

## Article

# A Diverging Species within the *Stewartia gemmata* (Theaceae) Complex Revealed by RAD-Seq Data

Hanyang Lin <sup>1</sup>, Wenhao Li <sup>2</sup> and Yunpeng Zhao <sup>2,\*</sup>

<sup>1</sup> Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, School of Life Sciences, Taizhou University, Taizhou 318000, China; hylin@tzc.edu.cn

<sup>2</sup> Systematic & Evolutionary Botany and Biodiversity Group, MOE Key Laboratory of Biosystems Homeostasis & Protection, College of Life Sciences, Zhejiang University, Hangzhou 310058, China; wenhao\_li@zju.edu.cn

\* Correspondence: ypzha@zju.edu.cn

**Abstract:** Informed species delimitation is crucial in diverse biological fields; however, it can be problematic for species complexes. Showing a peripatric distribution pattern, *Stewartia gemmata* and *S. acutisepala* (the *S. gemmata* complex) provide us with an opportunity to study species boundaries among taxa undergoing nascent speciation. Here, we generated genomic data from representative individuals across the natural distribution ranges of the *S. gemmata* complex using restriction site-associated DNA sequencing (RAD-seq). Based on the DNA sequence of assembled loci containing 41,436 single-nucleotide polymorphisms (SNPs) and invariant sites, the phylogenetic analysis suggested strong monophyly of both the *S. gemmata* complex and *S. acutisepala*, and the latter was nested within the former. Among *S. gemmata* individuals, the one sampled from Mt. Tianmu (Zhejiang) showed the closest evolutionary affinity with *S. acutisepala* (which is endemic to southern Zhejiang). Estimated from 2996 high-quality SNPs, the genetic divergence between *S. gemmata* and *S. acutisepala* was relatively low (an  $F_{st}$  of 0.073 on a per-site basis). Nevertheless, we observed a proportion of genomic regions showing relatively high genetic differentiation on a windowed basis. Up to 1037 genomic bins showed an  $F_{st}$  value greater than 0.25, accounting for 8.31% of the total. After SNPs subject to linkage disequilibrium were pruned, the principal component analysis (PCA) showed that *S. acutisepala* diverged from *S. gemmata* along the first and the second PCs to some extent. By applying phylogenomic analysis, the present study determines that *S. acutisepala* is a variety of *S. gemmata* and is diverging from *S. gemmata*, providing empirical insights into the nascent speciation within a species complex.

**Keywords:** nascent species; peripatric distribution; RAD-seq; phylogenomic analysis; taxonomic implication



**Citation:** Lin, H.; Li, W.; Zhao, Y. A Diverging Species within the *Stewartia gemmata* (Theaceae) Complex Revealed by RAD-Seq Data. *Plants* **2024**, *13*, 1296. <https://doi.org/10.3390/plants13101296>

Academic Editors: Alex Troitsky and Jaakko Hyvönen

Received: 18 March 2024

Revised: 4 May 2024

Accepted: 7 May 2024

Published: 8 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Species are the fundamental units of biodiversity, yet the ways in which species have been defined have been a longstanding topic since the epoch of Charles Darwin [1]. Historically, various species concepts and criteria for defining species have been proposed and discussed [2–9], and a general agreement has been reached that species are lineages in terms of biological, evolutionary, ecological, or phenetic aspects, in either a broad or strict sense [10]. Therefore, informed species delimitation should be based on integrated evidence [11,12]. If speciation events occurred a long time ago, species may well fit certain species concepts with strong evidence; however, issues can become complicated when species are still at the nascent stage (i.e., speciation is not complete) [13]. Some examples are as follows: (1) lineages diversify with indistinguishable morphological changes, also known as ‘cryptic/sibling species’ [14] or ‘non-adaptive radiation’ [15]; (2) the presence of discernible phenotypic divergence but little genetic differentiation, which is not rare in adaptive radiations where lineage boundaries are porous [16–18]; (3) lineages with

morphological and genetic distinctions arise from a former metapopulation, i.e., the case of progenitor-derivative relationships, in which paraphyletic assemblages of populations are predicted to be common [10]. The practice of species delimitation within species complexes that fall into the abovementioned situations will be challenging yet important given that it would largely facilitate our understanding of the patterns and causes at the initial step of speciation [19].

Showing a disjunct distribution and a species diversity disparity between eastern Asia and eastern North America (ca. 18 vs. 2 spp.), *Stewartia* L. (an early-diverged lineage in Theaceae) provides us with an ideal system for studying speciation mechanisms and local adaptation, along with its taxonomic difficulties [20–24]. Among Asian *Stewartia* species, *S. sinensis* Rehder & E. H. Wilson was recognized as having the largest distribution area and highest phenotypic diversity [21,22]. According to *Flora of China* (FOC), *S. sinensis* harbors four varieties, namely *S. sinensis* var. *sinensis*, *S. sinensis* var. *acutisepala* (P. L. Chiu & G. R. Zhong) T. L. Ming & J. Li (= *S. acutisepala* P. L. Chiu & G. R. Zhong), *S. sinensis* var. *brevicalyx* (S. Z. Yan) T. L. Ming & J. Li (= *S. brevicalyx* S. Z. Yan), and *S. sinensis* var. *shensiensis* (Hung T. Chang) T. L. Ming & J. Li [22]. With an integrated taxonomic approach, recent research decoupled *S. sinensis* into two distinct species based on lines of evidence from phylogenetic affinities, climatic niche volumes, and key morphological traits [25]. The geographically northern lineage (including part of *S. sinensis* var. *sinensis* and *S. sinensis* var. *shensiensis* recognized by FOC) comprised the species entity of *S. sinensis*, while the geographically southern lineage (including part of *S. sinensis* var. *sinensis*, *S. sinensis* var. *acutisepala*, and *S. sinensis* var. *brevicalyx* recognized by FOC) was assigned the name of *S. gemmata* S. S. Chien & W. C. Cheng, a legitimate name with resurrection [22,26]. Despite the clarification of the species boundary between *S. sinensis* and *S. gemmata*, the taxonomic relationships among *S. gemmata*, *S. acutisepala*, and *S. brevicalyx* (the *S. gemmata* complex, hereafter) remain elusive (Table S1). In the biogeographic sense, *S. gemmata* mainly inhabits temperate broadleaf and mixed forests in eastern to southern China (mainly Zhejiang, Anhui, Jiangxi, Fujian, Hunan, Guangxi, and Guangdong), *S. acutisepala* is endemic to mountainous regions in southern Zhejiang of eastern China, and *S. brevicalyx* can only be found at Mt. Tianmu (northern Zhejiang) and its adjacent areas according to the literature [20,25,27]. Therefore, *S. acutisepala* shows a peripatric distribution with *S. gemmata*, and *S. brevicalyx* is sympatric with *S. gemmata* but with extremely limited distribution (if it exists). Morphologically, *S. acutisepala* mainly differs from *S. gemmata* in the exfoliation pattern of barks, while no evident differences were found between *S. gemmata* and *S. brevicalyx*, as documented in a recently revised regional flora [28]. A molecular dating analysis suggested that *S. acutisepala* diverged from *S. gemmata* ca. 2 million years ago as its sister group, though the robustness of this inference remains to be tested given that only a few individuals were included [24]. Therefore, we hypothesize that *S. acutisepala* is a variety of *S. gemmata* or a sister species of *S. gemmata*, and *S. brevicalyx* is likely conspecific with *S. gemmata*. In summary, the *S. gemmata* complex (especially *S. gemmata* and *S. acutisepala*) is a good candidate for exploring the existence of nascent species and the underlying speciation mechanisms, and a deeper investigation into species boundaries within this species complex is called for.

Taking advantage of the rapid development of DNA sequencing technologies, systematists are now embracing nuclear genomic data to decipher the evolutionary history of diverse organisms [29]. Specifically, an increasing body of studies has harnessed the power of single-nucleotide polymorphisms (SNPs) to resolve recalcitrant phylogenetic relationships [30]. Compared to alternative genomic markers, SNPs are abundant genome-wide and easy to collect. Emerging techniques for the de novo sequencing and identification of SNPs, such as restriction site-associated DNA sequencing (RAD-seq), are increasingly welcomed in evolutionary research [30–33]. By discovering and screening SNPs in a fraction of the genome, RAD-seq could generate massive SNPs with relatively low economic and time costs, which have been proven efficient in empirical species delimitation studies [34–39].

Therefore, utilizing RAD-seq data to trace the evolutionary imprints of the *S. gemmata* complex is promising.

To test whether *S. acutisepala* is a variety or a sister species of *S. gemmata*, we sampled representative individuals of the *S. gemmata* complex across its native ranges and generated genomic data using RAD-seq to ask if the two species represent distinct evolutionary lineages, and if so, how divergent they are from one another.

## 2. Results

### 2.1. Assembled Data

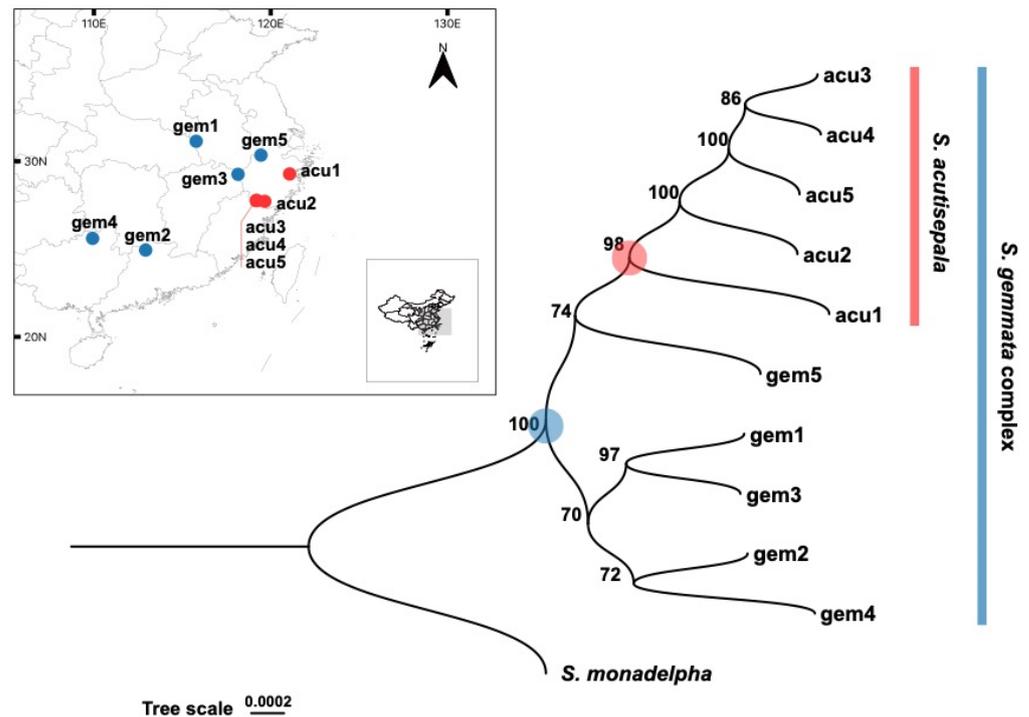
After filtering the adapter sequences and low-quality reads, the RAD-seq generated an average of 3.75 Gb (2.75–5.45 Gb) of clean data for each sample (Table 1). With its default settings, the ipyrad pipeline retained a total of 41,436 SNPs, of which each was shared by at least four individuals. The sequence alignment of assembled loci (including invariant sites) was 1,808,596 bp long. A detailed statistical summary of assembled loci among 11 accessions is available in Table S2.

**Table 1.** Information on plant accessions used for RAD-seq in this study.

Species	Locality	Voucher No.	Longitude	Latitude	ID	Clean Data (Gb)
<i>S. acutisepala</i>	Mt. Tiantai, Tiantai Co., Taizhou, China	H. Lin 16232	121.068863	29.278347	acu1	2.96
<i>S. acutisepala</i>	Wuyanling National Nature Reserve, Taishun Co., Wenzhou, China	H. Lin HZU13869	119.669257	27.716245	acu2	4.72
<i>S. acutisepala</i>	Mt. Baishanzu, Qingyuan Co., Lishui, China	H. Lin 21013	119.197744	27.762519	acu3	3.95
<i>S. acutisepala</i>	Mt. Baishanzu, Qingyuan Co., Lishui, China	H. Lin 21014	119.196900	27.761892	acu4	5.45
<i>S. acutisepala</i>	Mt. Baishanzu, Qingyuan Co., Lishui, China	H. Lin 21015	119.197056	27.762336	acu5	4.80
<i>S. gemmata</i>	Mt. Tiantangzhai, Jinzhai Co., Lu'an, China	W. Li LWH201704	115.787014	31.135476	gem1	3.39
<i>S. gemmata</i>	Mt. Mang, Yizhang Co., Chenzhou, China	H. Lin 16154	112.930632	24.940206	gem2	2.75
<i>S. gemmata</i>	Mt. Gutian, Kaihua Co., Quzhou, China	H. Lin HZU13983	118.152874	29.255328	gem3	3.87
<i>S. gemmata</i>	Huaping National Nature Reserve, Longsheng Co., Guilin, China	H. Lin 17499	109.929795	25.606376	gem4	3.42
<i>S. gemmata</i>	Mt. Tianmu, Linan Co., Hangzhou, China	X. Zheng ZXM00054	119.448134	30.348420	gem5	2.85
<i>S. monadelphpha</i>	Mt. Ohdai, Yoshino, Nara, Japan	S. Sakaguchi & D. Takahashi SS111-3	135.877671	34.352534	mon	3.09

### 2.2. Phylogenetic Relationship of *S. gemmata* and *S. acutisepala*

Among 484 tested models of DNA substitution, TPM3u + F + I was determined to be the best-fit model of the sequence alignment according to the Bayesian Information Criterion (4,984,151.533). The consensus maximum likelihood (ML) phylogenetic tree (log-likelihood = −2,491,895.601; total tree length = 0.013) suggested a strong monophyly of the *S. gemmata* complex (bootstrap support (BS) = 100; Figure 1).



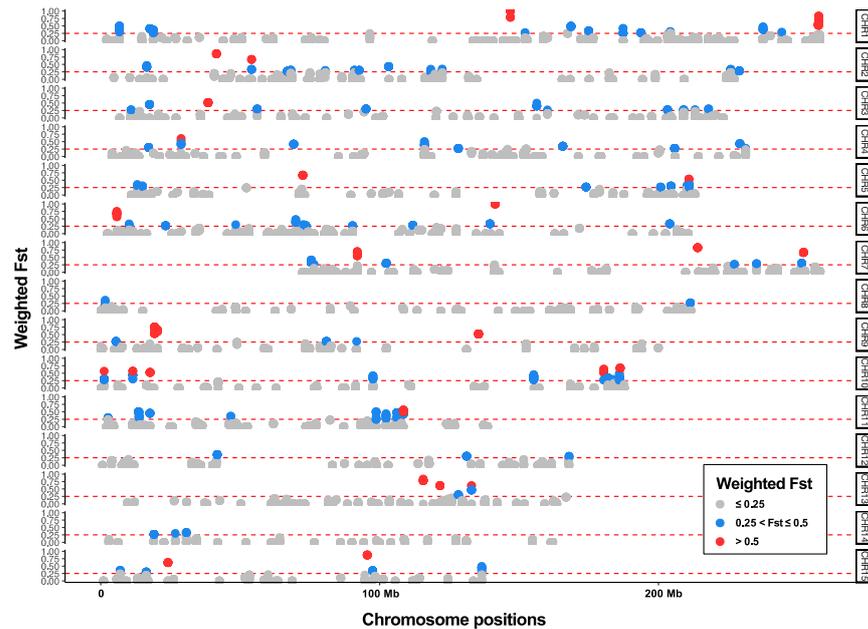
**Figure 1.** Geographic distribution of studied individuals of *Stewartia gemmata* complex and the phylogenetic tree showing the evolutionary affinities among these plants. The distribution ranges (shaded) of *S. gemmata* and *S. acutisepala* were drawn based on [22,25,28]. In the consensus maximum likelihood (ML) tree, ‘gem’ stands for *S. gemmata* (blue) and ‘acu’ stands for *S. acutisepala* (red). The bootstrap support value is shown for each node. The scale bar is in the unit of the number of substitutions per site. The map was retrieved from the Geospatial Data Cloud, Computer Network Information Center, Chinese Academy of Sciences (<http://www.gscloud.cn/>; accessed on 12 November 2023). The accession IDs are consistent with those in Table 1.

Within the *S. gemmata* complex, four *S. gemmata* individuals formed a clade (BS = 70). Among them, gem1 (sampled from Mt. Tiantangzhai, Anhui) was sister to gem3 (sampled from Mt. Gutian, Zhejiang) (BS = 97), and gem2 (sampled from Mt. Mang, Hunan) showed a sister relationship with gem4 (sampled from Huaping National Nature Reserve, Guangxi) (BS = 72). This clade was sister to another clade (BS = 74) consisting of one *S. gemmata* individual (gem5; sampled from Mt. Tianmu, Zhejiang) and all *S. acutisepala* individuals (Figure 1).

The five studied *S. acutisepala* individuals formed a strongly supported clade (BS = 98), which was nested within other *S. gemmata* individuals (Figure 1). Within the *S. acutisepala* clade, acu1 (sampled from Mt. Tiantai) and acu2 (sampled from Wuyanling National Nature Reserve) diverged sequentially, and the three accessions from Mt. Baishanzu (acu3, acu4, and acu5), showed close affinities (BS = 100).

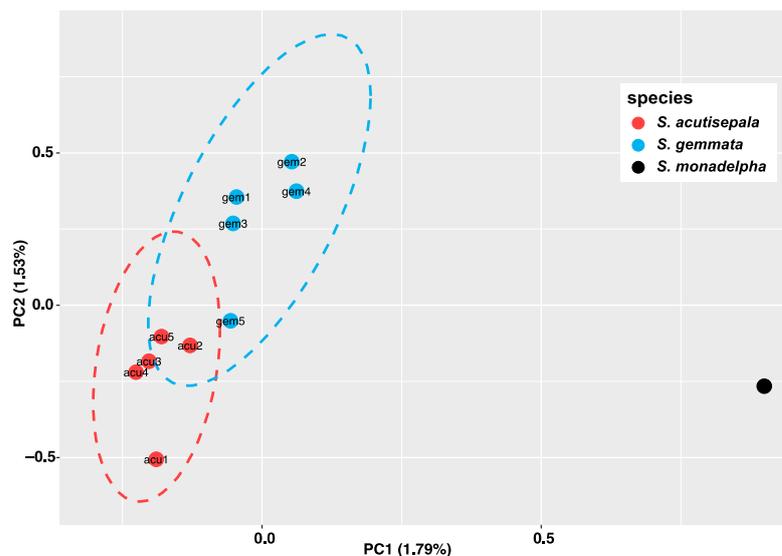
### 2.3. Genetic Differentiation within the *S. gemmata* Complex

The filtering procedure resulted in a dataset containing 2996 high-quality SNPs. Based on these SNPs, the overall weighted Weir and Cockerham’s  $F_{st}$  between *S. gemmata* and *S. acutisepala* was estimated at 0.073 on a per-site basis. Among 12,482 genomic bins, 11,445 bins (91.69%) showed an  $F_{st}$  value equal to or less than 0.25, while 1037 bins showed an  $F_{st}$  value greater than 0.25 (8.31%). More specifically, 787 bins (6.31%) showed an  $F_{st}$  value greater than 0.25 but equal to or less than 0.5, and 250 bins (2.00%) showed an  $F_{st}$  value greater than 0.5 (Figure 2).



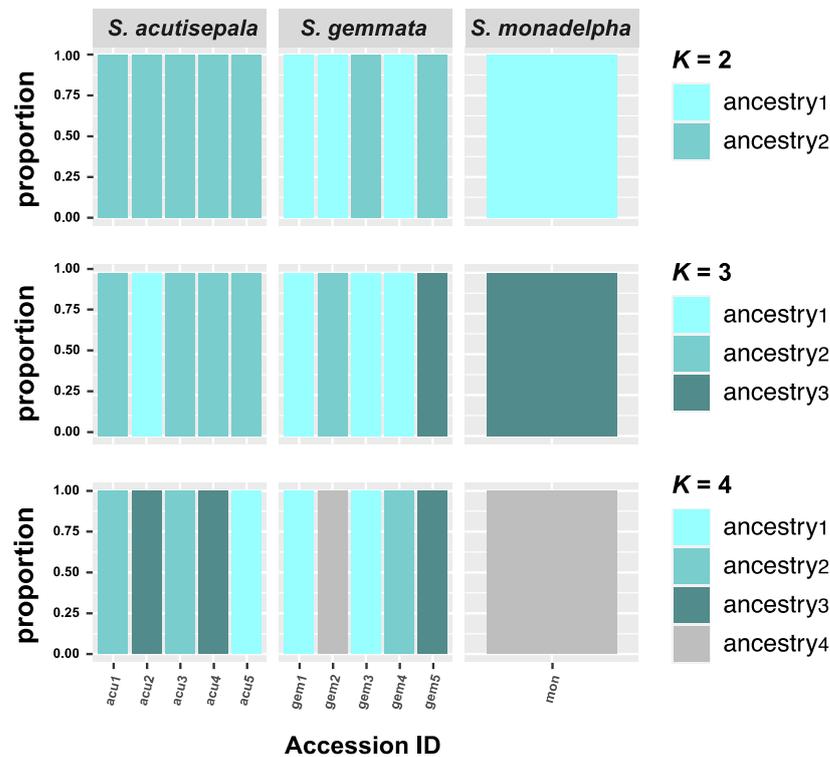
**Figure 2.** Genetic differentiation between *S. gemmata* and *S. acutisepala* measured using weighted Weir and Cockerham's  $F_{st}$  on a windowed basis. The dashed line indicates an  $F_{st}$  value greater than 0.25 in each pseudo-chromosome. The red, blue, and gray genomic bins indicate  $F_{st}$  values  $> 0.50$ ,  $0.50-0.25$ , and  $\leq 0.25$ , respectively.

After SNP sites that were subject to linkage disequilibrium (LD) were pruned, a total of 1673 SNPs were retained. The results of the principal component analysis (PCA) showed that the first and the second PCs (PC1 and PC2) accounted for 1.79% and 1.53% of the total genetic variation, respectively. Along both the PC1 and the PC2 axes, the distribution of *S. acutisepala* was slightly overlapped with that of *S. gemmata* (95% confidence level with a  $t$ -distribution), showing genetic divergence (Figure 3). Among *S. gemmata* individuals, gem5 (sampled from Mt. Tianmu, Zhejiang) showed the closest genetic affinity with *S. acutisepala* individuals, which is consistent with the results of phylogenetic analysis (Figure 1).



**Figure 3.** The distribution of five *S. gemmata* individuals (blue dots), five *S. acutisepala* individuals (red dots), and one *S. monadelpha* individual (the black dot) along the first and the second principal components (PCs) of genetic variations. PC1 explained 1.79% of the total variation, and PC2 explained 1.53% of the total variation. The accession IDs are consistent with those in Table 1. The ellipse draws a 95% confidence level with a  $t$ -distribution.

The ADMIXTURE analysis suggested a clear and unmixed ancestry for all studied individuals ( $Q > 0.99$ ) (Figure 4). Under the  $K = 2$  scenario, *S. gemmata* showed two ancestries; one was shared by *S. monadelpha*, and the other was shared by all *S. acutisepala* individuals. Under the  $K = 3$  scenario, all three ancestries were present in *S. gemmata*, two of which were found in *S. acutisepala*. Under the  $K = 4$  scenario, a similar pattern also emerged. *S. gemmata* individuals possessed four ancestries, three of which could be found in *S. acutisepala*.



**Figure 4.** The ADMIXTURE results showing the ancestries of 11 studied individuals through the  $K = 2$  to  $K = 4$  scenarios. Under each scenario, the individual ancestries of *S. acutisepala* can all be found within *S. gemmata*, and no genetic admixtures were found among individuals (ancestry proportion ( $Q$ )  $> 0.99$ ). The accession IDs are consistent with those in Table 1.

### 3. Discussion

The taxonomic rank of *S. acutisepala* has been discussed for a long time. Some scholars recognized *S. acutisepala* as a different species from *S. gemmata* (formerly *S. sinensis*) [27,28,40], while others argued that *S. acutisepala* was a variety under *S. gemmata* (formerly *S. sinensis*) [21,22]. This dispute is commonplace in taxonomic research since presumed nascent species defined by phenotypic evidence have often been treated as subspecies [41]. More importantly, none of these taxonomic works put emphasis on the genetic differentiation between these two taxa. As supported by our phylogenomic analysis, a progenitor–derivative relationship between *S. gemmata* and *S. acutisepala* is clear (Figure 1) [10]. Compared to other plant species endemic to eastern Asia, the extent of overall genetic divergence between *S. acutisepala* and *S. gemmata* ( $F_{st} = 0.073$ ) is relatively little and comparable to a value among populations rather than between species [42,43]. For example, the inter-lineage  $F_{st}$  of *Cercidiphyllum japonicum* populations was determined to be 0.07, while that between *Cercidiphyllum japonicum* and its sister species (*Cercidiphyllum magnificum*) was inferred to be 0.72 [44]. Taking the observed genomic divergence and inferred individual ancestry together (Figures 2–4), we deduce that *S. acutisepala* is in its infant speciation from *S. gemmata*. Therefore, we recommend treating *S. acutisepala* as a variety of *S. gemmata* as indicated by genetic evidence, which is supported by FOC [22].

According to some phylogenetic species concepts, the monophyly of lineages should never be overstated [7]. However, the paraphyly of the progenitor species is not rare in peripatric or budding speciation, or speciation events that follow long-distance dispersals [45–47]. Given enough time waiting for the genetic coalescence to happen, it is predictable that gene trees for both *S. acutisepala* and *S. gemmata* will transition from paraphyly to reciprocal monophyly [10,47]. Distinguishable morphological differences have been and are considered important, even essential, in most if not all “scientific” taxonomies as well [48]. In the case of distinguishing *S. acutisepala* from *S. gemmata* in the field, a rough observation of barks may help. The bark exfoliations of *S. acutisepala* are reddish-brown and membranous, while those of *S. gemmata* are yellowish or grayish and hard papery [27,28]. Moreover, we are looking forward to the taxonomic revisions of the *S. gemmata* complex (including *S. gemmata* and *S. acutisepala*) in the upcoming editions of other regional floras by taking the genetic evidence shown here into consideration.

Besides *S. acutisepala*, *S. brevicalyx* was also presumed to be part of the *S. gemmata* complex [25]. Unfortunately, we are not able to include any *S. brevicalyx* accessions in the molecular analysis. According to the literature where the name *S. brevicalyx* was published, the author stated that it is morphologically distinct from *S. gemmata* in having widely ovate to subcordate (vs. ovate) bracts and two nearly orbicular or reniform (vs. ovate) outer sepals [20]. During our field surveys at Mt. Tianmu and its adjacent mountainous areas, where *S. brevicalyx* was reported to be exclusively found, we failed to mark any individuals that perfectly fit the descriptions in the protologue of *S. brevicalyx*. It is most likely that *S. brevicalyx* is conspecific with *S. gemmata* and the type specimen of *S. brevicalyx* only represents an extreme within the phenotypic continuums of *S. gemmata* [28], but still, we encourage a phylogenetic study on historical materials from herbarium specimens [49] to further determine the species entity and the phylogenetic position of *S. brevicalyx*.

Though limited by the type of genomic data, in which only a fraction of the genome was investigated, and the number of studied individuals, the present study provides crucial clues for future studies regarding the speciation mechanisms of *S. acutisepala*. In speciation genomics, genomic regions with higher divergence than expected (i.e., the genomic island of divergence/differentiation) were intensely explored to determine evolutionary events (such as gene flow, modes of selection, the strength of drift, and the formation of loci related to the emergence of reproductive isolation) associated with the speciation process [50–53], which has been preliminarily revealed here (Figure 2). As the cost of de novo sequencing and assembly of chromosomal-level genomes continues to decrease [54], a reference genome of the *S. gemmata* complex along with whole-genome resequencing data from its natural populations is desperately needed to reveal a comprehensive genomic landscape of *S. acutisepala*, which would undoubtedly facilitate our understanding of the speciation mechanism of this nascent species [55].

## 4. Materials and Methods

### 4.1. Plant Sampling and DNA Extraction

Healthy plant leaves were collected from ten representative individuals of the *S. gemmata* complex (including five of *S. gemmata* and five of *S. acutisepala*) and one individual of *S. monadelpha* as the outgroup (Table 1). The sampling sites of the *S. gemmata* complex can well represent its natural distribution [25]. Specifically, *S. gemmata* accessions were collected from five sites from eastern to southern China, spanning approximately ten degrees of longitude and five degrees of latitude, and *S. acutisepala* accessions were collected from three sites in Zhejiang province, eastern China (Figure 1). Leaf tissue was dried in silica gel for later DNA extraction. Voucher specimens were curated in the Herbarium of Zhejiang University (HZU).

Total genomic DNA was isolated using the Plant DNAzol™ Reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer’s protocol. The degradation degree of DNA products was assessed using agarose gel electrophoresis, and the concentration of

isolated DNA was measured with a Qubit® 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). The DNA products were frozen at  $-20^{\circ}\text{C}$  until use.

#### 4.2. RAD-Seq and Data Assembly

Qualified DNA samples (total DNA mass  $> 300$  ng, DNA concentration  $> 10$  ng/ $\mu\text{L}$ , and no evident degradation) were delivered to the sequencing institution (Novogene Co., Ltd., Beijing, China) for RAD-seq. DNA samples were fragmented by the *EcoR* I enzyme. DNA fragments of the desired length were gel-purified before adapter ligation and DNA cluster preparation. The detailed method of RAD-seq can be found in the work of Emerson et al. [56]. The RAD-seq libraries were sequenced using the Illumina HiSeq™2000 platform (Illumina, San Diego, CA, USA) to generate 150 bp paired-end reads. After sequencing, raw reads were filtered by removing adapter sequences and low-quality reads (reads with quality lower than 50% or reads in which ambiguous bases made up more than 10% of the raw read). The filtering procedure was performed at the sequencing institution with in-house scripts.

Then, we assembled the clean reads using ipyrad (v. 0.9.90) [57] with default settings. To exclude genetic data from non-nuclear inheritances (e.g., plastome), we used the genome sequence of *Camellia sinensis* [58] as the reference during the data mapping process. The ipyrad pipeline generates a set of outputs for downstream analysis, including a sequence alignment file that contains the full dataset from all assembled loci and a vcf file that stores SNP information. The alignment file would be used in phylogenetic inference, and the vcf file would be used in the analysis of population structure.

#### 4.3. Inference of Phylogenetic Tree

Based on the sequence alignment generated above, we reconstructed the consensus ML phylogenetic tree using IQ-TREE (v. 2.2.0.3) with 1000 ultra-fast bootstrap replicates (-B 1000) [59,60]. The best-fit DNA substitution model was determined automatically with ModelFinder as supported by IQ-TREE (-m MFP) [61]. The phylogenetic tree was visualized using tvBOT (v. 2.6) [62] with Bézier curves.

#### 4.4. Analysis of Genetic Divergence

Prior to the analysis of genetic divergence, we conducted data filtering on the original vcf file generated from the ipyrad pipeline using vcftools (v. 0.1.16) [63]. We retained SNP loci that met the following criteria: (1) depth ranged from 25% to 75% in depth quartiles (Figure S1); (2) biallelic SNPs; (3) SNPs covered more than 70% of all individuals (--minDP 29 --maxDP 68 --min-alleles 2 --max-alleles 2 --max-missing 0.7). Then, we estimated the genetic divergence between *S. gemmata* and *S. acutisepala* (weighted Weir and Cockerham's  $F_{st}$ ) with vcftools on both the per-site basis (--weir-fst-pop) and the windowed basis (--weir-fst-pop --fst-window-size 100 --fst-window-step 10) using vcftools.

Further, we pruned SNP sites that were subject to high levels of LD using plink (v. 1.90) (--indep-pairwise 10 10 0.4) [64]. Then, we conducted a PCA with plink. Also, we inferred individual ancestries using ADMIXTURE (v. 1.3.0) [65] through the number of clusters ( $K$ ) of two to four. The patterns of genetic divergence within the *S. gemmata* complex were visualized using the R package ggplot2 (v. 3.4.4) [66].

## 5. Conclusions

Based on genome-wide SNPs and a suite of phylogenomic analyses, our results indicate that *S. acutisepala* should be recognized as a variety of *S. gemmata* given that they represent distinct evolutionary lineages and some distinguishable phenotypic differences are already documented. *S. acutisepala* is genetically diverging from *S. gemmata*. Though paraphyly is currently present in *S. gemmata*, both taxa will be monophyletic given adequate time. The present study highlights the power of RAD-seq and SNPs in solving recalcitrant taxonomic issues, especially in the presence of nascent speciation.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/plants13101296/s1>: Figure S1: The site depth summary of SNPs generated from the ipyrad pipeline; Table S1: Morphological characteristics, distribution areas, and taxonomic treatments of the *Stewartia gemmata* complex; Table S2: Summary statistics of assembled loci among 11 accessions.

**Author Contributions:** Conceptualization, H.L. and Y.Z.; methodology, H.L. and W.L.; software, H.L. and W.L.; validation, H.L. and W.L.; formal analysis, H.L.; investigation, H.L. and Y.Z.; resources, Y.Z. and H.L.; data curation, H.L.; writing—original draft preparation, H.L.; writing—review and editing, H.L. and Y.Z.; visualization, H.L.; supervision, Y.Z.; project administration, H.L. and Y.Z.; funding acquisition, H.L. and Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Nos. 32100164 and 31461123001), the National Key Research and Development Program of China (No. 2017YFA0605104), and the Key Technology Research and Development Program of Zhejiang Province (No. 2023C03138).

**Data Availability Statement:** The data presented in this study are openly available in GenBank under the BioProject accession number PRJNA1089257.

**Acknowledgments:** The authors thank Xiaoming Zheng, Shota Sakaguchi, and Daiki Takahashi for their assistance in collecting leaf samples in this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Darwin, C. *On the Origin of Species by Means of Natural Selection*; John Murray: London, UK, 1859.
2. Mayr, E. *Systematics and the Origin of Species*; Columbia University Press: New York, NY, USA, 1942.
3. Simpson, G.G. The species concept. *Evolution* **1951**, *5*, 285–298. [[CrossRef](#)]
4. Michener, C.D. Diverse approaches to systematics. *Evol. Biol.* **1970**, *4*, 1–38.
5. Van Valen, L. Ecological species, multispecies, and oaks. *Taxon* **1976**, *25*, 233–239. [[CrossRef](#)]
6. Donoghue, M.J. A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist* **1985**, *88*, 172–181. [[CrossRef](#)]
7. De Queiroz, K. Species concepts and species delimitation. *Syst. Biol.* **2007**, *56*, 879–886. [[CrossRef](#)] [[PubMed](#)]
8. Wilkins, J.S. *Species: A History of the Idea*; University of California Press: Berkeley, CA, USA, 2009.
9. Hong, D.Y. Gen-morph species concept—A new and integrative species concept for outbreeding organisms. *J. Syst. Evol.* **2020**, *58*, 725–742. [[CrossRef](#)]
10. Freudenstein, J.V.; Broe, M.B.; Folk, R.A.; Sinn, B.T. Biodiversity and the species concept—Lineages are not enough. *Syst. Biol.* **2017**, *66*, 644–656. [[CrossRef](#)] [[PubMed](#)]
11. Wiens, J.J. Species Delimitation: New approaches for discovering diversity. *Syst. Biol.* **2007**, *56*, 875–878. [[CrossRef](#)] [[PubMed](#)]
12. Edwards, D.L.; Knowles, L.L. Species detection and individual assignment in species delimitation: Can integrative data increase efficacy? *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20132765. [[CrossRef](#)]
13. Liu, J. The integrative species concept and species on the speciation way. *Biodivers. Sci.* **2016**, *24*, 1004–1008. [[CrossRef](#)]
14. Bickford, D.; Lohman, D.J.; Sodhi, N.S.; Ng, P.K.L.; Meier, R.; Winker, K.; Ingram, K.K.; Das, I. Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* **2007**, *22*, 148–155. [[CrossRef](#)]
15. Gittenberger, E. What about non-adaptive radiation? *Biol. J. Linn. Soc.* **1991**, *43*, 263–272. [[CrossRef](#)]
16. Seehausen, O. Hybridization and adaptive radiation. *Trends Ecol. Evol.* **2004**, *19*, 198–207. [[CrossRef](#)] [[PubMed](#)]
17. Shaffer, H.B.; Thomson, R.C. Delimiting species in recent radiations. *Syst. Biol.* **2007**, *56*, 896–906. [[CrossRef](#)] [[PubMed](#)]
18. Barley, A.J.; White, J.; Diesmos, A.C.; Brown, R.M. The challenge of species delimitation at the extremes: Diversification without morphological change in Philippine sun skinks. *Evolution* **2013**, *67*, 3556–3572. [[CrossRef](#)]
19. Pinheiro, F.; Dantas-Queiroz, M.V.; Palma-Silva, C. Plant species complexes as models to understand speciation and evolution: A review of South American studies. *CRC Crit. Rev. Plant Sci.* **2018**, *37*, 54–80. [[CrossRef](#)]
20. Yan, S. On the Chinese genera *Stewartia* L. and *Hartia* Dunn. *Acta Phytotaxon. Sin.* **1981**, *19*, 462–471.
21. Li, J. A systematic study on the genera *Stewartia* and *Hartia* (Theaceae). *Acta Phytotaxon. Sin.* **1996**, *34*, 48–67.
22. Ming, T.; Bartholomew, B. *Stewartia*. In *Flora of China*; Wu, Z.Y., Raven, P., Hong, D.Y., Eds.; Science Press: Beijing, China; Missouri Botanical Garden Press: St. Louis, MO, USA, 2007; pp. 424–429.
23. Yu, X.; Gao, L.; Soltis, D.E.; Soltis, P.S.; Yang, J.; Fang, L.; Yang, S.; Li, D. Insights into the historical assembly of East Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family. *New Phytol.* **2017**, *215*, 1235–1248. [[CrossRef](#)]

24. Lin, H.Y.; Hao, Y.J.; Li, J.H.; Fu, C.X.; Soltis, P.S.; Soltis, D.E.; Zhao, Y.P. Phylogenomic conflict resulting from ancient introgression following species diversification in *Stewartia* s.l. (Theaceae). *Mol. Phylogenet. Evol.* **2019**, *135*, 1–11. [[CrossRef](#)]
25. Lin, H.Y.; Gu, K.J.; Li, W.H.; Zhao, Y.P. Integrating coalescent-based species delimitation with ecological niche modeling delimited two species within the *Stewartia sinensis* Complex (Theaceae). *J. Syst. Evol.* **2022**, *60*, 1037–1048. [[CrossRef](#)]
26. Chien, S.; Cheng, W. *Stewartia gemmata*. *Contrib. Biol. Lab. Sci. Soc. China Bot. Ser.* **1931**, *6*, 66–69.
27. Chiu, P.; Zhong, G. Brief note on the genus *Stewartia* in Zhejiang. *Plant Divers.* **1988**, *10*, 1–3.
28. Chen, Z. *Stewartia*. In *New Edition of Flora of Zhejiang*; Editorial Board of Flora of Zhejiang, Ed.; Zhejiang Science and Technology Publishing House: Hangzhou, China, 2021; pp. 312–315.
29. Zimmer, E.A.; Wen, J. Using nuclear gene data for plant phylogenetics: Progress and prospects II. Next-gen approaches. *J. Syst. Evol.* **2015**, *53*, 371–379. [[CrossRef](#)]
30. Leaché, A.D.; Oaks, J.R. The utility of single nucleotide polymorphism (SNP) data in phylogenetics. *Annu. Rev. Ecol. Evol. Syst.* **2017**, *48*, 69–84. [[CrossRef](#)]
31. Baird, N.A.; Etter, P.D.; Atwood, T.S.; Currey, M.C.; Shiver, A.L.; Lewis, Z.A.; Selker, E.U.; Cresko, W.A.; Johnson, E.A. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **2008**, *3*, e3376. [[CrossRef](#)] [[PubMed](#)]
32. Seeb, J.E.; Carvalho, G.; Hauser, L.; Naish, K.; Roberts, S.; Seeb, L.W. Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Mol. Ecol. Resour.* **2011**, *11*, 1–8. [[CrossRef](#)]
33. Andrews, K.R.; Good, J.M.; Miller, M.R.; Luikart, G.; Hohenlohe, P.A. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* **2016**, *17*, 81–92. [[CrossRef](#)] [[PubMed](#)]
34. Herrera, S.; Shank, T.M. RAD Sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Mol. Phylogenet. Evol.* **2016**, *100*, 70–79. [[CrossRef](#)]
35. Curto, M.; Schachtler, C.; Puppo, P.; Meimberg, H. Using a new RAD-sequencing approach to study the evolution of *Micromeria* in the Canary islands. *Mol. Phylogenet. Evol.* **2018**, *119*, 160–169. [[CrossRef](#)]
36. Li, L.; Zhang, J.; Lu, Z.; Zhao, J.; Li, Q. Genomic data reveal two distinct species from the widespread alpine ginger *Roscoea tibetica* Batalin (Zingiberaceae). *J. Syst. Evol.* **2021**, *59*, 1232–1243. [[CrossRef](#)]
37. Parker, E.; Dornburg, A.; Struthers, C.D.; Jones, C.D.; Near, T.J. Phylogenomic species delimitation dramatically reduces species diversity in an Antarctic adaptive radiation. *Syst. Biol.* **2022**, *71*, 58–77. [[CrossRef](#)] [[PubMed](#)]
38. Obiol, J.F.; Herranz, J.M.; Paris, J.R.; Whiting, J.R.; Rozas, J.; Riutort, M.; González-Solís, J. Species delimitation using genomic data to resolve taxonomic uncertainties in a speciation continuum of pelagic seabirds. *Mol. Phylogenet. Evol.* **2023**, *179*, 107671. [[CrossRef](#)] [[PubMed](#)]
39. Du, Z.; Cheng, J.; Xiang, Q. RAD-seq data provide new insights into biogeography, diversity anomaly, and species delimitation in eastern Asian–North American disjunct clade *Benthamidia* of *Cornus* (Cornaceae). *J. Syst. Evol.* **2024**, *62*, 1–19. [[CrossRef](#)]
40. Jin, X.; Ji, Z.; Gao, J. Validation of *Rhododendron sparsifolium* and *Stewartia acutisepala*, endemic to China. *Nord. J. Bot.* **2009**, *27*, 370–371. [[CrossRef](#)]
41. Mayr, E. Of what use are subspecies? *Auk* **1982**, *99*, 593–595.
42. Hartl, D.L.; Clark, A.G. *Principles of Population Genetics*; Sinauer: Sunderland, UK, 2007.
43. Qiu, Y.; Lu, Q.; Zhang, Y.; Cao, Y. Phylogeography of East Asia’s Tertiary relict plants: Current progress and future prospects. *Biodivers. Sci.* **2017**, *25*, 24–28. [[CrossRef](#)]
44. Zhu, S.; Chen, J.; Zhao, J.; Comes, H.P.; Li, P.; Fu, C.; Xie, X.; Lu, R.; Xu, W.; Feng, Y.; et al. Genomic insights on the contribution of balancing selection and local adaptation to the long-term survival of a widespread living fossil tree, *Cercidiphyllum japonicum*. *New Phytol.* **2020**, *228*, 1674–1689. [[CrossRef](#)] [[PubMed](#)]
45. Rieseberg, L.H.; Brouillet, L. Are many plant species paraphyletic? *Taxon* **1994**, *43*, 21–32. [[CrossRef](#)]
46. Hudson, R.R.; Coyne, J.A. Mathematical consequences of the genealogical species concept. *Evolution* **2002**, *56*, 1557–1565.
47. Anacker, B.L.; Strauss, S.Y. The geography and ecology of plant speciation: Range overlap and niche divergence in sister species. *Proc. R. Soc. B Biol. Sci.* **2014**, *281*, 20132980. [[CrossRef](#)] [[PubMed](#)]
48. Atran, S. *Cognitive Foundations of Natural History: Towards an Anthropology of Science*; Cambridge University Press: Cambridge, UK, 1990.
49. Burbano, H.A.; Gutaker, R.M. Ancient DNA genomics and the renaissance of herbaria. *Science* **2023**, *382*, 59–63. [[CrossRef](#)]
50. Harr, B. Genomic islands of differentiation between house mouse subspecies. *Genome Res.* **2006**, *16*, 730–737. [[CrossRef](#)] [[PubMed](#)]
51. Nosil, P.; Funk, D.J.; Ortiz-Barrientos, D. Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **2009**, *18*, 375–402. [[CrossRef](#)] [[PubMed](#)]
52. Cruickshank, T.E.; Hahn, M.W. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* **2014**, *23*, 3133–3157. [[CrossRef](#)]
53. Sun, Q.; Wang, R.; Chen, X. Genomic island of divergence during speciation and its underlying mechanisms. *Biodivers. Sci.* **2022**, *30*, 21383. [[CrossRef](#)]
54. Marks, R.A.; Hotaling, S.; Frandsen, P.B.; VanBuren, R. Representation and participation across 20 years of plant genome sequencing. *Nat. Plants* **2021**, *7*, 1571–1578. [[CrossRef](#)] [[PubMed](#)]
55. Bock, D.G.; Cai, Z.; Elphinstone, C.; González-Segovia, E.; Hirabayashi, K.; Huang, K.; Keais, G.L.; Kim, A.; Owens, G.L.; Rieseberg, L.H. Genomics of plant speciation. *Plant Commun.* **2023**, *4*, 100599. [[CrossRef](#)] [[PubMed](#)]

56. Emerson, K.J.; Merz, C.R.; Catchen, J.M.; Hohenlohe, P.A.; Cresko, W.A.; Bradshaw, W.E.; Holzapfel, C.M. Resolving postglacial phylogeography using high-throughput sequencing. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16196–16200. [[CrossRef](#)]
57. Eaton, D.A.R.; Overcast, I. Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* **2020**, *36*, 2592–2594. [[CrossRef](#)]
58. Zhang, X.; Chen, S.; Shi, L.; Gong, D.; Zhang, S.; Zhao, Q.; Zhan, D.; Vasseur, L.; Wang, Y.; Yu, J.; et al. Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant *Camellia sinensis*. *Nat. Genet.* **2021**, *53*, 1250–1259. [[CrossRef](#)] [[PubMed](#)]
59. Hoang, D.T.; Chernomor, O.; von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **2018**, *35*, 518–522. [[CrossRef](#)] [[PubMed](#)]
60. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)] [[PubMed](#)]
61. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermini, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589. [[CrossRef](#)]
62. Xie, J.; Chen, Y.; Cai, G.; Cai, R.; Hu, Z.; Wang, H. Tree Visualization by One Table (TvBOT): A web application for visualizing, modifying and annotating phylogenetic trees. *Nucleic Acids Res.* **2023**, *51*, W587–W592. [[CrossRef](#)]
63. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)]
64. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
65. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [[CrossRef](#)]
66. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag: New York, NY, USA, 2016.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.