Biol Invasions manuscript No.
(will be inserted by the editor)

Genetic diversity of the invasive alien plant White ginger (*Hedychium coronarium*) in three National Natural Parks of the Colombian Andean region

Camilo Cano-Salazar · Esteban Gañán-Gómez · María G. Ramírez González · Javier Correa Álvarez · Diego F. Villanueva-Mejía

18 May 2020

Abstract *Hedychium coronarium* (Zingiberaceae) is a perennial rhizomatous plant considered as a potential invasive alien species in Colombia. In the present research, we examined the genetic diversity among *H. coronarium* individuals from three Andean National Natural Parks (Las Orquídeas (LO), Tatamá (TT) and Otún Quimbaya (OQ)). Two PCR-based molecular marker techniques, inter simple sequence repeats (ISSR) and random amplified polymorphic DNA (RAPD) were used to assess the genetic diversity in *H. coronarium* individuals. For thirty individuals, ISSR analysis yielded 82 fragments of which 71 were polymorphic for a percentage of 86.43%. Four RAPD mixes produced 31 fragments of which 29 were polymorphic for a percentage of 93.65%. The results of Nei's gene diversity (h) shows relatively low genetic diversity in individuals of LO in both marker systems (0.16 ± 0.21 for ISSR and 0.12 ± 0.21 for RAPD), whereas the highest genetic diversity for ISSR and RAPD markers was observed in TT and OQ respectively (0.19 ± 0.21 for ISSR and 0.23 ± 0.23 for RAPD). Dendrograms based on unweighted pair group method with arithmetic mean (UPGMA) for both marker systems grouped the individuals by National Natural Park. This clustering was further supported by principle coordinates analysis (PCoA). For RAPD, analysis of molecular variance (AMOVA) displayed high genetic variability occurred among populations rather than within populations, whereas for ISSR the variability occurred at a similar level. The results of this investigation show significant genetic variability and potential of genetic structure in the sampled National Natural Parks and will contribute to a better understanding of the invasion process of *H. coronarium* in Colombia.

Keywords Colombian Andean region · Genetic diversity · *Hedychium coronarium* · Invasive alien plant · ISSR · RAPD

C. Cano-Salazar · E. Gañán-Gómez · J. Correa · D. F. Villanueva-Mejía (✉)
Department of Biological Sciences, Universidad EAFIT, Medellín, Antioquia, Colombia
E-mail: dvillanu@eafit.edu.co

M. G. Ramírez (✉)
Santuario de Fauna y Flora Otún Quimbaya, Pereira, Risaralda, Colombia
E-mail: maria.ramirez@parquesnacionales.gov.co

Introduction

Globalization has brought with it a substantial increase in the number of invasive alien species, which greatly affect the biodiversity of ecosystems worldwide, compromising their integrity and hindering their conservation (Bartz and Kowarik, 2019). Biological invasions by these species are considered among the largest causes of biodiversity loss worldwide (Keane and Crawley, 2002; Linders et al, 2019); affecting both the functioning of ecosystems and the services they provide (Lim et al, 2014; Bartz and Kowarik, 2019). Homogenization of forests due to invasive plants not only represents the displacement of native species, but also significant economic losses associated with plant management and control, especially in areas dedicated to the conservation of biodiversity (Gallien et al, 2010). Within these invasive species, many wild gingers (*Hedychium* spp.) have proven to be a major threat, as their asexual reproduction through rhizome networks obstructs their removal from infestation sites (Acevedo-Rodríguez et al, 2019).

Hedychium coronarium J. Koenig (Zingiberaceae) is a perennial and rhizomatous plant native to the Hi-

Table 1 Plant material and geographic sampling locations.

National Natural Park	Individuals per sampling site	Latitude	Longitude	Altitude (msnm)	Habitat description
SFF Otún Quimbaya	OQ1.5 - OQ1.17	4° 43' 46.5"	75° 35' 04.5"	1858	Subandean forest Highly disturbed by human activity Moist soils
	OQ2.3 - OQ2.1	4° 43' 23.3"	75° 34' 01.8"	1959	
	OQ3.4	4° 43' 39.8"	75° 34' 36.8"	1907	
	OQ4.12 - OQ4.27	4° 43' 46.6"	75° 34' 44.5"	1903	
	OQ5.8	4° 43' 46.3"	75° 34 '40.7"	1924	
	OQ6.1 - OQ6.24	4° 43' 43.4"	75° 34' 37.4"	1870	
PNN Tatamá	TT1.3 - TT1.4	5° 14' 44.8"	76° 4' 19"	1331	Hilly moist slopes Shady and humid places River banks
	TT2.3 - TT2.4 - TT2.9	5° 12' 57.8"	76° 4' 48.4"	1569	
	TT3.3 - TT3.5 - TT3.6	5° 13' 31.1"	76° 4' 54.4"	1399	
	TT4.2 - TT4.4	5° 14' 04.4"	76° 5' 21.9"	1333	
PNN Las Orquídeas	LO1.6 - LO1.9	6° 32' 46.02"	76° 20' 11.8"	593	Grasslands and waterlogged areas High light habitats
	LO2.2 - LO2.3	6° 32' 16.94"	76° 19' 14.88"	801	
	LO3.3 - LO3.5	6° 30' 50.57"	76° 17' 19.62"	1062	
	LO4.7 - LO4.1	6° 31' 07.69"	76° 15' 04.37"	1364	
	LO5.1 - LO5.3	6° 31' 06.75"	76° 15' 14.4"	1441	

malayas, India, and southern China (Acevedo-Rodríguez et al, 2019). In Colombia, it is highly distributed and is regarded as an alien plant with high potential of being invasive (Cárdenas et al, 2010; Aguilar-Garavito et al, 2017). *H. coronarium* reproduces sexually but mainly vegetatively through rhizomes forming dense and extensive thickets that suppress the establishment of native plants (Acevedo-Rodríguez et al, 2019). This has been exacerbated by ignorance of the danger they represent on the part of the inhabitants of affected areas, partly because of the medicinal and ornamental applications of the plant (Barreto et al, 2012). Hence, it is becoming a managing problem for some National Natural Parks and protected areas for conservation due to its environmental impacts on biodiversity and the difficulty of control and prevention of its dispersal.

Genetic diversity and evolutionary changes are key points to understand biological invasions, expansion mechanisms and source populations (Sakai et al, 2001). High genetic diversity is not always required for invasion. Indeed, invasive populations tend to exhibit low genetic variation due to founder effects (Ward et al, 2008; Yang et al, 2012). However, these populations can overcome low genetic variability, either by increasing genetic diversity through multiple introductions and hybridizations or exploiting strategies such as phenotypic plasticity and asexual reproduction (Frankham, 2005; Le Roux and Wieczorek, 2009). In this regard, genetic diversity may contribute significantly to invasiveness enhancing the ability of populations to evolve and adapt to new habitats.

The use of the dominant markers ISSR and RAPD has proven to be a useful method to evaluate the genetic diversity/relationship among and within *Hedychium* species. Das et al (2011) highlights the discrimination capacity of ISSR markers as well as considering it as a fast and sensitive method for detecting genetic variations in Zingiberaceae in PCR-based fingerprinting works. Furthermore, Basak et al (2014) recommended the combined use of ISSR and RAPD markers as the most comprehensive framework for future studies in genetic comparative analyses.

This study aimed to explore the genetic diversity and variation of *H. coronarium* in three Colombian National Natural Parks located in the Andean region, in order to understand its geographic patterns of invasion and get insights about its potential for colonization and establishment. The present work marks the first attempt to determine intraspecific genetic diversity in White ginger populations distributed out of their native range.

Materials and methods

Plant material collection

Ten individuals were sampled from each National Natural Park (Las Orquídeas (LO), Tatamá (TT) and Otún Quimbaya (OQ)) for a total of thirty (Table 1). For each haphazardly selected plant, fresh and young leaves were collected and placed in plastic bags with silica gel. The distance between each collected plant was at least 5 m to minimize the chance of sampling ramets of the same clone. The samples were carried to laboratory and stored at -80°C until the DNA extraction could be performed.

Table 2 Characteristics of ISSR and mixes of RAPD primers.

Marker	Primer name	Primer sequence[a] (5'–3')	Ta (°C)	Total no of bands	No of polymorphic bands	Percentage Polymorphism	PIC	MI	Rp
ISSR									
	807	(AG)8T	51	11	9	81.82	0.31	2.82	3.93
	816	(CA)8T	48	8	7	87.50	0.20	1.39	1.93
	825	(AC)8T	54	11	9	81.82	0.41	3.68	5.53
	826	(AC)8C	56	13	13	100	0.40	5.18	8.67
	HB12	(CAC)3GC	40	8	7	87.50	0.30	2.07	2.87
	UBC-841	(AG)7AYC	48	10	10	100	0.37	3.74	5.93
	UBC-888	BDB(CA)7	50	12	9	75	0.35	3.12	4.73
	UBC-890	HVH(GT)7	52	9	7	77.78	0.12	0.83	0.93
	Total			82	71				
	Mean			10.25	8.875	86.43	0.31	2.85	4.32
RAPD									
	SIGMA-D-01	AACGCCGCC							
A	OPD-07	TTGGCACGGG	42	8	8	100	0.44	3.55	5.73
	OPA-13	CAGCACCCAC							
	OPE-07	AGATGCAGCC							
B	OPF-14	TGCTGCAGGT	44	7	6	85.71	0.34	2.03	3.13
	OPA-02	TGCCGAGCTG							
	SIGMA-D-14	CTCGCTCCA							
C	OPA-03	AGTCAGCCAC	44	7	7	100	0.32	2.22	3.27
	OPH-13	GACGCCACAC							
	OPD-08	GTGTGCCCCA							
D	OPD-02	GGACCCAACC	45	9	8	88.89	0.39	3.08	4.40
	SIGMA-D-P	GGACCGGTG							
	Total			31	29				
	Mean			7.75	7.25	93.65	0.37	2.72	4.13

[a]Y = C + T, B = G + T + C, D = G + A + T, H = A + T + C, V = G + A + C.
Ta, Annealing temperature; PIC, Polymorphism information content; MI, Marker index; Rp, Resolving power.

DNA extraction

Total genomic DNA was extracted from thirty tender leaf samples using the cetyltrimethylammonium bromide (CTAB) protocol described by Doyle and Doyle (1987) with few modifications. Quantity and quality of purified DNA were determined spectrophotometrically by measuring the ratios A260/A280 and A260/A230 using a Nanodrop 2000C (Thermo Scientific, USA) and electrophoretically using 0.8% Agarose gel. All the samples were diluted at 50 ng/µl as a working concentration for RAPD and ISSR amplifications.

ISSR amplification

After screening 11 ISSR primers, PCR amplification of genomic DNA was carried out using 8 primers, which were chosen based on the clear and reproducible banding pattern they produced within the expected range and without presence of non-specific bands (Table 2). The reaction mixture was performed in a total volume of 10 µl containing 25 ng DNA, 1x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.625 U of Taq DNA polymerase (Thermo Scientific, USA) and 4 pmol of each primer (Sigma Aldrich). PCR amplifications were carried out in a T1000 Touch thermocycler (Bio-Rad, USA). Thermal cycling conditions were as follows: initial denaturation at 94 °C for 4 min, 35 cycles consisting of denaturation at 94 °C for 45 s, annealing at specific temperature for 1 min, and extension at 72 °C for 2 min followed by final extension step of 10 min at 72 °C (Table 2). The amplification products were kept at 4 °C throughout. Amplicons were loaded with SYBR dye (NEB, USA) and separated in 2% agarose gel at 60 V in 1X TBE for 180 min. Gel pictures were photographed under a UV transilluminator (Enduro GDS, USA) and images were processed in ImageJ.

RAPD amplification

PCR amplification of the genomic DNA was carried out using 4 mixes of RAPD primers (Table 2). The reaction mixture was performed in a total volume of 10 µl containing 25 ng DNA, 1x PCR buffer, 3 mM $MgCl_2$, 0.3 mM dNTPs, 0.6 U of Taq DNA polymerase (Thermo Scientific, USA) and 3 pmol of each primer (Sigma Aldrich). PCR amplifications were carried out in a T1000 Touch thermocycler (Bio-Rad, USA). Thermal cycling condi-

tions were as follows: initial denaturation at 94 °C for 4 min, 35 cycles each consisting of denaturation at 94 °C for 1 min, annealing at specific temperature for 1 min, and extension at 72 °C for 2 min followed by final extension step of 5 min at 72 °C (Table 2). The amplification products were stored at 4 °C. DNA amplified products were separated under the same conditions specified for ISSR amplification.

Data analysis

For the two types of marker systems, the amplified DNA products were scored as 1/0 for the presence or absence of bands in each sample (Ng and Tan, 2015). To compare the efficiency of each primer, number of polymorphic bands, percentage polymorphism, polymorphism information content (PIC), marker index (MI) and resolving power (Rp) were estimated. Polymorphism information content (PIC) was calculated as PIC = 1 - Σp_i^2, where p_i is the frequency of the i_{th} allele (Lemos et al, 2019). Resolving power was computed using the formula Rp = ΣIB (IB, band information = 1 - [2*(0.5-P)], where P is the proportion of individuals containing the band (Prevost and Wilkinson, 1999). POPGENE32 software was used to calculate the basic parameters for genetic diversity such as the observed number of alleles (n_a), effective number of alleles (n_e), Nei's gene diversity (h), Shannon's information index (I), number of polymorphic loci (NPL), and percentage of polymorphic loci (PPL) (Yeh et al, 1997).

Cluster analysis

Level of similarity among individuals was determined and compiled in a matrix of genetic similarity using the DICE's coefficient (Nei and Li, 1979; Kosman and Leonard, 2005). A dendrogram representing the genetic relationships for each marker system was performed by applying UPGMA (unweighted pair group method average) method using DARwin software V6.021 (Perrier and Jacquemoud-Collet, 2006). To check the goodness of fit for thirty individuals of *H. coronarium* to a specific cluster in the UPGMA algorithm, the correlation between the original similarity indices and cophenetic values was calculated preforming a Mantel test using 10000 permutations in DARwin (Mantel, 1967). Principal coordinates analysis (PCoA) and analysis of molecular variance (AMOVA) based on 9999 permutations were performed in the Excel extension GenAlEx V6.5 (Peakall and Smouse, 2012). To obtain the correlation coefficient between geographic and genetic distances (the so called "isolation-by-Distance" hypothesis) among all individuals, a Mantel test was performed also in GenAIEx using 999 permutations (Mantel, 1967).

Results

Efficiency of polymorphism detection

The intraspecific genetic diversity of *H. coronarium* was evaluated for three National Natural Parks located in the Andean region, where 10 leaf samples were taken from each for a total of 30. The 8 selected ISSR markers previously described were tested to verify production of specific and informative DNA profiles (Das et al, 2011). The 8 primers generated a total of 82 loci, out of which 71 were polymorphic (86.43 %)(Table 2). With an average of 8.875 fragments per primer, the polymorphic products ranged from 7 (ISSR 816, ISSR HB12, and ISSR UBC-890) to a maximum of 13 (ISSR 826). After evaluating the informativeness in the other marker system, 31 fragments were amplified by the 4 RAPD mixes, of which 29 were polymorphic for a total of 93.65%. The highest number of polymorphic fragments was 8 (RAPD A and RAPD D) and the lowest was 6 (RAPD B). The polymorphism information content (PIC) for the ISSR dataset ranged from 0.12 for ISSR UBC-890 to 0.41 for ISSR 825 with an average of 0.31, whereas for RAPD dataset varied from 0.32 for RAPD C to 0.44 for RAPD A with an average of 0.37. For marker index (MI), the

Table 3 Genetic diversity parameters for ISSR primers applied to three populations of *Hedychium coronarium*.

Population	Sample size	n_a	n_e	h	I	NPL	PPL	G_{ST}	Nm
SFF Otún Quimbaya	10	1.55 ± 0.50	1.29 ± 0.36	0.17 ± 0.20	0.26 ± 0.28	45	54.88		
PNN Tatamá	10	1.51 ± 0.50	1.33 ± 0.38	0.19 ± 0.21	0.28 ± 0.30	42	51.22		
PNN Las Orquídeas	10	1.44 ± 0.50	1.30 ± 0.42	0.16 ± 0.21	0.24 ± 0.30	36	43.90		
Mean	10	1.50 ± 0.50	1.31 ± 0.39	0.17 ± 0.21	0.26 ± 0.29	41	50		
Total gene diversity	30	1.87 ± 0.34	1.51 ± 0.38	0.29 ± 0.19	0.44 ± 0.25			0.405	0.735

n_a, Observed number of alleles; n_e, Effective number of alleles; h, Nei's gene diversity; I, Shannon's Information index; NPL, Number of polymorphic loci; PPL, Percentage of polymorphic loci.
G_{ST} = coefficient of gene differentiation.
Nm = estimate of gene flow.

Table 4 Genetic diversity parameters for RAPD primers applied to three populations of *Hedychium coronarium*.

Population	Sample size	n_a	n_e	h	I	NPL	PPL	G_{ST}	Nm
SFF Otún Quimbaya	10	1.55 ± 0.51	1.42 ± 0.43	0.23 ± 0.23	0.34 ± 0.32	17	54.84		
PNN Tatamá	10	1.58 ± 0.50	1.32 ± 0.38	0.19 ± 0.20	0.29 ± 0.29	18	58.06		
PNN Las Orquídeas	10	1.26 ± 0.45	1.23 ± 0.40	0.12 ± 0.21	0.17 ± 0.29	8	25.81		
Mean	10	1.46 ± 0.49	1.32 ± 0.40	0.18 ± 0.21	0.27 ± 0.30	14.33	46.24		
Total gene diversity	30	1.94 ± 0.25	1.70 ± 0.29	0.39 ± 0.14	0.56 ± 0.18			0.534	0.436

n_a, Observed number of alleles; n_e, Effective number of alleles; h, Nei's gene diversity; I, Shannon's Information index; NPL, Number of polymorphic loci; PPL, Percentage of polymorphic loci.
G_{ST} = coefficient of gene differentiation.
Nm = estimate of gene flow.

average ISSR value (2.85) was almost the same as the RAPD value (2.72). Average Rp value was also similar in both systems, where ISSR Rp (4.32) was slightly higher than RAPD Rp (4.13).

Gene diversity

For the ISSR dataset studied among the three National Natural Parks, the percentage of polymorphic loci (*PPL*) ranged from 43.9% in LO to 54.88% in OQ, with an average value of 50% (Table 3). The individuals of OQ displayed the highest (1.55 ± 0.5) observed number of alleles (n_a) and the value for LO was the lowest (1.44 ± 0.5); with an average of 1.5 ± 0.5. With lower variability than n_a, the effective number of alleles (n_e) ranged from 1.29 ± 0.36 (OQ) to 1.33 ± 0.38 (TT) with an average of 1.31 ± 0.39. Nei's gene diversity (h) and Shannon's information index (I) were the highest (h=0.19 ± 0.21 and I=0.28 ± 0.3) in TT and the lowest values (h=0.16 ± 0.21 and I=0.24 ± 0.3) were in LO. In the RAPD system, an average of 46.24% of the loci were polymorphic (*PPL*) for the three National Natural Parks, where the minimum (25.81%) and the maximum (58.06%) values were detected in LO and TT respectively (Table 4). The observed number of alleles (n_a) displayed the highest value (1.58 ± 0.5) in TT and the lowest (1.26 ± 0.45) in LO with an average of 1.46 ± 0.49. The effective number of alleles (n_e) varied from 1.23 ± 0.4 (LO) to 1.42 ± 0.43 (OQ) with an average of 1.32 ± 0.40. Nei's gene diversity (h) and Shannon's information index (I) were the highest (h=0.23 ± 0.23 and I=0.34 ± 0.32) in OQ and the lowest (h=0.12 ± 0.21 and I=0.17 ± 0.29) in LO.

Genetic similarity analysis

The DICE genetic similarity values and hierarchical clustering based on UPGMA were used to determine the relationship among individuals (Nei and Li, 1979). Resulting clusters for each marker system were represented as UPGMA dendrograms reconstructed by DARwin V6.021. The ISSR dataset grouped genotypes in three major clusters that correspond to the three National Natural Parks sampled (Fig.1A). The same clustering pattern was indicated by RAPD analysis, but it displays a different tree topology (Fig. 1B).

Principal coordinates analysis for ISSR and RAPD also exhibited a pattern where clusters grouped genotypes of the same National Natural Park (Fig. 2). The first three axes of the ISSR analysis explained the 55.56% of the total variation (Axis 1 - 30.56%, Axis 2 - 17.54%, Axis 3 - 7.46%) (Table 5). Based on the RAPD data, the 68.57% of the total variation was explained by the first three axes (Axis 1 - 36.9%, axis 2 - 23.01% and axis 3 - 8.66%) (Table 5).

Analysis of molecular variance (AMOVA) for ISSR indicates that 51% of the genetic variance is due to differentiation within populations and 49% among populations (Fig. 3). The estimator of genetic differentiation among populations Φ_{ST} for this set of markers was 0.494 (Table 6). The AMOVA based on RAPD analysis showed that 63% of the genetic variance can be accounted for the differences among populations and the remaining 37% is owing to the differences within populations (Fig 3). For RAPD markers, the Φ_{ST} was also calculated being a value of 0.63.

Discussion

Recognizing the genetic diversity status of the plant populations may provide valuable information to understand the invasion process. Genetic diversity and variation of *H. coronarium* in the Colombian Andean region was measured using two marker systems.

There are several criteria and recommendations when scoring PCR/based markers on a gel in order to minimize human and stochastic errors. Scoring only considers those bands that are clear, unambiguous, and reproducible (Ng and Tan, 2015). Here, the analyses were carried out based on scoring without replicates,

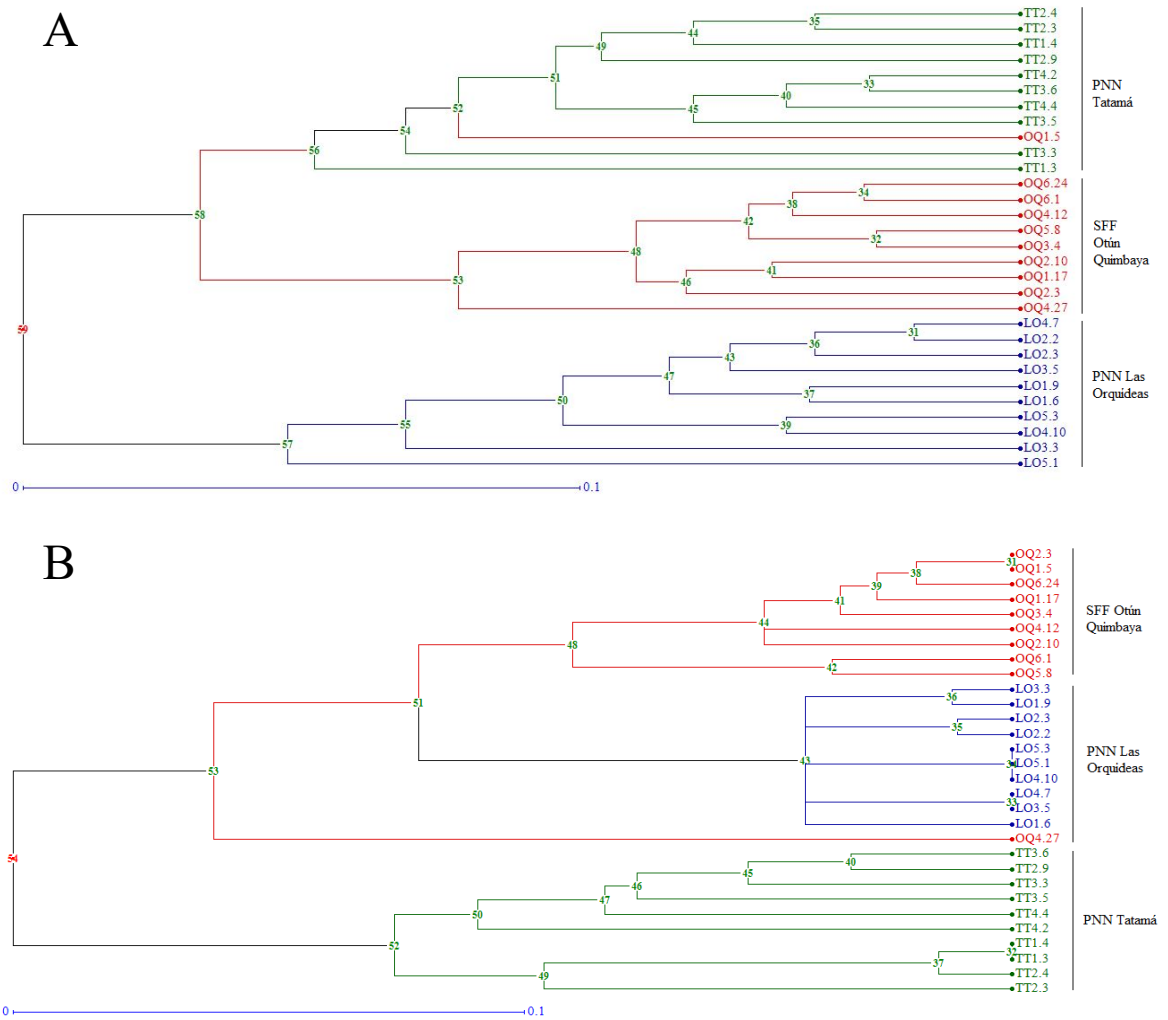
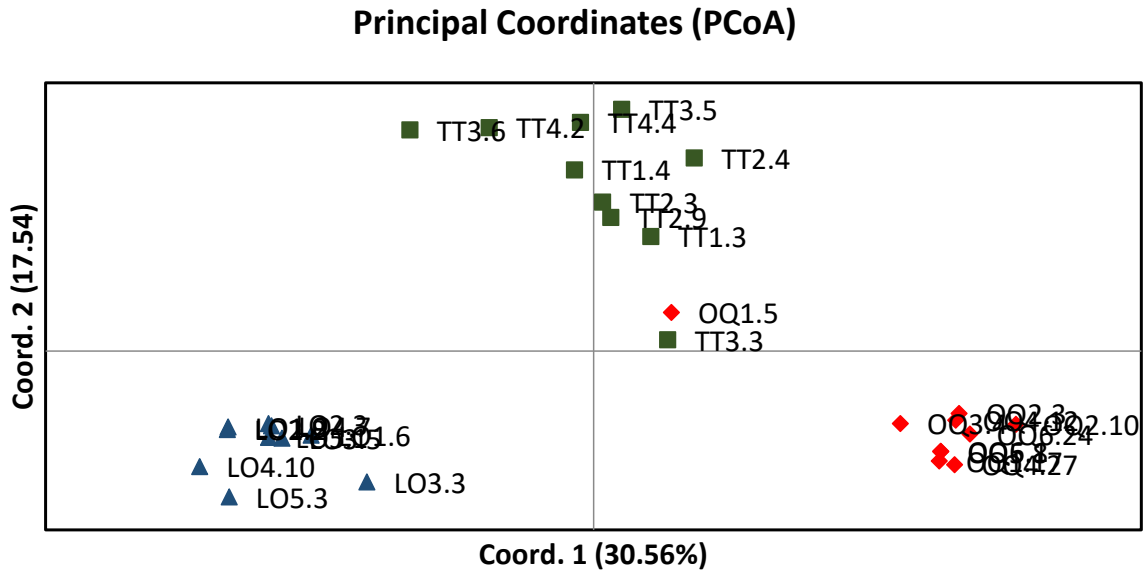


Fig. 1 Dendrograms of the UPGMA cluster analysis among the 30 individuals of *H. coronarium* using (A) ISSR ($r = 0.893$) and (B) RAPD data ($r = 0.93$). Numbers indicate DICE genetic distances.

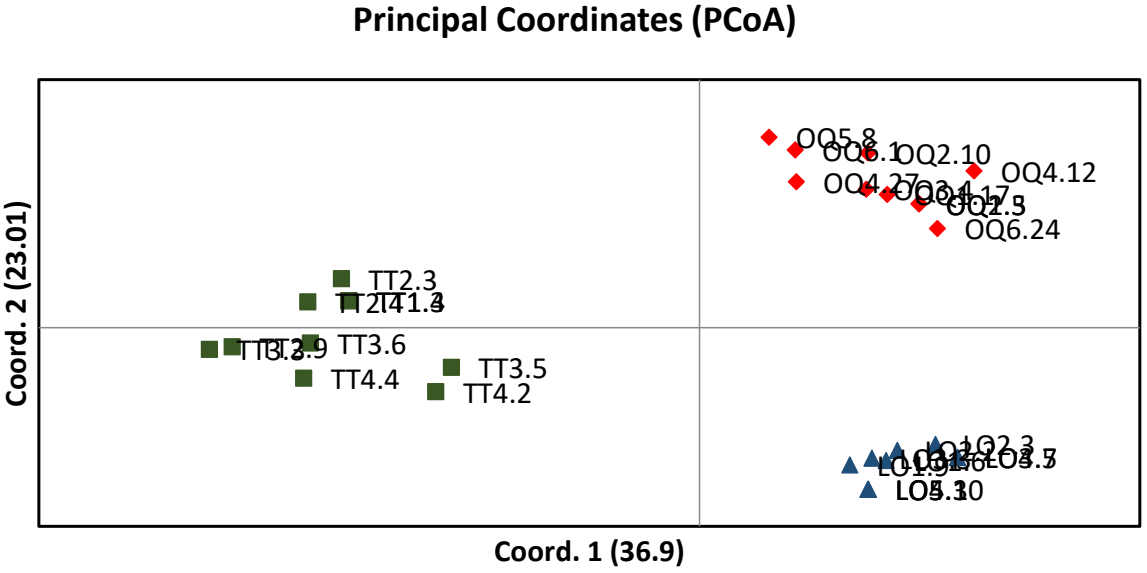
which could not be finished on time due to the COVID-19 contingency. Thus, reproducibility and unambiguity could not be guaranteed for the bands, and the presented results are to be interpreted with caution.

Regarding detection, both marker systems display a high polymorphism percentage (86.43% for ISSR and 93.65% for RAPD). The probability of finding a marker in two different states (present/absent) in two randomly selected individuals (PIC) is also high in both systems, except for ISSR 816 and ISSR UBC-890. Data from these two markers were therefore dismissed for all the analyses of genetic similarity and variation. Considering the high capacity of the markers to differentiate between genotypes (resolving power), these results clearly support the previously reported information (Basak et al, 2014; Ray et al, 2019). Previous studies on the assessment of genetic diversity using similar RAPD and ISSR markers have reported congruent values of polymorphism information content (PIC) and resolving power (Rp) to those here shown (Das et al, 2011; Singh et al, 2012; Basak et al, 2014; Rajkumari and Sanatombi, 2019; Ray et al, 2019; Saha et al, 2019). Regarding that these studies have been carried out in population of *Hedychium* species and plants of the family Zingiberaceae, our results are highly supported by the efficiency of detecting polymorphism in *H. coronarium*.

Because dominant polymorphisms usually exhibit dominance, genotypes possessing two homologous copies of a fragment cannot be distinguished from genotypes possessing only one (heterozygotes). ISSR and RAPD markers normally underestimate the number of genotypes at any given locus (Kliber and Eckert, 2005). However, this bias should be greatly reduced when multiple loci are analyzed (Patamsytė et al, 2013). Although as many primers as possible were used, the number of loci generated by each system is relatively low (82 for



(A)



(B)

Fig. 2 Principal coordinate analyses based on (A) ISSR and (B) RAPD data of *H. coronarium* from three populations.

ISSR and 31 for RAPD), therefore the markers yield information at the phenotype rather than the genotype level (Patamsytė et al, 2013).

For each diversity parameter, the National Natural Park Las Orquídeas (LO) has the lowest values in both marker systems, suggesting that LO has the lowest intraspecific genetic variability among its individuals relative to the other parks. G_{ST}, being an analogous statistical to F_{ST}, describes the partitioning of genetic diversity within and among populations (Holsinger and Weir, 2009). Multi-population G_{ST} displays intermediate values (0.405 for ISSR and 0.534 for RAPD), thereby suggesting that evaluated populations are not homogeneous. This result provides insights into the existence of

a relative genetic structure of the populations. Comparing the two marker systems, the G_{ST} for ISSR dataset is slightly lower than RAPD value, which can be explained by its discrimination power among individuals, detecting greater variability within populations.

Table 5 Percentage of variation explained by first three axes in PCoA.

Axis	ISSR			RAPD		
	1	2	3	1	2	3
% variance	30.56	17.54	7.46	36.90	23.01	8.66
% cumulative	30.56	48.10	55.56	36.90	59.91	68.57

To date, few studies have examined the genetic diversity of alien species in Colombia and there have been no attempts to carry out the genetic diversity and variation of invasive alien plants in any region of this country. Conversely, several studies have been made not only to assess the genetic diversity of *H. coronarium* in its native range of distribution, but also of many plants of the same family such as Curcuma (Das et al, 2011; Singh et al, 2012). Our results for genetic diversity reflect similar and comparable findings as reported in those studies also using PCR-based molecular markers. For instance, Basak et al (2014), Ray et al (2019) and Saha et al (2019) found high genetic variation in populations of *H. coronarium* in the northeast of India displaying similar values for Nei's gene diversity (h); Shannon's information index (I) and number of polymorphic loci (NPL). Moreover, Ray et al (2019) also found a clustering pattern by geographic location as it is shown in the present results of clustering analysis.

Although the UPGMA analysis provides a different tree topology for each marker system, both dendrograms display a clustering pattern that grouped the individuals according to the three sampled National Natural Parks (Fig. 1). These differences between analyses might be the result of limitations in the resolution that each marker yields in terms of individuals or populations. According to Ray et al (2019), ISSR provide high resolution at the level of individuals for *H. coronarium*. In contrast, RAPD markers used in this study were not suitable to find differences within populations. This can explain the unresolved relations for the LO population (Fig. 1). Nevertheless, these results suggest that there is enough genetic distance among the studied populations for appropriate clustering. Dendrograms are also highly supported by the cophenetic correlation coefficients, which are r=0.89 for ISSR and r=0.93 for RAPD.

The PCoA analysis showed a similar cluster patterns found by UPGMA dendrograms (Fig. 2). These results further support the idea of populations are genetically different because they are grouped by specific geographic locations. However, the results for the ISSR analysis should be interpreted with caution because it showed 55.56% of cumulative eigenvalues explaining the total variability of the individuals (Table 5). RAPD exhibit higher cumulative eigenvalues (68,57%) explaining the

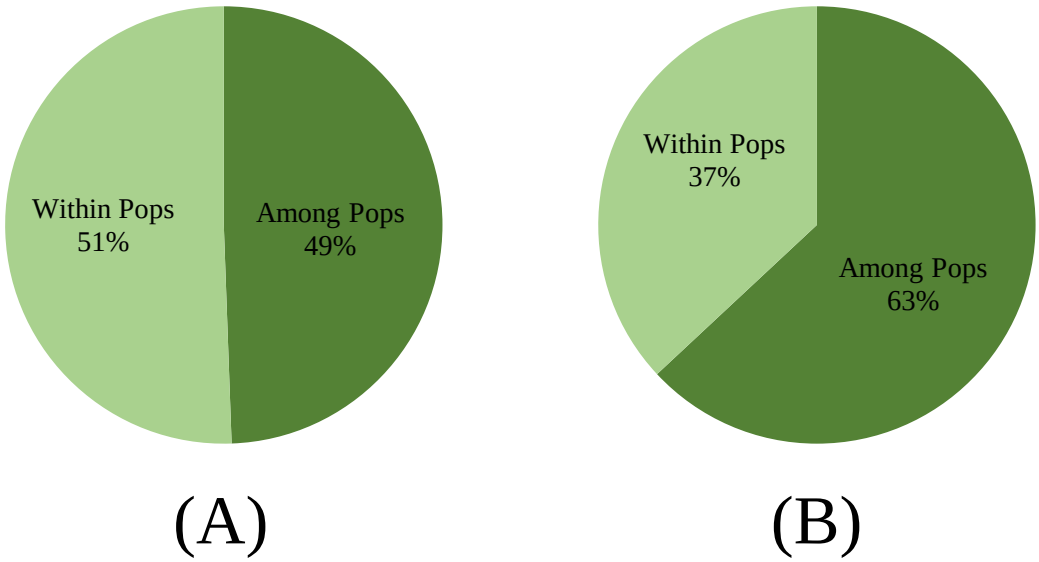


Fig. 3 Analysis of molecular variance (AMOVA) of three populations of *H. coronarium* based on (A) ISSR and (B) RAPD marker data.

Table 6 Analysis of molecular variance (AMOVA) based on ISSR and RAPD markers in *H. coronarium*.

Source	df	SS	MS	Est. Var.	Variance (%)	P-value	Φ_{ST}
ISSR							
Among Pops	2	137.20	68.60	6.22	49	0.0001	0.494
Within Pops	27	171.90	6.37	6.37	51		
Total	29	309.10		12.59	100		
RAPD							
Among Pops	2	93.40	46.70	4.41	63	0.0001	0.630
Within Pops	27	69.80	2.59	2.59	37		
Total	29	163.20		7.00	100		

Significant level based on 9999 permutations; P-value = probability of obtaining a more extreme component by chance alone.

total variability of the individuals (Table 5). These results seem to be consistent with previous analyses using ISSR, in which the genetic variability within populations, rather than among populations, is more easily detected (Singh et al, 2012; Ray et al, 2019).

The analysis of molecular variance (AMOVA) indicates the distribution of genetic variance among the sampled individuals. For our ISSR markers, 51% of the variance can be attributed to differences within populations, and 49% to differences among populations (Fig. 3). RAPD results exhibit a slightly different distribution (37% within populations and 63% among populations) (Fig. 3). These results may further indicate that there is a relative high differentiation among populations. Additionally, the different distribution of the genetic variance for ISSR and RAPD may be explained in part by the low sensibility of RAPD to reveal genetic diversity between individuals. AMOVA also provides the Φ_{ST} as an alternative estimator of genetic differentiation among populations and is analogous to the classical F_{ST} (Wright, 1931) (Table 6). Based on these values for both genetic marker systems (0.494 for ISSR and 0.63 for RAPDS) we found the sampled populations to be genetically distinct. These analyses support the hypothesis that *H. coronarium* in Colombia displays population structure.

Mantel tests between the genetic distances provided by ISSR markers and the geographic distances reveal a correlation coefficient of 0.60. Although this value is not high enough to prove the isolation by distance hypothesis, more populations in a broader geographic space should be evaluated to improve the resolution of the analyses (Diniz-Filho et al, 2013).

Conclusion

This study is the first piece of research that has examined the genetic diversity of the invasive plant *H. coronarium* outside its native range, specifically of those populations inside three National Natural Parks located in the Colombian Andean region. The results of this investigation show significant genetic variability of the sampled populations, which furthermore are genetically structured according to their geographic locations. However, these findings are limited by the lack of replicate of the electrophoresis gels for scoring bands and the low number of loci analyzed in each marker system used.

Further studies need to be carried out in order to confirm these results. It is recommended that these studies take into account more geographic locations spanning a broader invasion range. Both PCR-based markers used in this study are suitable to survey the genetic diversity of *H. coronarium*. However, it would be interesting to use haplotypic markers such as cpDNA regions to complement the analyses of genotypic geographical distribution patterns.

The findings of this research will contribute to a better understanding of the invasion process of *H. coronarium* in Colombia, and are directed to improve current management practices, and help devise better control strategies. Beyond the collection of more and better genetic and evolutionary data, ecological studies should be conducted to get the updated information on the colonization mechanisms and the development of the invasion in the country and elsewhere.

Acknowledgements The authors are grateful to Parques Nacionales Naturales de Colombia for providing logistic support with field work; to Universidad EAFIT for financial support and extending all infrastructural facilities; and to Daniel Montoya-Londoño for his support in the initial phases of this investigation and for revisioning this manuscript.

Conflict of interest

The authors declare that there is no conflict of interest or involvement in any organization or entity with any financial interest, to submit this manuscript.

References

Acevedo-Rodríguez P, Djeddour D, Rojas-Sandoval J (2019) *Hedychium coronarium* (white butterfly ginger lily). CABI. https://www.cabi.org/isc/datasheet/26678. Accessed 24 April 2020

Aguilar-Garavito M (2017) *Hedychium coronarium*. In: Baptiste MP, García LM, Castellanos-Castro C, González R et al (eds) Plantas exóticas con alto potencial de invasión en Colombia. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, pp 229–234

Barreto RW, Ellison CA, Seier MK, Evans HC (2012) Biological Control of Weeds with Plant Pathogens: Four Decades On. In: Abrol DP, Shankar U (eds) Integrated Pest Management: Principles and Practice. CABI, Massachusetts, pp 299-334

Bartz R, Kowarik I (2019) Assessing the environmental impacts of invasive alien plants: A review of assessment approaches. NeoBiota 43:69–99

Basak S, Ramesh AM, Kesari V, Parida A, Mitra S, Rangan L (2014) Genetic diversity and relationship of *Hedychium* from Northeast India as dissected using PCA analysis and hierarchical clustering. Meta gene 2:459–468

Cárdenas D, Castaño N, Cárdenas-Toro J (2010) Análisis de riesgo de especies de plantas introducidas para Colombia. In: Baptiste MP, Castaño N, Cárdenas D, Gutiérrez FP et al (eds) Análisis de riesgo y propuesta de categorización de especies introducidas para Colombia. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, pp 56–57

Das A, Kesari V, Satyanarayana VM, Parida A, Rangan L (2011) Genetic relationship of *Curcuma* species from Northeast India using PCR-based markers. Molecular biotechnology 49:65–76

Diniz-Filho JAF, Soares TN, Lima JS, Dobrovolski R et al (2013). Mantel test in population genetics. Genetics and molecular biology 36:475–485

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bull 19:11–15

Frankham R (2005) Resolving the genetic paradox in invasive species. Heredity 94:385-385

Gallien L, Münkemüller T, Albert CH, Boulangeat I et al (2010) Predicting potential distributions of invasive species: where to go from here? Diversity and Distributions 16:331–342

Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}. Nature Reviews Genetics 10:639–650

Keane RM and Crawley M J (2002). Exotic plant invasions and the enemy release hypothesis. Trends in ecology and evolution 17:164-170

Kliber A, Eckert CG (2005) Interaction between founder effect and selection during biological invasion in an aquatic plant. Evolution 59:1900–1913

Kosman E, Leonard KJ (2005) Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid, and polyploid species. Molecular ecology 14:415–424

Le Roux J, Wieczorek AM (2009) Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. Annals of Applied Biology 154:1–17

Lemos SCM, Silveira RLR, Buuron SK et al (2019) Determining the Polymorphism Information Content of a Molecular Marker. Gene 144175

Lim J, Crawley MJ, De Vere N et al (2014) A phylogenetic analysis of the British flora sheds light on the evolutionary and ecological factors driving plant invasions. Ecology and evolution 4:4258–4269

Linders TEW, Schaffner U, Eschen R, Abebe A et al (2019) Direct and indirect effects of invasive species: Biodiversity loss is a major mechanism by which an invasive tree affects ecosystem functioning. Journal of ecology 107:2660–2672

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences 76:5269-5273

Ng WL, Tan SG (2015) Inter-simple sequence repeat (ISSR) markers: are we doing it right? ASM Sci J 9:30–39

Patamsytė J, Rančelis V, Čėsnienė T, Kleizaitė V et al (2013) Clonal structure and reduced diversity of the invasive alien plant *Erigeron annuus* in Lithuania. Central European Journal of Biology 8:898–911

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics 28:2537–2539

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288–295

Perrier X, Flori A, Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (ed) Genetic diversity of cultivated tropical plants. Enfield, Science Publishers, Montpellier, pp 43–76

Perrier X, Jacquemoud-Collet JP (2006) DARwin software

Prevost A, Wilkinson MJ (1999) A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. Theoretical and Applied Genetics 98:107–112

Rajkumari S, Sanatombi K (2019) Genetic diversity analysis of *Hedychium* species based on RAPD and ISSR markers. Proceedings of the National Academy of Sciences, Indian Section B: Biological Sciences 89:623–629

Ray A, Jena S, Haldar T et al (2019) Population genetic structure and diversity analysis in *Hedychium coronarium* populations using morphological, phytochemical and molecular markers. Industrial Crops and Products 132:118–133

Saha K, Sinha RK, Sinha S (2019) Chromosomal and PCR-based molecular characterization of *Hedychium* spp. of Tripura, North-East India. Vegetos 32:521-531

Sakai AK, Allendorf FW, Holt JS, Lodge DM et al (2001) The population biology of invasive species. Annual Review of Ecology and Systematics 32:305–332

Singh S, Panda MK, Nayak S (2012) Evaluation of genetic diversity in turmeric (*Curcuma longa* L.) using RAPD and ISSR markers. Industrial Crops and Products 37:284–291

Ward SM, Reid SD, Harrington J, Sutton J et al (2008). Genetic variation in invasive populations of yellow toadflax (*Linaria vulgaris*) in the western United States. Weed science 56:394-399

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

Yang J, Tang L, Guan YL, Sun WB (2012) Genetic diversity of an alien invasive plant Mexican sunflower (*Tithonia diversifolia*) in China. Weed Science 60:552–557

Yeh FC, Yang RC, Boyle TBJ, Ye ZH et al (1997) POPGENE, the User-Friendly Shareware for Population Genetic Analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Alberta