

Evaluating the genetic diversity of *Erythropalum scandens* based on using inter-simple sequence repeat markers

Tianwei Yang (✉ 464082363@qq.com)

Guangxi Academy of Agricultural Science <https://orcid.org/0000-0002-2001-5440>

Xiangjun Zhang

Guangxi Academy of Agricultural Science

Shiyu Huang

Guangxi Academy of Agricultural Science

Manrong Gao

Guangxi University Guangxi Agriculture College

Ting Li

Guangxi Academy of Agricultural Science

Shangwen Zhang

Guangxi Academy of Agricultural Science

Research Article

Keywords: *Erythropalum scandens* Blume, system optimization, inter-simple sequence repeat, genetic diversity

Posted Date: November 3rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-2189943/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Genetic Resources and Crop Evolution on March 21st, 2023. See the published version at <https://doi.org/10.1007/s10722-023-01567-y>.

Abstract

Erythralum scandens Blume, an emerging medicinal plant with great potential for drug development, also possesses high edible value. In this study, we investigated the genetic diversity of the germplasm of *E. scandens* obtained from different geographical locations using inter simple sequence repeat (ISSR) markers. For this purpose, 18 ISSR primer pairs with a distinct background and adequate polymorphism were selected. We established an optimal ISSR–PCR reaction system (20 μL) with the following parameters: 1 μL DNA template (60 $\text{ng}\cdot\mu\text{L}^{-1}$), 1.2 μL primers (10 $\mu\text{mol}\cdot\mu\text{L}^{-1}$), 10 μL MasterMix, and 7.8 μL H_2O . A total of 183 loci were amplified using the 18 primer pairs, of which 121 (66.12%) indicated polymorphism. Moreover, 34 germplasms of *E. scandens* exhibited genetic similarity coefficients ranging from 0.7104 to 0.9563, genetic distances ranging from 0.0447 to 0.3420, Nei's genetic diversity index of 0.1946, and Shannon's information index of 0.2982, suggesting high intraspecific genetic diversity. UPGMA cluster and PCoA analyses distinguished the germplasm of *E. scandens* obtained from Guangxi from those collected from Guangdong, Hainan, Fujian, and Guizhou. However, the Mantel correlation analysis revealed that the genetic variation among the 34 germplasms of *E. scandens* was not significantly related to geographical distance. The analysis of the genetic background of wild and cultivated germplasms of *E. scandens* can help guide variety selection and breeding. Furthermore, the present study revealed the genetic background and affinities among 34 germplasms of *E. scandens*. Overall, our findings lay the foundation for the conservation and utilization of germplasm resources, identification and classification of varieties, and variety selection and improvement of *E. scandens* at the molecular level.

1. Introduction

Erythralum scandens Blume, a perennial vine of the genus *Erythralum* belonging to the family Olacaceae, is native to the tropical and subtropical regions of the world. In China, *E. scandens* is primarily distributed in Guangxi, Guangdong, Yunnan, and Hainan (Zhang et al. 2020). According to the Flora of China (Chinese Flora Committee 1988), *E. scandens* is pharmaceutically important, neutral, bitter, and targets the veins of the liver and kidneys. In addition, its stem possesses diuretic properties and can be used to treat jaundice, rheumatism, and ostealgia, whereas its roots dispel edema and bruises. The antigout effects of the stem and leaf extracts of *E. scandens* have been associated with the promotion of uric acid metabolism, anti-inflammation, and protection or improvement of renal function (Xu et al. 2019). Furthermore, coumarins, flavonoids, phenols, triterpenoids, polysaccharides, and other phytochemicals contained in the alcoholic extracts of *E. scandens* can significantly reduce serum uric acid and creatinine levels, thereby enhancing renal function, promoting uric acid excretion, and reducing uric acid levels in the body. The alcoholic extracts of *E. scandens* also exhibit low acute toxicity and exert significant inhibitory and protective effects in both rat and mouse models of hyperuricemia. Therefore, this species is an excellent potential alternative for the treatment and prevention of hyperuricemia (Pan et al. 2020; Huang et al. 2017).

In addition to its therapeutic applications, the young leaves of *E. scandens* are used as a vegetable with a pleasant aroma and unique flavor and can be consumed as fresh, fried, stuffed, boiled in soup, congee,

and pickled. Moreover, the nutritional value of its edible shoots is high, with protein, fat, fiber, vitamin B1, vitamin B2, and vitamin C contents comparable to those of common nutritive vegetables such as lettuce, choy sum, mustard, water spinach, sweet potato leaves, pumpkin seedlings, boxthorn leaves, and cauliflower. *E. scandens* also exhibits high levels of minerals, such as copper, zinc, iron, calcium, and phosphorus, and its calcium content is as high as 1040 mg·kg⁻¹, making it a calcium-rich vegetable (700–2300 mg·kg⁻¹) (Huang et al. 2021).

In recent years, the demand for *E. scandens* has also been increasing owing to its widely recognized edible and medicinal values. However, the supply of seedlings of *E. scandens* is insufficient to meet the present market demands, and the breeding technology is still evolving. As a result, growers primarily rely on cutting wild branches for propagation, which severely damages the wild resources and environment of *E. scandens* (Huang et al. 2021). Furthermore, varieties suitable for edible and medicinal purposes have not yet been identified and developed. In addition, few reports are available on the germplasm resources of *E. scandens*, resulting in inefficient germplasm collection, conservation, and seedling breeding. Therefore, genetic-diversity studies on the germplasm of *E. scandens* obtained from different geographical locations are necessary to improve the collection and conservation of germplasm resources, evaluation, and efficient utilization, as well as the selection and breeding of superior varieties to establish the theoretical basis for the breeding and improvement research programs on *E. scandens*.

Molecular markers—one of the most important tools used for analyzing genetic diversity—have been widely used in several plant species. The most commonly used molecular markers include simple sequence repeat (SSR), inter SSR (ISSR), start codon-targeted polymorphism (SCoT), and random amplified polymorphic DNA (RAPD) markers (Bernard et al. 2020; Sirangi et al. 2020). Of these, ISSR markers, which are used for amplifying simple, repetitive sequences, exhibit several advantages, such as high polymorphism, adequate stability, low cost, and simple operation (Shi et al. 2010; Li et al. 2020). They have been successfully applied for analyzing the genetic diversity of medicinal plants such as *Cassia tora* (Vikas and Krishna 2018), sand ginger (Subositi et al. 2020), and *Dendrobium huoshanense* (Tikendra et al. 2019).

In this study, ISSR primers were selected by optimizing the ISSR–PCR reaction system for analyzing the germplasms of *E. scandens* obtained from different geographical locations. In addition, a genetic diversity analysis was conducted to evaluate the genetic relatedness among the germplasms of *E. scandens* at the molecular level. Overall, our findings provide the foundation for the conservation and utilization of germplasm resources, identification and classification of varieties, and variety selection and improvement of *E. scandens*.

2. Materials And Methods

2.1. Sampling

The samples were obtained from 34 germplasm resources (Table 1). The geographical distribution of the germplasms is shown in Fig. 1. The phenotypes of two cultivated germplasm (*E. scandens* No. 1 and 2)

are shown in Fig. 2.

Table 1
Geographical location and number of 34 germplasm of *E. scandens*.

Germplasm No.	Source	Longitude	Latitude	
1	Daxin, Guangxi 1	E107.1793	N22.8611	Artificial cultivation
2	Daxin, Guangxi 2	E107.1827	N22.8584	Artificial cultivation
3	Baise, Guangxi	E106.5221	N23.9378	Wild Type
4	Cenxi, Guangxi	E110.9863	N22.9429	Wild Type
5	Mashan, Guangxi	E108.1807	N23.7422	Wild Type
6	Shangsi, Guangxi	E108.0072	N22.1384	Wild Type
7	Lipu, Guangxi	E110.3974	N24.5111	Wild Type
8	Hepu, Guangxi	E109.2564	N21.6761	Wild Type
9	Xingye, Guangxi	E109.8692	N22.7186	Wild Type
10	Cangwu, Guangxi	E111.5313	N23.8507	Wild Type
11	Liuzhou, Guangxi	E109.3222	N24.3421	Wild Type
12	Guiping, Guangxi	E110.0446	N23.4137	Wild Type
13	Zhaoping, Guangxi	E110.8441	N24.1786	Wild Type
14	Hechi, Guangxi	E108.0623	N24.7527	Wild Type
15	Rong'an, Guangxi	E109.4358	N25.2973	Wild Type
16	Long'an, Guangxi	E107.7237	N23.2184	Wild Type
17	Longmen, Guangdong	E114.2465	N23.7528	Wild Type
18	Xinyi, Guangdong	E110.9405	N22.3722	Wild Type
19	Huaiji, Guangdong	E112.1868	N23.9501	Wild Type
20	Haifeng, Guangdong	E115.3026	N22.9926	Wild Type
21	Luoding, Guangdong	E111.5359	N22.7741	Wild Type
22	Lianjiang, Guangdong	E110.2791	N21.6419	Wild Type
23	Baoting, Hainan	E109.7365	N18.6754	Wild Type
24	Ledong, Hainan 1	E109.1631	N18.8206	Wild Type
25	Ledong, Hainan 2	E109.1802	N18.7293	Wild Type
26	Qiongzong, Hainan 1	E109.7842	N19.0441	Wild Type
27	Qiongzong, Hainan 2	E109.8284	N19.0456	Wild Type
28	Anxi, Fujian	E118.2247	N25.1091	Wild Type

29	Germplasm No.	Fuqing, Fujian 1	E119.3792	N25.7962	Wild Type
30		Fuqing, Fujian 2	E119.3677	N25.7891	Wild Type
31		Luodian, Guizhou	E106.7623	N25.3987	Wild Type
32		Libo, Guizhou	E107.7951	N25.4448	Wild Type
33		Anlong, Guizhou 1	E105.4569	N25.1375	Wild Type
34		Anlong, Guizhou 2	E105.4877	N25.1548	Wild Type

In this study, we used the ISSR primers UBC800–864 out of the 100 primers published by the University of British Columbia, Canada (<http://www.biotech.ubc.ca/services/naps/primers.html>).

2.2. Genomic DNA extraction

Genomic DNA was extracted using a modified CTAB method (Safeena et al. 2021), and its quality was examined using 0.8% agarose gel electrophoresis. The concentration and purity of DNA were determined using an ultra-micro UV–Vis spectrophotometer (Thermo NanoDrop One, America). Subsequently, high-quality DNA was diluted to appropriate concentrations for PCR amplification, and the remaining DNA was stored at -20°C .

2.3. Optimization of the ISSR–PCR amplification

The three factors affecting the ISSR–PCR amplification, including the amount of template, primers, and 2× ES Taq MasterMix (CW0690M; Kangwei Century Biotechnology Co. Ltd. China), were optimized using the L9 (34) orthogonal test table (Table 2). The PCR (20 μL) conditions were as follows: pre-denaturation at 94°C for 3 min, denaturation at 94°C for 30 s, annealing at 57°C for 30 s (35 cycles), and extension at 72°C for 5 min. The results of the orthogonal test were analyzed using intuitive analysis (He et al. 1998). Finally, the number and definition (clarity) of amplified bands were scored to identify the most suitable reaction conditions.

Table 2
Orthogonal test factors and levels.

Level	Factor		
	DNA (ng)	Primer ($10\ \mu\text{mol}\cdot\mu\text{L}^{-1}$, μL)	2×Es Taq Master Mix (μL)
1	30	0.6	9
2	40	0.8	10
3	50	1.0	11
4	60	1.2	12

2.4. Primer screening

We performed the initial screening of 64 ISSR primers using the optimized ISSR–PCR system. The primers that resulted in unclear bands in the initial screening were subjected to gradient PCR (annealing

temperature: 49°C, 51°C, 53°C, 55°C, 57°C, and 59°C) for re-screening. The primers with clear bands and high polymorphism in the two rounds of screening were selected for validating the ISSR–PCR system.

2.5. PCR amplification

The reaction conditions for ISSR–PCR amplification was as follows: initial denaturation at 94°C for 2 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing for 30 s (annealing temperature of primers), extension at 72°C for 30 s, and final extension at 72°C for 5 min. Finally, the results were analyzed using agarose gel electrophoresis.

2.6. Data processing

The number and position of DNA bands were converted into binary data (0 and 1), where the presence or absence of bands at the same locus with the same mobility was recorded as “1” and “0,” respectively, to form a (0, 1) matrix. Thereafter, the genetic diversity indexes were calculated and clustering analysis was performed for 34 germplasms of *E. scandens* using PopGene 32, GenAlEx 6.51b2 and STRUCTURE software. Finally, the correlation between geographical location and germplasm was determined using the RStudio software.

3. Results

3.1. Quality and quantity of genomic DNA extracted from *E. scandens* germplasms

The optical density (A₂₆₀/A₂₈₀) and concentration of genomic DNA isolated from 34 germplasms of *E. scandens* were in the range of 1.81–1.97 and 298.1–1693.9 ng·μL⁻¹, respectively. Agarose gel electrophoresis depicted single bands with no smears for all samples, indicating that the extracted DNA was of good quality and could be used for ISSR analysis (Fig. 3).

3.2. Optimization of the ISSR–PCR system

The PCR amplification results suggested that the different orthogonal combinations amplified bands with large differences in the number and the degree of clarity (Fig. 4). Among them, treatment 16 exhibited the best amplification, with the highest score of 16 in the visual analysis method (Table 3). In general, a good amplification profile resulted in high scores. The mean *K_i* of each factor at different levels and the extreme difference of the mean between different levels of the same factor was derived from the scoring result *R*. The degree of influence of the three factors on the ISSR–PCR amplification of *E. scandens* germplasms was ranked using the *R*-value: DNA template amount > primer amount > Mix mixture amount. Finally, the following optimal reaction composition (20 μL) was determined using the *K_i* value: 1 μL DNA template (60 ng·μL⁻¹), 1.2 μL primer (10 μmol·μL⁻¹), 10 μL MasterMix, and 7.8 μL H₂O.

Table 3
Combinations of orthogonal tests and derived scores.

Treatment No.	DNA (ng)	Primer (10 $\mu\text{mol}\cdot\mu\text{L}^{-1}$, μL)	2×ES Taq Master Mix (μL)	Score
T 1	30	1.0	12	11.2
T 2	30	0.6	9	10.8
T 3	30	1.2	10	12.0
T 4	30	0.8	11	11.7
T 5	40	0.8	12	11.2
T 6	40	0.6	10	10.2
T 7	40	1.2	9	11.7
T 8	40	1.0	11	11.8
T 9	50	0.8	9	12.0
T10	50	0.6	11	12.2
T11	50	1.2	12	13.2
T12	50	1.0	10	14.2
T13	60	0.6	12	14.2
T14	60	0.8	11	14.0
T15	60	1.0	9	15.0
T16	60	1.2	10	15.7
K ₁	11.4	11.8	12.4	
K ₂	11.2	12.2	13.0	
K ₃	12.9	13.0	12.4	
K ₄	14.7	13.1	12.4	
R	3.5	1.3	0.6	

Note: K1, K2, K3, and K4 are the mean values of scores at different levels for each factor, and R is the extreme difference in the mean values of scores between different levels for the same factor.

3.3. Polymorphic amplification and genetic diversity analyses using ISSR markers

The initial screening of primers UBC800–864 and the re-screening of a few primers (using gradient PCR) indicated the primer pairs and annealing temperatures suitable for ISSR–PCR amplification of *E. scandens* germplasms (Table 4). A total of 18 primer pairs resulting in bands with distinct backgrounds, adequate

stability, and high polymorphism were identified. In total, 183 markers were amplified using the 18 ISSR primers, including 121 polymorphic markers (66.12%), with an average of 10.2 markers and 6.7 polymorphic markers per primer pair, an average allele number (N_a) of 1.66, and an average effective allele number (N_e) of 1.32.

Table 4
Statistics of ISSR primer pairs and their polymorphism and genetic diversity indexes.

ISSR Primer	Annealing temperature	Total site number	Polymorphic site	Percentage of polymorphic site	N_a	N_e	H	I
UBC808	59°C	8	2	25.00%	1.25	1.13	0.0696	0.1032
UBC811	57°C	8	6	75.00%	1.75	1.33	0.2037	0.3196
UBC812	53°C	9	5	55.56%	1.56	1.30	0.1797	0.2738
UBC822	51°C	9	6	66.67%	1.67	1.32	0.1940	0.3000
UBC823	51°C	10	7	70.00%	1.70	1.42	0.2403	0.3601
UBC824	49°C	12	8	66.67%	1.67	1.29	0.1746	0.2744
UBC825	51°C	11	6	54.55%	1.55	1.30	0.1771	0.2671
UBC826	59°C	6	3	50.00%	1.50	1.22	0.1491	0.2360
UBC827	49°C	13	12	92.31%	1.92	1.39	0.2365	0.3686
UBC834	59°C	12	9	75.00%	1.75	1.46	0.2676	0.4005
UBC835	51°C	11	6	54.55%	1.55	1.19	0.1291	0.2079
UBC836	55°C	12	8	66.67%	1.67	1.26	0.1696	0.2714
UBC842	55°C	11	9	81.82%	1.82	1.43	0.2381	0.3565
UBC849	49°C	7	3	42.86%	1.43	1.22	0.1290	0.1943
UBC850	55°C	10	7	70.00%	1.70	1.32	0.1815	0.2795
UBC855	55°C	10	7	70.00%	1.70	1.38	0.2144	0.3235
UBC856	55°C	13	10	76.92%	1.77	1.39	0.2308	0.3528
UBC859	49°C	11	7	63.64%	1.64	1.40	0.2279	0.3397
Total		183	121	/	/	/	/	/
Average		10.2	6.7	64.29%	1.66	1.32	0.1946	0.2982

The genetic diversity indexes of Guangxi, Guangdong, Hainan, Fujian, and Guizhou regions are shown in Table 5. We observed that the number of samples, Nei's genetic diversity index (H), and Shannon's information index (I) were higher in germplasms collected from Guangxi than those collected from other regions.

Table 5
Statistics of genetic diversity indicators for each regional subgroup.

Region	Sample size	Polymorphic site	Na	Ne	H	I
Guangxi	16	100	1.55	1.31	0.1838	0.2761
Guangdong	6	76	1.42	1.28	0.1574	0.2317
Hainan	5	66	1.36	1.25	0.1383	0.2029
Fujian	3	41	1.22	1.16	0.0908	0.1325
Guizhou	4	55	1.30	1.18	0.1077	0.1615

3.4. Correlation between genetic similarity coefficients and genetic distances of *E. scandens* germplasms

The ISSR data of 34 germplasms of *E. scandens* were analyzed using PopGen32 (Fig. 5). The genetic similarity coefficients and genetic distances among the germplasms of *E. scandens* based on ISSR markers were 0.7104–0.9563 (mean: 0.7995) and 0.0447–0.3420 (mean: 0.2247), respectively. The lowest genetic similarity coefficient was observed for the germplasms obtained from both Cengxi City and Shangsi County, Guangxi (0.7104) with Ledong County, Hainan, and the germplasm obtained from them showed the maximum genetic distance (0.3420), indicating a high genetic variation and distance between them. The genetic similarity coefficient between the two germplasms collected from Qiongzhang County, Hainan, was the highest (0.9563), and their genetic distance was the lowest, indicating less genetic variation and relatedness between them.

3.5. ISSR marker-based clustering analysis

The genetic similarity coefficients of the 34 germplasms of *E. scandens* were analyzed using UPGMA clustering and mapped using the software PopGen32 and MEGA7. The results of cluster analysis (Fig. 6) revealed that the 34 germplasms of *E. scandens* were classified into three clusters. Cluster 1 contained 93.75% of Guangxi and 33.3% of Fujian germplasms, cluster 3 contained 50% of Guangdong and 20% of Hainan germplasms, and cluster 2 included germplasms collected from Guangxi (6.25%), Guangdong (50%), Hainan (80%), Fujian (66.7%), and Guizhou (100%) regions. The genetic similarity coefficients among the germplasms of *E. scandens* in Guangxi were high, with a few differences in genetic backgrounds among the germplasms of *E. scandens* from Guangdong, Hainan, Fujian, and Guizhou.

3.6. ISSR marker-based principal coordinates analysis (PCoA)

The ISSR data of the 34 germplasms of *E. scandens* were transformed into genetic distances for PCoA analysis using the GenAlEx software (Fig. 7). The two-dimensional PCoA plot of the 34 germplasm resources of *E. scandens* was constructed using the first and second principal components as the horizontal and vertical coordinates, respectively. The two principal components explained 10.45% and

8.87% of the variations, respectively. In addition, the germplasms of *E. scandens* were divided into five groups based on geographical locations, namely Guangxi, Guangdong, Hainan, Fujian, and Guizhou, and the germplasm of each group can be roughly clustered together. The germplasms of *E. scandens* from Guangxi could be distinguished from those collected from Guangdong, Hainan, Fujian, and Guizhou, whereas those from Guangdong, Hainan, Fujian, and Guizhou were not strictly divided by geographical location.

3.7. Structure analysis

The Bayesian clustering model was performed to evaluate the population structure of 34 *E. scandens* samples. The optimal cluster value (K) was two (Fig. 8a), with the highest values of both $\text{LnP}(K)$ (log probability of data, -1921.17) and ΔK (5.27) obtained from the STRUCTURE. The results showed that it was more appropriate to divide 34 germplasms into two groups. Two run results ($K = 2$ and $K = 3$) are shown in Fig. 8b, the main difference between the results of $K = 2$ and $K = 3$ is the germplasm division in Guangdong, Fujian, Hainan and Guizhou. Structure analysis results are generally consistent with UPGMA clustering analysis.

3.8. Correlation analysis of genetic and geographical distances

The geographical distances between sampling sites were calculated using RStudio based on the latitude and longitude information of each sampling site. Subsequently, the correlation between geographic and genetic distances among different germplasms was analyzed using Mantel's correlation test [Mantel's $r = 0.1176$; $P = 0.084$] (Fig. 9)], and no significant correlation between geographic and genetic distances was observed among germplasms of *E. scandens* obtained from different locations.

4. Discussion

The study of genetic diversity is an important way to understand the genetic variation and relatedness of germplasm resources, which can promote the effective conservation, management, and utilization of medicinal plant species (Soltis and Soltis 1991). *E. scandens* is the only species belonging to the genus *Erythrolalum* of the family Olacaceae. Field visits and research revealed that most of the *E. scandens* germplasms were extremely similar in morphology, making it difficult to distinguish between different germplasms. In this regard, DNA molecular markers, one of the most important techniques for examining genetic diversity, can rapidly identify and characterize germplasms that cannot be accurately distinguished phenotypically, regardless of the developmental stage and growing environment of the plant (Shahid et al. 2012). Each molecular marker has its advantages and disadvantages. However, ISSR markers combine the advantages of several molecular markers such as high polymorphism, reproducibility, low cost, ease of use, and no need for genome sequencing (Singh et al. 2021). Therefore, the analysis of the genetic diversity of 34 germplasms of *E. scandens* from different locations using ISSR markers provided the foundation for the identification of *E. scandens* germplasm resources for breeding.

Accurate and effective application of ISSR marker technology requires the establishment of a stable and reliable ISSR–PCR amplification (Deng et al. 2019). However, ISSR–PCR amplification is affected by several factors, and each factor exhibits different effects on the system based on the sample (Li et al. 2020; Huang et al. 2011; Mohamad et al. 2017; Jamil et al. 2022).

The widespread use of PCR MasterMix has resulted in a shift from the traditional five-component PCR reaction system (Mg^{2+} , DNA template, primers, DNA polymerase, and dNTPs) to the three-component PCR system (MasterMix, DNA template, and primers), thereby reducing the additional steps and making PCR results highly stable and reliable with enhanced reproducibility (Khodaei et al. 2021; Akhtar et al. 2021; Sheikh et al. 2021). In this study, the optimization of the ISSR–PCR system was performed using a three-component, four-level orthogonal test. Finally, the optimal ISSR–PCR system (20 μ L) was established as follows: 1 μ L DNA template ($60\text{ ng}\cdot\mu\text{L}^{-1}$), 1.2 μ L primer ($10\text{ }\mu\text{mol}\cdot\mu\text{L}^{-1}$), 10 μ L MasterMix, and 7.8 μ L H_2O .

In total, 18 primer pairs were identified from 64 ISSR universal primers using the optimized ISSR–PCR system. The ISSR primer pairs with high polymorphism and adequate stability were used for analyzing the genetic diversity of *E. scandens*. Thus, in this study, PCR amplification was performed for 34 *E. scandens* germplasm with different geographical distributions using 18 pairs of ISSR primers, followed by data transformation for genetic diversity analysis. We amplified a total of 121 polymorphic markers (66.12%).

High genetic diversity indexes (H and I) indicated high genetic diversity and an increasingly complex genetic background (Zhang et al. 2011). In addition, the germplasms of *E. scandens* exhibited H and I values of 0.1946 and 0.2982, respectively, indicating some genetic variation among the germplasms from Guangxi, Guangdong, Hainan, Fujian, and Guizhou regions of China, with high genetic diversity. Furthermore, H and I were higher for Guangxi germplasms than for other regions, according to the statistical results of genetic diversity indexes based on different provinces.

According to the UPGMA clustering analysis of genetic similarity coefficients of the 34 germplasms of *E. scandens*, the germplasms of *E. scandens* for Guangxi were classified in the same cluster and could be distinguished from those in Guangdong, Hainan, Fujian, and Guizhou. Subsequently, PCoA was performed using the genetic distances among the 34 germplasms of *E. scandens*. The coordinates of the germplasm of *E. scandens* from each province could be roughly distributed together. The germplasms from Guangxi could be separated from those from other provinces. In contrast, the germplasms from Guangdong, Hainan, Fujian, and Guizhou were cross-distributed on the main coordinate map and did not exhibit an obvious distribution pattern, which was consistent with the results of UPGMA cluster analysis. The results of the Mantel correlation test revealed no significant correlation between the geographical and genetic distances among the germplasms. Moreover, the genetic variation among the 34 germplasms of *E. scandens* was not significantly related to the geographical distance, which was consistent with the results of the genetic diversity of the homolog of *E. scandens*—*Malania oleifera* Chun & S.K. Lee (Lai 2006).

The kinship (genetic) analysis can rapidly screen target germplasms with similar genetic backgrounds from numerous germplasm resources, which improves the screening efficiency of good germplasms (Sulima et al. 2017; Gupta et al. 2021) and facilitates the rapid advancement in breeding. In this study, 34

germplasms of *E. scandens* included two cultivated and 32 wild germplasms. The two cultivated germplasm (*E. scandens* No. 1 and 2) were obtained from the wild germplasm in Daxin County, Guangxi. *E. scandens* No. 1 exhibited outstanding growth potential and number of buds, whereas *E. scandens* No. 2 exhibited distinctive aroma and taste (Zhang et al. 2020). The results of cluster analysis suggested that *E. scandens* No. 1 could be distinguished from other germplasms from Guangxi, indicating that the genetic background of *E. scandens* No. 1 was different from those of other germplasms, which may be associated with its high-yield and high-quality germplasm characteristics. Nonetheless, the germplasm *E. scandens* No. 2 and those from Guangxi Lipu City and Guangxi Hepu County (No. 7 and 8) were clustered into a small category, indicating that the genetic backgrounds of these three germplasms were closely related and may exhibit similar phenotypic characteristics. Therefore, we must focus on the phenotypic traits related to aroma and taste in the two wild germplasms, *E. scandens* No. 7 and 8.

The medicinal components of *E. scandens* are primarily attributed to old stems and roots (Xu et al. 2019; Huang et al. 2021). Therefore, phenotypic characteristics, such as root biomass, stem thickness, and medicinal components, must be analyzed. In contrast, edible components are primarily distributed in newly sprouted shoots and branches (Long et al. 2017), which can be associated with phenotypic characteristics such as shoot number, taste, and aroma. The combination of phenotypes in future studies and the database of ISSR markers used in this study can further improve the screening efficiency of excellent germplasms.

The collection of *E. scandens* germplasm resources is difficult owing to their narrow distribution and severely damaged wild resources. Several locations from where the data were recorded previously are no longer available. Therefore, the ecological conservation of the germplasm resources of *Erythronium* is necessary. The number of germplasm samples and the level of genetic diversity among germplasms found in the wild in Guangxi were higher than those collected from Guangdong, Hainan, Fujian, and Guizhou. Nonetheless, the germplasms from each region can be propagated and transplanted to enhance the genetic exchange between regions and enrich genetic diversity. However, the samples collected in the present study were limited. Thus, future studies must focus on collecting and analyzing more germplasms to enrich the germplasm resource database. In addition, an extensive analysis of multiple molecular markers should be conducted to comprehensively and accurately evaluate the genetic diversity and relatedness of *E. scandens*, thereby facilitating the screening of excellent germplasms and meeting the increasing market demands.

5. Conclusions

In the present study, an optimal ISSR–PCR system was established for analyzing the genetic diversity of *E. scandens*. Eighteen ISSR primer pairs that could amplify genetic markers with distinct backgrounds and adequate polymorphism in 34 germplasms of *E. scandens* were selected. The germplasm resources from Guangxi, Guangdong, Hainan, Fujian, and Guizhou in China exhibited a relatively high level of genetic diversity. However, a few genetic differences were observed between the germplasms of Guangxi and those of Guangdong, Hainan, Fujian, and Guizhou, but the genetic variation among the 34 germplasms

was not significantly correlated with geographical distance. Thus, the 34 *E. scandens* germplasms could be propagated and transplanted onto each other to enhance genetic exchange between regions and improve genetic diversity in each region, thereby elevating the overall genetic diversity of *E. scandens*. The kinship analysis of wild and cultivated germplasms of *E. scandens* indicated that wild germplasms exhibited genetic backgrounds similar to those of cultivated varieties, which could be used for variety selection and breeding. Thus, in this study, we used ISSR markers to study the genetic diversity of *E. scandens* and revealed the genetic background and affinities of 34 *E. scandens* germplasms. The findings of this study lay the foundation for further research focusing on the conservation and utilization of *E. scandens* germplasm resources, identification and classification of varieties, and variety selection and improvement at the molecular level.

Declarations

Author Contributions: Yang, T.W. and Zhang, S.W.: Conceptualization, Methodology, Experimental design, Formal analysis, Writing—Original draft preparation; Huang S.Y.: Resources, Investigation; Gao M.R., Li, T.: Software, Data curation, Visualization; Zhang X.J.: Writing—Review and Editing, Funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Guangxi Key Research and Development Program, grant number Guike21220042; Basic scientific research business special project of Guangxi Academy of Agricultural Sciences, grant number Guinongke2017YM27; Basic scientific research business special project of Guangxi Academy of Agricultural Sciences, grant number Guinongke2022JM51; Guangxi Characteristic Crop Test Station 'Guangxi Long'an Chinese Herbal Medicine Test Station', grant number GuiTS2022002; Guangxi Soil and Water Conservation Society's innovation projects in key areas, grant number 202009001 and The APC was funded by Guangxi Key Research and Development Program.

Acknowledgments: We thank Nanning Current Science Biotechnology Co., Ltd. for editing this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Akhtar N, Hafiz IA, Hayat MQ, Potter D, Abbasi NA, Habib U, Hussain A, Hafeez H, Bashir MA, Malik SI (2021) ISSR-Based Genetic Diversity Assessment of Genus *Jasminum* L. (Oleaceae) from Pakistan. *Plants* (Basel, Switzerland) 10:1270-1270.
2. Bernard A, Barreneche T, Lheureux F, Dirlewanger E (2020) SSR genetic diversity assessment of the INRAE's walnut (*Juglans* spp.) germplasm collection. *Acta Hort* 1297:377-384.
3. Chinese Flora Committee (1988) Chinese Academy of Sciences Flora of China Science Press: Beijing, China 1988, 24:46.

4. Deng Q, Deng QX, Liu L, Tao L (2019) DNA extraction and optimization of ISSR-PCR reaction system for *pyracantha*. IOP Conf Ser: Earth Environ Sci 237:052025.
5. Gupta P, Mishra A, Lal RK, Dhawan SS (2021) DNA fingerprinting and genetic relationships similarities among the accessions/species of *Ocimum* using SCoT and ISSR markers system Mol Biotechnol 63:1-12.
6. He ZW, Liu YS, Chen LH, Cao MH (1998) Orthogonal design-direct analysis for pcr optimization. Bulletin of Hunan Medical University 04:76-77.
7. Huang SY, Zhang XJ, Li T, Zhang SW (2021) The current situation and development measures of Guangxi's industry of *Erythralum scandens* as a vegetable with edible and medicinal uses. Chinese Cucurbits and Vegetables 34:109-115.
8. Huang XX, Hu J, Wang Y, Song SD, Zhu YF, Zhu SJ (2011) Optimization of ISSR-PCR system for cultivar verification in watermelon (*Citrullus lanatus* var. *lanatus*). Seed Sci Technol 39:293-302.
9. Huang YH, Li XX, Pan QD, Li WN, Huang YN (2017) Acute toxicity and effects of *Erythralum scandens* Bl ethanol extracts on hyperuricemic in mice. Chinese Journal of Ethnomedicine and Ethnopharmacy 26:52-54.
10. Jamil A, Razzaq K, Rajwana IA, Naz A, Akhtar G, Ullah S, Faried HN, Hussain SB, Li YZ, Amin M, Sher MA, Altaf MT, Hassan H, Ahmed MAA, Alfarraj S, Ansari MJ (2022) Characterization of indigenous phalsa (*Grewia subinequalis*) genotypes using morphological traits and ISSR markers. J King Saud Univ Sci 34:102237.
11. Khodae L, Azizinezhad R, Etminan AREza, Khosroshahi M (2021) Assessment of genetic diversity among Iranian *Aegilops triuncialis* accessions using ISSR, SCoT and CBDP markers. J Genet Eng Biotechnol 19:1-9.
12. Lai JY (2006) Study on conservation biology of rare and precious plant *Malania oleifera*. PhD, Sichuan University, Chengdu China pp 114-115.
13. Li JJ, Peng JQ, Wu Y, Wang YL, Liao TZ, Ma B, Xue C, Yu LQ, Xie QY, Wang HY, Cao JW (2020) Genetic diversity among the species of *Michelia* in China using ISSR. Int J Agric Biol 24:413-419.
14. Long WG, Li SP, An JC, Zhu CS (2017) Analysis and evaluation of nutritional components in *Erythralum scandens* Blume. Food Research And Development 38:124-127.
15. Mohamad A, Alhasnawi AN, Kadhimi AA, Isahak A, Radziah CMZC (2017) DNA Isolation and Optimization of ISSR-PCR Reaction System in *Oryza sativa* L. Int J Adv Sci Eng Inf Technol 7:2264-2272
16. Pan QD, Huang YH, Mo XX, Tan WT, Lu HF (2020) Pretest and TLC analysis the chemical components in stems of *Erythralum scandens* Bl. Chinese Journal of Ethnomedicine and Ethnopharmacy 29:16-21+26.
17. Safeena MIS, Dissanayake Y, Zakeel MCM, Warnakula L, Cooray R, Dayarathna DARK (2021) An improved method for efficient recovery of high quality DNA from date palm (*Phoenix dactylifera* L , *Arecaceae*). MethodsX 8:101384-101384.

18. Shahid M, Srivastava M, Ali Sobia, Pathak N, Rastogi S, Ali SS (2012) An overview of molecular marker techniques. *Adv Life Sci* 1:13-19.
19. Sheikh ZN, Sharma V, Shah RA, Sharma N, Summuna B, Al-Misned FA, El-Serehy HA, Mir JI (2021) Genetic diversity analysis and population structure in apricot (*Prunus armeniaca* L.) grown under north-western himalayas using ISSR markers. *Saudi J Biol Sci* 28:5986-5992.
20. Shi A, Kantartzi S, Mmbaga M, Chen P (2010) Development of ISSR PCR markers for diversity study in dogwood (*Cornus* spp.). *Agric Biol J North Am* 1:189-194.
21. Singh WA, Singh NS, Devi EJ, Handique PJ, Devi HS (2021) Collection and characterization of banana gene pools (*Musa* spp.) in Manipur (NE India) using PCR–RFLP and RAPD and ISSR markers *Braz J Bot* 44:1-14.
22. Sirangi S, Jogam P, Nemali G, Ajmeera R, Raju VS (2020) Intraspecific genetic variation in *Corynandra chelidonii* (Angiosperms: Cleomaceae) as revealed by SCoT, ISSR and RAPD analyses *J Plant Biotechnol* 47:289-297.
23. Soltis PS, Soltis DE (1991) Genetic variation in endemic and widespread plant species. *Aliso* 13:215-223.
24. Subositi D, Kurnianingrum N, Mujahid R, Widiyastuti Y (2020) *Kaempferia galanga* L. A medicinal plant used by Indonesian ethnic groups: genetic diversity based on Inter-Simple Sequence Repeats (ISSR). *Agrivita* 42:45-52.
25. Sulima P, Prinz K, Przyborowski JA (2017) Genetic Diversity and Genetic Relationships of Purple Willow (*Salix purpurea* L.) from Natural Locations. *Int J Mol Sci* 19:105.
26. Tikendra L, Amom T, Nongdam P (2019) Molecular genetic homogeneity assessment of micropropagated *Dendrobium moschatum* Sw. - A rare medicinal orchid using RAPD and ISSR markers. *Plant Gene* 19:100196-100196.
27. Vikas K, Krishna RB (2018) Population authentication of the traditional medicinal plant *Cassia tora* L. based on ISSR markers and FTIR analysis. *Sci Rep* 8:10714.
28. Xu CY, Wei GY, Zhu D, Wang LQ, Zhou QM, Jiang WZ (2019) Experimental study on the anti-gout effect of aqueous extract from the stems and leaves of *Erythralum scandens*. *China Pharmacy* 30:3418-3422.
29. Zhang SW, Li T, Shi Q, Pan YN, Yu WH, Meng P, Zhang XJ (2020) New *Erythralum scandens* Blume varieties with medicinal and edible vegetable— ‘Guichicangteng no1’ and ‘Guichicangteng no2’. *China Vegetables* 10:92-95.
30. Zhang W, Luo JX, Gu YJ, Hu TX (2011) Allozyme variation of genetic diversity in natural populations of *Jatropha curcas* L. germplasm from different areas in southwest China. *Chinese Journal of Plant Ecology* 35:330-336.

Figures



Figure 1

Geographical distribution of *E. scandens* germplasm. Red, orange, purple, yellow, and blue dots represent Guangxi, Guangdong, Hainan, Fujian region, and Guizhou, respectively.



Figure 2

Phenotypes of two cultivated germplasm (*E. scandens* No. 1 and 2)

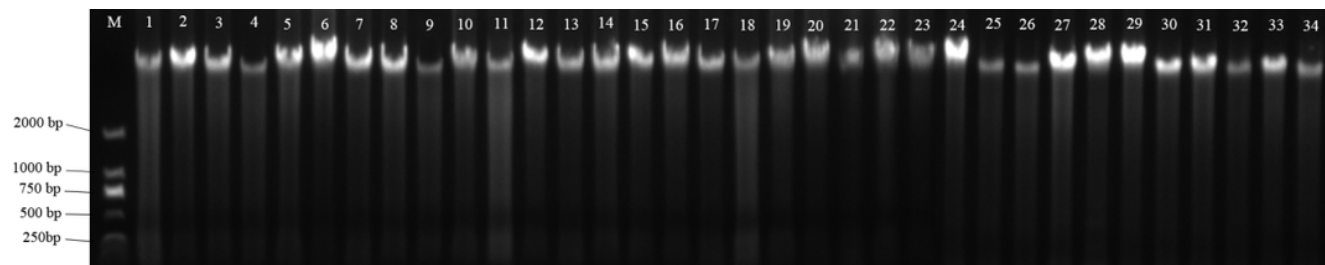


Figure 3

Agarose gel electrophoresis of 34 germplasms of *E. scandens*. M, DL2000 DNA Marker; 1 – 34, 34 germplasms of *E. scandens*.

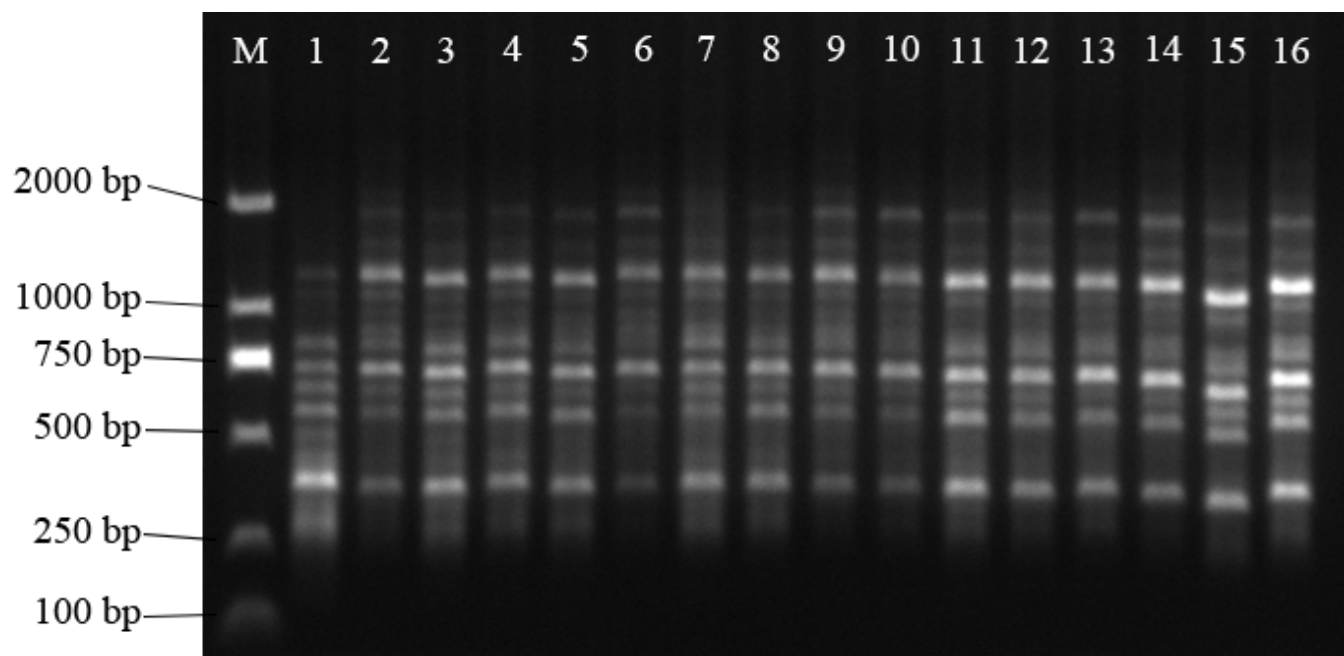


Figure 4

Orthogonal test and electrophoretic detection of ISSR-PCR amplification. M, DL 2000 DNA Marker; T 1–16, Treatments 1–16.

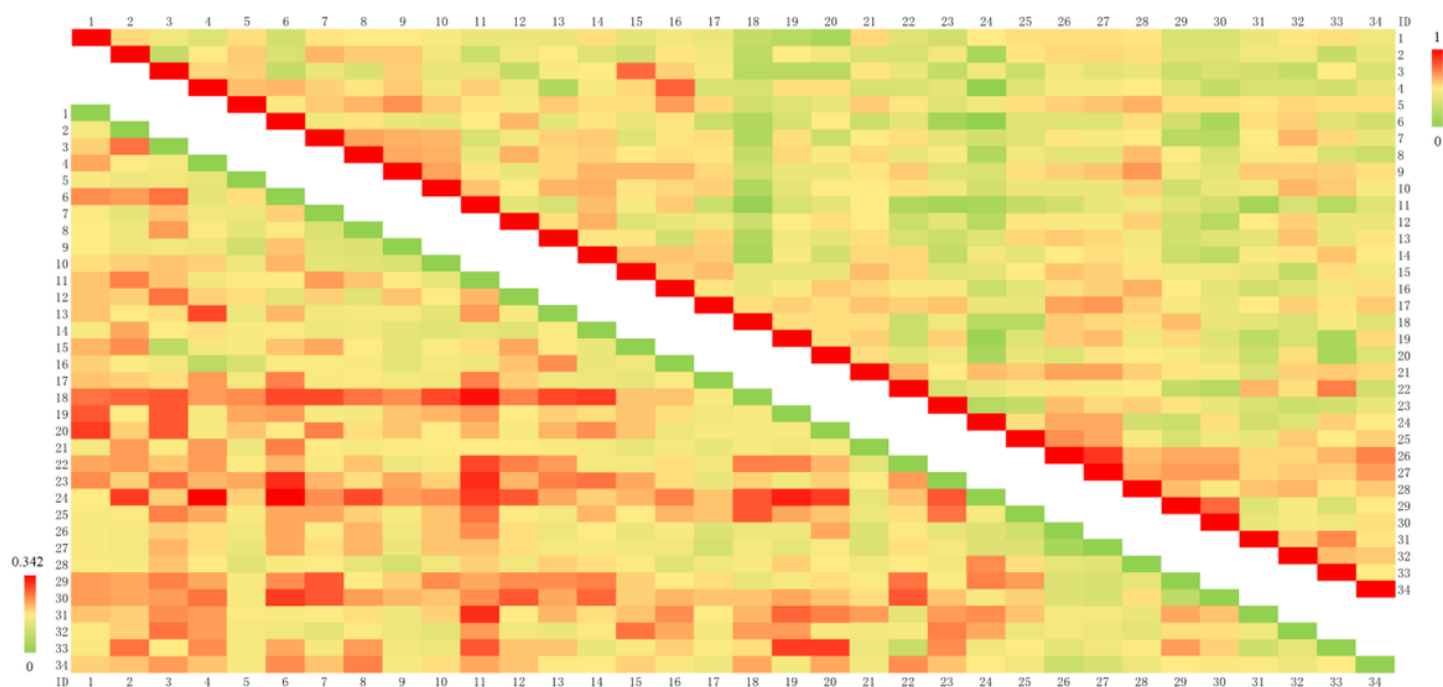


Figure 5

Genetic distance and similarity coefficient matrix of 34 germplasms of *E. scandens* based on ISSR markers. Genetic similarity coefficient is indicated above the diagonal line and genetic distance below it; 1–34, 34 germplasms of *E. scandens*.

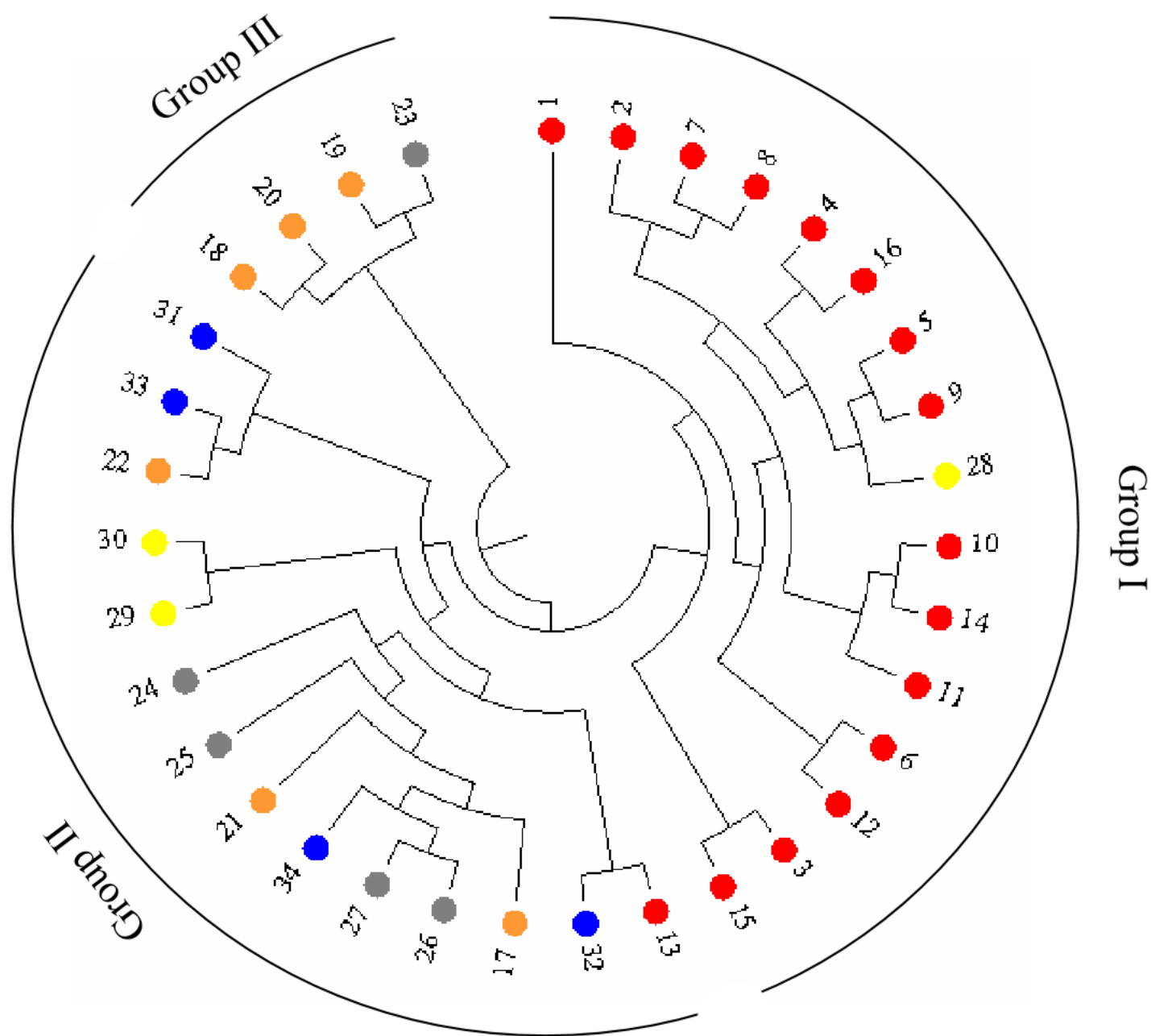


Figure 6

ISSR marker-based UPGMA clustering analysis. 1–34, 34 germplasms of *E. scandens*; Red, orange, purple, yellow, and blue dots represent Guangxi, Guangdong, Hainan, Fujian region, and Guizhou, respectively.

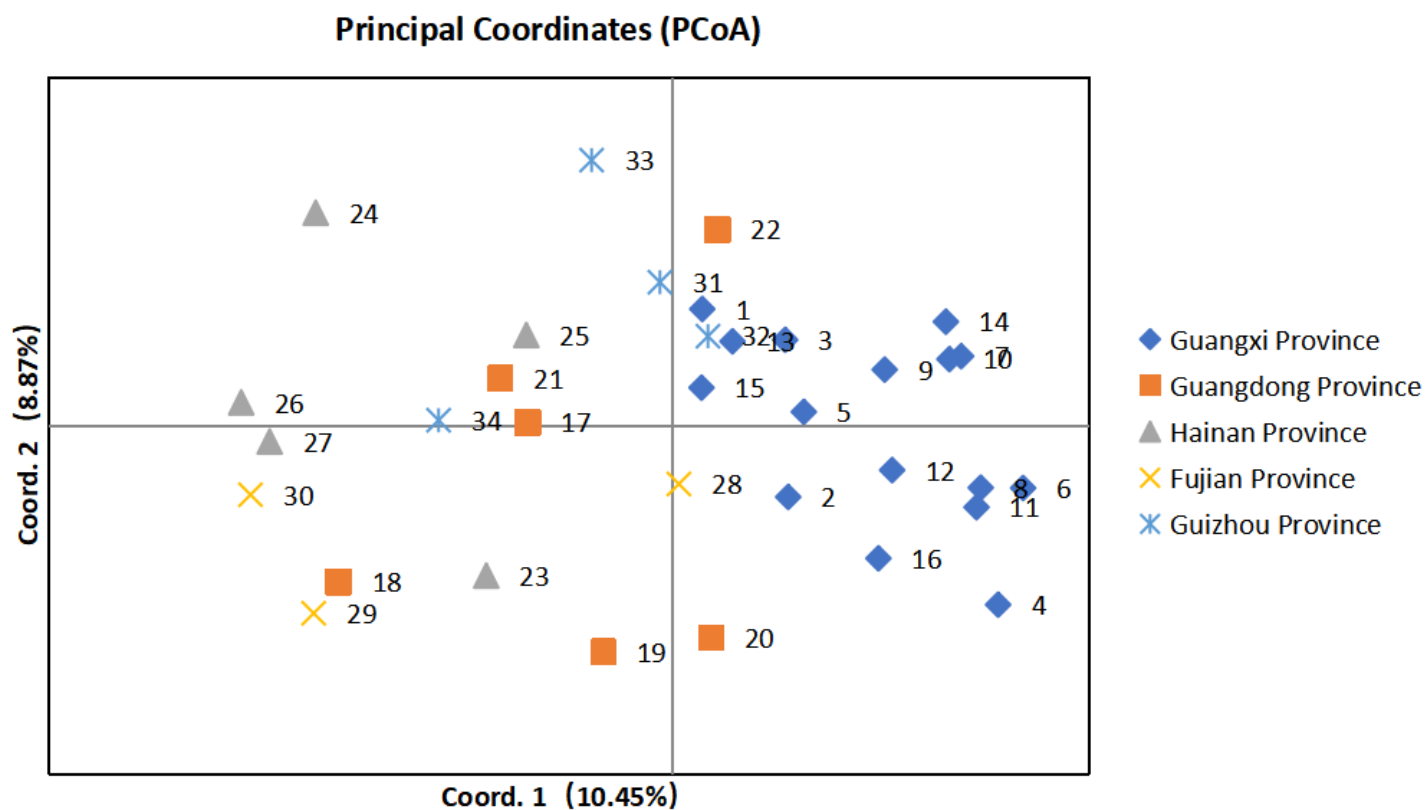


Figure 7

PCoA analysis based on genetic distance. 1–34, 34 germplasms of *E. scandens*.

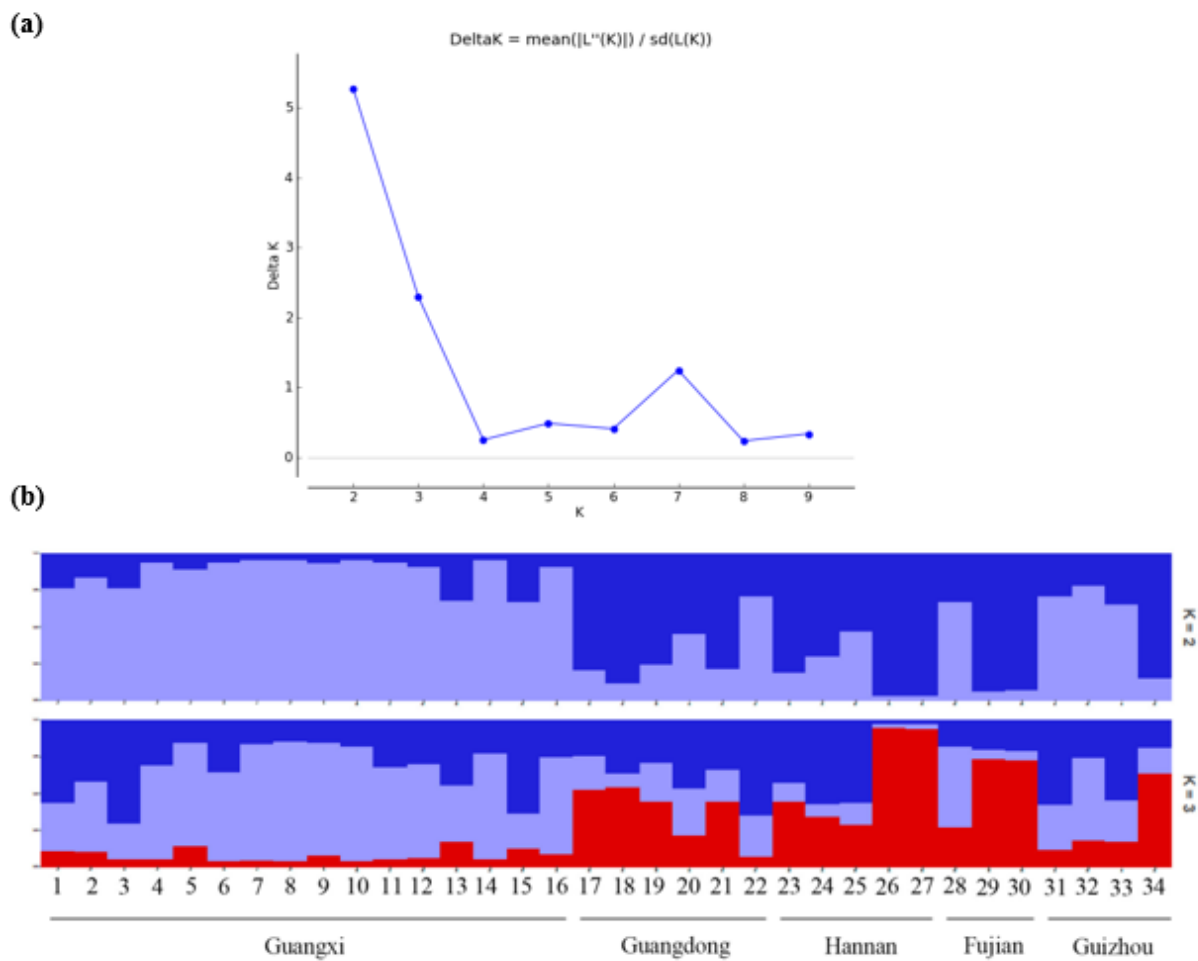


Figure 8

(a) Estimated K values for the structure analysis of *E. scandens* germplasm. **(b)** Two run results (K= 2 and K = 3) by STRUCTURE. 1–34, 34 germplasms of *E. scandens*.

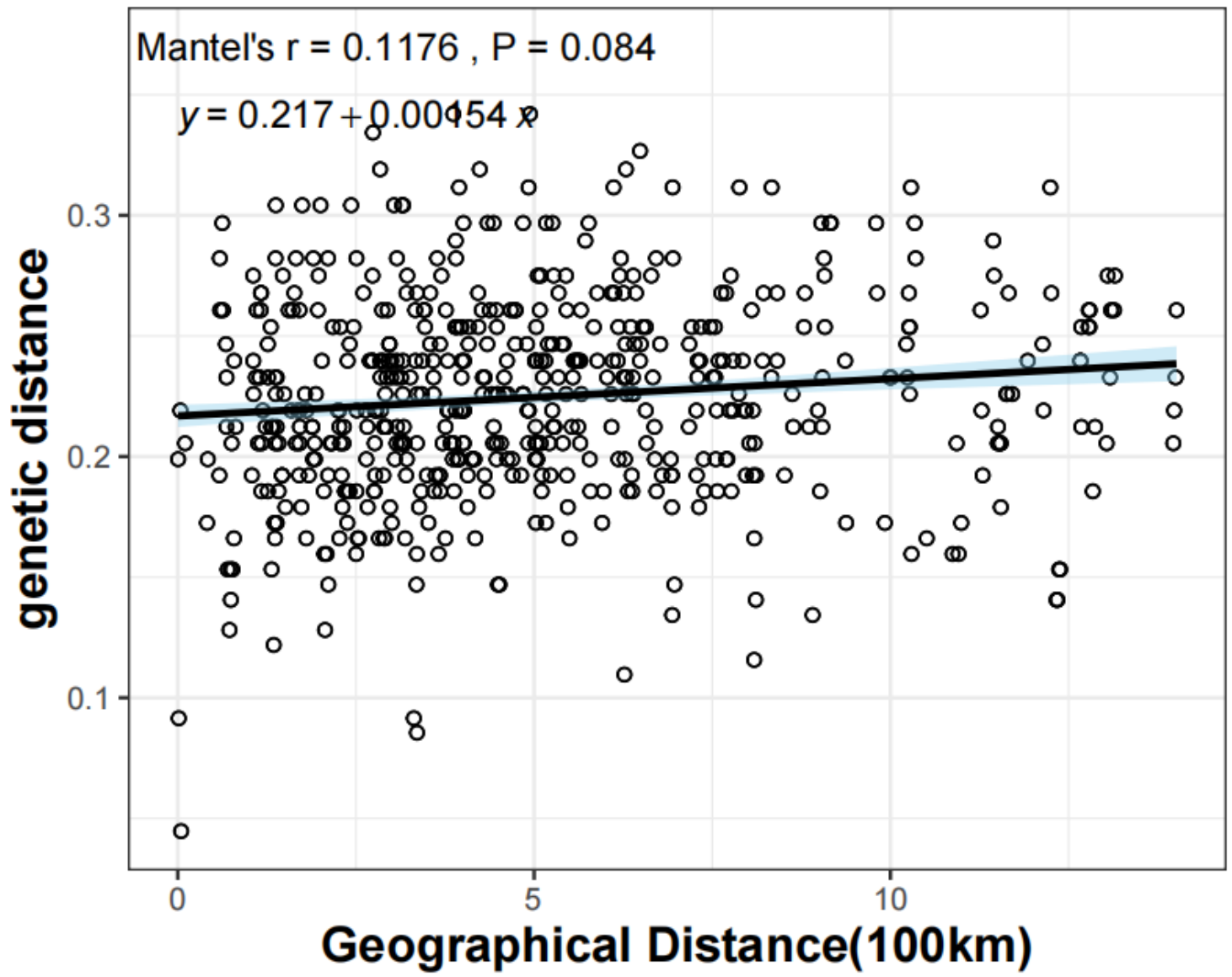


Figure 9

Mantel test to determine the correlation between geographical and genetic distances.