eastern mainland, western mainland, and Hainan Island. We found that Hainan Island and its surrounding areas are genetic diversity centers for the species. The low genetic diversity of the eastern coastal populations indicates that these populations need protection. By combining genetic and environmental data, we determined the influence of geographic and environmental differences on population differentiation. By combining the results of evolutionary history and gene flow, we infer that the Chinese *R. tomentosa* originated from Southeast Asia, and that dramatic climatic change and geological processes during the Pleistocene were significant causes of bottlenecking and population differentiation. These findings are of major importance for the genetic conservation and utilization of *R. tomentosa*. Future work should expand the scope of the research and make more use of WGS data, which provides more genetic information.

### Abbreviations

| | |
|---|---|
| RAD-seq | Restriction site Associated DNA sequencing |
| NGS Next | Generation Sequencing |
| WGS Whole | Genome Sequencing |
| SNP | Single Nucleotide Polymorphism |
| HWE | Hardy-Weinberg Equilibrium |
| LD | Linkage Disequilibrium |
| $H_O$ | Observed Heterozygosity |
| $H_E$ | Expected Heterozygosity |
| Π | Nucleotide Diversity |
| $F_{IS}$ | Inbreeding coefficient |
| $F_{ST}$ | Genetic diferentiation coefcient |
| PCA | Principal Component Analysis |
| IBD | Isolation By Distance |
| IBE | Isolation By Environment |
| RDA | Redundancy Analysis |
| VIF | Variance Inflation Factor |
| PSMC′ | Pairwise Sequentially Markovian Coalescent |
| BIC | Bayesian Information Criterion |
| MCMC | Markov chain Monte Carlo |
| CVH | Chinese Virtual Herbarium |
| GBIF | Global Biodiversity Information Facility |
| ROC | Receiver Operating Characteristic |
| AUC | Area Under the Curve |
| MH | Mid-Holocene |
| LGM | Last Glacial Maximum |
| LIG | Last Interglacial |
| SSR | Simple Sequence Repeat |
| ISSR | Inter-Simple Sequence Repeat |
| MYA | Million Years Ago |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06364-6.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

### Acknowledgements

We thank Su-fang Chen for discussing some key points for this paper initially. We would like to thank Xin-lei Zhao, Zhi-liang Peng, and Yi Bai for their kind help in collecting the samples. We would also like to thank the reviewers for their constructive comments, which significantly improved our study.

### Author contributions

Bo-yong Liao and Yong-quan Li designed the project. Long-yuan Wang, Fang-qiu Zhang, and Li-xia Peng collected plant material, and Xing-ming Xu and Su-jiao Liao analyzed the data. Xing-ming Xu and Su-jiao Liao wrote the manuscript. Qiao-mei Qin, Chun-yan He, Xin Ding, Wei Wu, and Bryan Drew reviewed the manuscript. All authors read and approved the final manuscript.

### Funding

This research was funded by the Guangdong Province Forestry Science and Technology Innovation Project (2020KJCX011), the Scarce and Quality Economic Forest Engineering Technology Research Center (2022GCZX002) and the Science and Technology Program of Guangzhou (202201011461).

### Data availability

All data important to understand the manuscript are attached as supplementary material. Raw data have been uploaded to GenBank. shallow WGS data (BioProject: PRJNA912363; accession numbers: SRR26683076–SRR26683105), and RAD-seq data (BioProject: PRJNA912363; accession numbers: SRR22820469–SRR22820795).

## Declarations

### Ethics approval and consent to participate

Our research materials are derived from natural populations, and the sample collection work is approved by the forestry administration. All our collections and experiments are conducted in accordance with relevant institutional, national, and international norms and regulations.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

**Author details**
[1]College of Horticulture and Landscape Architecture, Zhongkai University of Agriculture and Engineering, Guangzhou, Guangdong 510220, China
[2]Guangdong Eco-engineering Polytechnic, Guangzhou, Guangdong 510630, China
[3]Department of Biology, University of Nebraska-Kearney, Kearney, NE 68849, USA
[4]Department of Education of Guangdong Province, Guangdong Provincial Engineering Technology Research Center for High-quality, Rare, and Characteristic Economic Forest and Fruit Trees in Regular Higher Education Institutions, No.501 of Zhongkai Road, Guangzhou, Guangdong 510225, China

Received: 6 May 2024 / Accepted: 7 March 2025
Published online: 15 March 2025

**References**

1. Meyer J. Threat of invasive alien plants to native flora and forest vegetation of Eastern Polynesia. Pac Sci. 2004;58(3):357–75.
2. Si SB, Cao FX, Peng JQ, Xu RX. Delineation of the Northern line of Myrtle community distribution in China and its relationship with Climatic factors. Zhongnan Linye Keji Daxue Xuebao. 2012;32(03):162–5.
3. Ren H, Peng SL, Dai ZM, Liang XD, Cai XA, Lin RB. Ecological and biological characteristics of *Wikstroemia indica*. Yingyong Shengtai Xuebao. 2002(12):1529–32.
4. Wei MS, Chen ZH, Ren H, Yin ZY. Reproductive ecology of *Rhodomyrtus tomentosa* (Myrtaceae). Yingyong Shengtai Xuebao. 2009;27(2):154–60.
5. Zhao ZG, Cheng W, Guo JJ. Resource utilization and artificial cultivation of *Rhodomyrtus tomentosa*. Guangxi Shengtai Xuebao. 2006(02):70–2.
6. Limsuwan S, Kayser O, Voravuthikunchai SP. Antibacterial activity of *Rhodomyrtus tomentosa* (Aiton) Hassk. leaf extract against clinical isolates of Streptococcus pyogenes. Evid-Based Compl Alt Med. 2012;2012.
7. Na-Phatthalung P, Chusri S, Suanyuk N, Voravuthikunchai SP. In vitro and in vivo assessments of *Rhodomyrtus tomentosa* leaf extract as an alternative anti-streptococcal agent in nile tilapia (*Oreochromis niloticus* L). J Med Microbiol. 2017;66(4):430–9.
8. Zhang YB, Li W, Jiang L, Yang L, Chen NH, Wu ZN, Li YL, Wang GC. Cytotoxic and anti-inflammatory active phloroglucinol derivatives from *Rhodomyrtus tomentosa*. Phytochemistry. 2018;153:111–9.
9. Na-Phatthalung P, Teles M, Voravuthikunchai SP, Tort L, Fierro-Castro C. Immune-related gene expression and physiological responses in rainbow trout (*Oncorhynchus mykiss*) after intraperitoneal administration of *Rhodomyrtus tomentosa* leaf extract: A potent phytoimmunostimulant. Fish Shellfish Immun. 2018;77:429–37.
10. Abd Hamid H, Mutazah R, Yusoff MM, Abd Karim NA, Razis AFA. Comparative analysis of antioxidant and antiproliferative activities of *Rhodomyrtus tomentosa* extracts prepared with various solvents. Food Chem Toxicol. 2017;108:451–7.
11. Ren H, Yang L, Liu N. Nurse plant theory and its application in ecological restoration in lower subtropics of China. Prog Nat Sci. 2008;18(2):137–42.
12. Liang KM, Yang XY, Zhang JE, Zhao BL, Luo H, Guo J. The nursing effect of Melastoma candidum, Rhodomyrtus tomentosa and Acacia farnesiana on the survival rate of Indian sandalwood (Santalum album) seedlings in South China. Shengtai Xuebao. 2014;33(3):480–7.
13. Detcharoen M, Bumrungsri S, Voravuthikunchai SP. Complete genome of Rose Myrtle, *Rhodomyrtus tomentosa*, and its population genetics in Thai Peninsula. Plants-Basel. 2023;12(8):1582.
14. Hue TS, Abdullah TL, Abdullah N, Sinniah UR. Genetic variation in *Rhodomyrtus tomentosa* (Kemunting) populations from Malaysia as revealed by inter-simple sequence repeat markers. Genet Mol Res. 2015;14:16827–39.
15. Sun L, Li J, Sun K, Wang H, Yang K, Chen Q, Lin M. Development and characterization of EST-SSR markers in *Rhodomyrtus tomentosa* Hassk. Based on transcriptome. Genet Resour Crop Evol. 2023;70(6):1691–705.
16. Gao Y, Wang SY, Luo J, Murphy RW, Du R, Wu SF, Zhu CL, Li Y, Poyarkov AD, Nguyen SN, Luan PT, Zhang YP. Quaternary palaeoenvironmental oscillations drove the evolution of the Eurasian Carassius auratus complex (*Cypriniformes, Cyprinidae*). J Biogeogr. 2012;39(12):2264–78.
17. Hewitt GM. Genetic consequences of Climatic oscillations in the quaternary. Philos Trans R Soc Lond Ser B Bio Sci. 2004;359(1442):183–95.
18. Manel S, Schwartz MK, Luikart G, Taberlet P. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol. 2003;18(4):189–97.
19. Husemann M, Ray JW, King RS, Hooser EA, Danley PD. Comparative biogeography reveals differences in population genetic structure of five species of stream fishes. Biol J Linn Soc. 2012;107(4):867–85.
20. Kyriazi P, Kornilios P, Nagy ZT, Poulakakis N, Kumlutaş Y, Ilgaz Ç, Avcı A, Göçmen B, Lymberakis P. Comparative phylogeography reveals distinct colonization patterns of C Retan snakes. J Biogeogr. 2013;40(6):1143–55.
21. Kremer A, Potts BM, Delzon S. Genetic divergence in forest trees: Understanding the consequences of climate change. Funct Ecol. 2014:22–36.
22. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE. 2008;3(10):e3376.
23. Davey JW, Blaxter ML. RADSeq: next-generation population genetics. Brief Funct Genomics. 2010;9(5–6):416–23.
24. Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Res. 2007;17(2):240–8.
25. Mousavi M, Tong C, Liu F, Tao S, Wu J, Li H, Shi J. De Novo SNP discovery and genetic linkage mapping in Poplar using restriction site associated DNA and whole-genome sequencing technologies. BMC Genomics. 2016;17:1–12.
26. Sarkar D, Kundu A, Das D, Chakraborty A, Mandal NA, Satya P, Karmakar PG, Kar CS, Mitra J, Singh NK. Resolving population structure and genetic differentiation associated with RAD-SNP loci under selection in Tossa jute (*Corchorus olitorius* L). Mol Genet Genomics. 2019;294:479–92.
27. Yan SY, Zhu P, Gong W, Wang JY, Wu KZ, Wu CY. Studies on genetic diversity of Juglans cultivar germplasms in Sichuan based on RAD-SNPs analysis. J Trop Subtrop Bot. 2019;27(1):19–28.
28. Zhong YD, Yang AH, Liu SJ, Liu LP, Li YQ, Wu ZX, Yu FX. RAD-Seq data point to a distinct split in *Liriodendron* (Magnoliaceae) and Obvious east–west genetic divergence in L. chinense. Forests. 2018;10(1):13.
29. Liu S, Hansen MM. PSMC (pairwise sequentially Markovian coalescent) analysis of RAD (restriction site associated DNA) sequencing data. Mol Ecol Resour. 2017;17(4):631–41.
30. Rogers SO, Bendich AJ. Extraction of DNA from plant tissues. Plant Mol Biology Man. 1989:73–83.
31. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. Mol Ecol. 2013;22(11):3124–40.
32. Li FP, Xu SQ, Xiao ZT, Wang JM, Mei Y, Hu HF, Li JY, Hou ZW, Zhao JL, Yang SH, Wang JH. Gap-free genome assembly and comparative analysis reveal the evolution and anthocyanin accumulation mechanism of *Rhodomyrtus tomentosa*. Hortic Res. 2023;10(3):uhad005.
33. Li H. Aligning sequence Reads, clone sequences and assembly contigs with BWA-MEM. ArXiv Preprint ArXiv:1303.3997. 2013.
34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and samtools. Bioinformatics. 2009;25(16):2078–9.
35. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, McVean G. The variant call format and vcftools. Bioinformatics. 2011;27(15):2156–8.
36. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
37. Jin JJ, Yu WB, Yang JB, Song Y, DePamphilis CW, Yi TS, Li DZ. GetOrganelle: a fast and versatile toolkit for accurate de Novo assembly of organelle genomes. Genome Biol. 2020;21:1–31.
38. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de Novo genome assemblies. Bioinformatics. 2015;31(20):3350–2.
39. Schmidt TL, Jasper ME, Weeks AR, Hoffmann AA. Unbiased population heterozygosity estimates from genome-wide sequence data. Meth Ecol Evol. 2021;12(10):1888–98.
40. Alexander DH, Novembre J, Lange K. Fast model-based Estimation of ancestry in unrelated individuals. Genome Res. 2009;19(9):1655–64.
41. Novembre J, Stephens M. Interpreting principal component analyses of Spatial population genetic variation. Nat Genet. 2008;40(5):646–9.
42. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S. Daly M depristo MA. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–303.
43. Hijmans RJ, Williams E, Vennes C, Hijmans MRJ. Package 'geosphere'. Spher Trigon. 2017;1(7):1–45.

44. Oksanen J, Kindt R, Legendre P, O Hara B, Stevens MHH, Oksanen MJ, Suggests MASS. The vegan package. Comm Ecol Package. 2007;10(631–637):719.
45. Peakall R, Smouse PE. GENALEX 6: genetic analysis in excel. Population genetic software for teaching and research. Mol Ecol Notes. 2006;6(1):288–95.
46. Dray S, Legendre P, Peres-Neto PR. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecol Model. 2006;196(3–4):483–93.
47. Yamada KD, Tomii K, Katoh K. Application of the MAFFT sequence alignment program to large data—reexamination of the usefulness of chained guide trees. Bioinformatics. 2016;32(21):3246–51.
48. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35(6):1547.
49. Nguyen L, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74.
50. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. BEAST 2: a software platform for bayesian evolutionary analysis. Plos Comput Biol. 2014;10(4):e1003537.
51. Schiffels S, Wang K. MSMC and MSMC2: the multiple sequentially Markovian coalescent. Methods Mol Biology (Clifton NJ). 2020;2090:147–66.
52. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. Twelve years of samtools and BCFtools. Gigascience. 2021;10(2):giab008.
53. Low YW, Rajaraman S, Tomlin CM, Ahmad JA, Ardi WH, Armstrong K, Athen P, Berhaman A, Bone RE, Cheek M, Cho NRW, Choo LM, Cowie LD, Crayn D, Fleck SJ, Ford AJ, Forster PL, Girmansyah D, Goyder DJ, Gray B, Heatubun CD, Lbrahim A, Lbrahim B, Jayasinghe HD, Kalat MA, Kathriarachchi HS, Kintamani E, Koh SL, Lai JTK, Lee SML, Leong PKF, Lim WH, Lum SKY, Mahyuni R, McDonald WJF, Metali F, Ranasinghe S, Repin R, Rustiami H, Simbiak VI, Sukri RS, Sunarti Siti, Trethowan LA, Trias-Blasi A, Vasconcelos TNC, Wanma JF, Widodo P, Wijesundara DSA, Worboys DSA, Worboys S, Yap JW, Yong KT, Khew GSW, Salojärvi J, Michael TP, Middleton DJ, Burslem DFRP, Lindqvist C. Lucas EJ Albert VA. Genomic insights into rapid speciation within the world's largest tree genus Syzygium. Nat Commun. 2022;13(1):5031.
54. Cornuet J, Santos F, Beaumont MA, Robert CP, Marin J, Balding DJ, Guillemaud T, Estoup A. Inferring population history with DIY ABC: a user-friendly approach to approximate bayesian computation. Bioinformatics. 2008;24(23):2713–9.
55. Pickrell J, Pritchard J. Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. 2012;8:e1002967.
56. Malinsky M, Matschiner M, Svardal H. Dsuite-Fast D-statistics and related admixture evidence from VCF files. Mol Ecol Resour. 2021;21(2):584–95.
57. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190(3–4):231–59.
58. Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. A statistical explanation of maxent for ecologists. Divers Distrib. 2011;17(1):43–57.
59. Swets JA. Measuring the accuracy of diagnostic systems. Science. 1988;240(4857):1285–93.
60. Xie C, Huang B, Jim CY, Han W, Liu D. Predicting differential habitat suitability of *Rhodomyrtus tomentosa* under current and future climate scenarios in China. For Ecol Manag. 2021;501:119696.
61. Abolmaali SM, Tarkesh M, Bashari H. MaxEnt modeling for predicting suitable habitats and identifying the effects of climate change on a threatened species, Daphne mucronata, in central Iran. Ecol Inf. 2018;43:116–23.
62. Ji RX, Yu X, Ren TM, Chang Y, Li Z, Xia XL, Yin WL, Liu C. Genetic diversity and population structure of *Caryopteris mongholica* revealed by reduced representation sequencing. BMC Plant Biol. 2022;22(1):297.
63. Janjua S, Peters JL, Weckworth B, Abbas FI, Bahn V, Johansson O, Rooney TP. Improving our conservation genetic toolkit: ddRAD-seq for SNPs in snow leopards. Conserv Genet Resour. 2020;12:257–61.
64. Zhu XL, Zou R, Qin HZ, Chai SF, Tang JM, Li YY, Wei X. Genome-wide diversity evaluation and core germplasm extraction in ex situ conservation: A case of golden *Camellia tunghinensis*. Evol Appl. 2023;16(9):1519–30.
65. Zhao Z, Song Q, Bai D, Niu S, He Y, Qiao D, Chen ZW, Li CY, Luo J, Li F. Population structure analysis to explore genetic diversity and geographical distribution characteristics of cultivated-type tea plant in Guizhou plateau. BMC Plant Biol. 2022;22(1):55.
66. Oztolan-Erol N, Helmstetter AJ, İnan A, Buggs RJ, Lucas SJ. Unraveling genetic diversity amongst European hazelnut (*Corylus Avellana* L.) varieties in Turkey. Front Plant Sci. 2021;12:661274.
67. Zhang XJ, Liu XF, Liu DT, Cao YR, Li ZH, Ma YP, Ma H. Genetic diversity and structure of *Rhododendron meddianum*, a plant species with extremely small populations. Plant Divers. 2021;43(6):472–9.
68. Huang Y, Li Y, Hong X, Luo S, Cai D, Xiao X, Huang Y, Zheng Y. Genetic variation for wild populations of the rare and endangered plant *Glyptostrobus pensilis* based on Double-Digest restriction Site-Associated DNA sequencing. Curr Issues Mol Biol. 2025;47(1):12.
69. Clugston JA, Ruhsam M, Kenicer GJ, Henwood M, Milne R, Nagalingum NS. Conservation genomics of an Australian cycad *Cycas calcicola*, and the absence of key genotypes in botanic gardens. Conserv Genet. 2022;23(3):449–65.
70. Zimmerman SJ, Aldridge CL, Oyler-McCance SJ. An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. BMC Genomics. 2020;21:1–16.
71. Hargreaves S, Maxted N, Hirano R, Abberton M, Skøt L, Ford-Lloyd BV. Islands as refugia of Trifolium repens genetic diversity. Conserv Genet. 2010;11:1317–26.
72. Wright S. Isolation by distance. Genetics. 1943;28(2):114.
73. Bertelsmeier C, Keller L. Bridgehead effects and role of adaptive evolution in invasive populations. Trends Ecol Evol. 2018;33(7):527–34.
74. Dai JX, Cao LJ, Chen JC, Yang F, Shen XJ, Ma LJ, Hoffmann AA, Chen M, Wei SJ. Testing for adaptive changes linked to range expansion following a single introduction of the fall webworm. Mol Ecol. 2024;33:e17038.
75. Excoffier L, Foll M, Petit RJ. Genetic consequences of range expansions. Annu Rev Ecol Evol Syst. 2009;40:481–501.
76. Wiberg R, Tyukmaeva V, Hoikkala A, Ritchie MG, Kankare M. Cold adaptation drives population genomic divergence in the ecological specialist, *Drosophila montana*. BioRxiv. 2021.
77. Sexton JP, Hangartner SB, Hoffmann AA. Genetic isolation by environment or distance: which pattern of gene flow is most common? Evolution. 2014;68(1):1–15.
78. Liu HJ, Li MH, Wang Z, Wang T, Su YJ. Local adaptation and demographic history of vulnerable conifer *Cephalotaxus Oliveri* in Southern China. J Syst Evol. 2024;62(3):457–74.
79. Liang GH. A study of the genesis of Hainan Island. Geol China. 2018;45(4):693–705.
80. Lin X, Feng C, Lin T, Harris AJ, Li Y, Kang M. Jackfruit genome and population genomics provide insights into fruit evolution and domestication history in China. Hortic Res. 2022;9:uhac173.
81. Meng HH, Zhang CY, Song YG, Yu XQ, Cao GL, Li L, Cai CN, Xiao JN, Zhou SS, Tan YH, Li J. Opening a door to the Spatiotemporal history of plants from the tropical Indochina Peninsula to subtropical China. Mol Phylogenet Evol. 2022;171:107458.
82. Wang XH, Li J, Zhang LM, He ZW, Mei QM, Gong X, Jian SG. Population differentiation and demographic history of the *Cycas taiwaniana* complex (Cycadaceae) endemic to South China as indicated by DNA sequences and microsatellite markers. Front Genet. 2019;10:1238.
83. Shi YF, Cui ZJ, Su Z. The quaternary glaciations and environmental variations in China. Shijiazhuang: Hebei Science and Technology; 2006. pp. 173–9.
84. Song G, Qu Y, Yin Z, Li S, Liu N, Lei F. Phylogeography of the *Alcippe morrisonia* (Aves: Timaliidae): long population history beyond late pleistocene glaciations. BMC Evol Biol. 2009;9:1–11.
85. Chen T, Wang ZH, Qiang XK, MA CY, Zhan Q. Magnetic properties of minerals recorded by the borehole WJ and late quaternary transgressions in the Taihu plain, Southern Yangtze Delta. Chin J Geophys. 2013;56(8):2748–59.


## Publisher's note