Evolution

Genetic diversity, outcrossing rate, and demographic history along a climatic gradient in the ruderal plant *Ruellia nudiflora* (Acanthaceae)

Diversidad genética, tasa de entrecruzamiento e historia demográfica a lo largo de un gradiente climático en la planta ruderal Ruellia nudiflora *(Acanthaceae)*

Carlos F. Vargas-Mendoza [a], Ilka Ortegón-Campos [b], Denis Marrufo-Zapata [b],
Carlos M. Herrera [c], Víctor Parra-Tabla [b,*]

[a] *Laboratorio de Variación Biológica y Evolución, Escuela Nacional de Ciencias Biológicas-Instituto Politécnico Nacional, Prolongación de Carpio y Plan de Ayala s/n, Col. Casco de Santo Tomás, 11340 México D.F., Mexico*
[b] *Departamento de Ecología Tropical, Campus de Ciencias Biológicas y Agropecuarias, Universidad Autónoma de Yucatán, Apartado Postal 4-116, 97000 Itzimná, Mérida, Yucatán, Mexico*
[c] *Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas, Isla de la Cartuja, Avenida Américo Vespucio s/n, 41092 Sevilla, Spain*

Abstract

Ruellia nudiflora has shown a high potential to easily invade disturbed areas. Outcrossing rate and genetic structure and diversity in this species were examined along a climatic gradient in the Yucatán Peninsula (Mexico) in order to understand the effects of environmental heterogeneity – isolation by environment (IBE) – as well as correlation in herkogamy on genetic structure, diversity, and demographic history in this species. Nine populations were sampled along a temperature-precipitation gradient with marked differences in vegetation, measuring the degree of herkogamy in each sampled plant. To evaluate genetic diversity and structure, amplified fragment length polymorphism (AFLP) markers from 76 plants in these populations were used, while the outcrossing rate was evaluated in 22 genetic families from six populations. All populations had high levels of within-population genetic diversity and high outcrossing rates. Virtually all populations were also undergoing demographic and spatial expansion. The partial Mantel test found a positive correlation between geographic and genetic distance, but none between mean temperature, precipitation, and herkogamy with genetic distance. In this context, *R. nudiflora* shows incipient evidence of genetic differentiation on a small geographical scale (isolation by distance, IBD) but not along the environmental gradient (IBE). Results suggest that under conditions of constant disturbance, *R. nudiflora* is able to continuously colonize new areas.

Resumen

Ruellia nudiflora ha mostrado un alto potencial para invadir fácilmente áreas perturbadas. Con la finalidad de comprender los efectos de la heterogeneidad ambiental (aislamiento por el ambiente, IBE) y de la correlación en la hercogamia, sobre la diversidad, estructura genética e historia demográfica de esta especie, se examinaron la tasa de entrecruzamiento, la estructura y diversidad genética a lo largo de un gradiente climático en la península de Yucatán (México). Se muestreó un total de 9 poblaciones a lo largo de un gradiente de temperatura-precipitación con marcadas diferencias en la vegetación, en donde además se midió el grado de hercogamia en cada planta. Para evaluar la diversidad y estructura genética, se usaron AFLP's en 76 plantas de estas poblaciones, mientras que la tasa de entrecruzamiento se evaluó en 22 familias genéticas de

* Corresponding author.
 E-mail address: ptabla@uady.mx (V. Parra-Tabla).

Peer Review under the responsibility of Universidad Nacional Autónoma de México.

6 poblaciones. Todas las poblaciones de estudio tuvieron niveles altos de diversidad genética, así como una tasa elevada de entrecruzamiento. Prácticamente todas las poblaciones experimentaron expansión demográfica y espacial. La prueba parcial de Mantel mostró una correlación positiva de la distancia geográfica con la distancia genética, pero no hubo relación de la temperatura promedio, precipitación y hercogamia con la distancia genética. En este contexto, *R. nudiflora* parece mostrar evidencia incipiente de una diferenciación genética en una pequeña escala geográfica (aislamiento por distancia, IBD), pero no en el gradiente ambiental. Los resultados sugieren que bajo condiciones de perturbación constante, *R. nudiflora* es capaz de colonizar áreas nuevas continuamente.

Introduction

Ruderal plant species are characterized by life-history traits which make them tolerant to disturbance (Grime, 1979). This functional group has been favored by anthropogenic disturbances in terrestrial ecosystems during the past century (Pino, Font, de Cáceres, & Molowny-Horas, 2009), and these habitat changes have opened new niches and resulted in demographic and range expansion for many species (Alexander, Van Kleunen, Ghezzi, & Edwards, 2012). In this way, ruderal plants may achieve invasive status and displace species in some areas (Bossdorf et al., 2005). Thus, as new areas become disturbed and are subsequently colonized by ruderal species, these species contribute to habitat modification (Alexander et al., 2012).

A common approach for evaluating the invasive capacity of ruderal plants has been to describe their evolutionary genetics and capacity to displace other species (Leiss, & Müller-Schäer, 2001; Müller-Schäer, Schaffner, & Steinger, 2004; Poulin, Weller, & Skal, 2005). In contrast, very little is known about the outcrossing rate, demographic status, and demographic expansion of these species. However, there is evidence that bioclimatic heterogeneity influences genetic diversity and structure in ruderal plants, and has even been found to have a selective effect on some genetic markers (Hamasha, Schmidt-Lebuhn, Durka, Schleuning, & Hensen, 2013 and references therein).

The herb *Ruellia nudiflora* (Engelman and Gray) Urban (Acanthaceae) is found from southern Texas and northern Tamaulipas to Central America and the Caribbean islands, where it occurs in disturbed areas and is considered an invasive native weed; however, its dispersal history remains unknown (Villaseñor, & Espinoza, 1998). *R. nudiflora* exhibits traits common to ruderal species, such as a high dispersal ability, high rates of seed germination and seedling survival in environments with low water availability, and adequate growth in a wide range of soil types (Cervera, & Parra-Tabla, 2009; Ortegón-Campos et al., 2012), like other species of the genus *Ruellia* (Meyer, & Lavergne, 2004). Furthermore, *R. nudiflora* grows in contrasting environments including deforested habitats as well as sites with varying levels of precipitation (Cervera, & Parra-Tabla, 2009; Munguía-Rosas, Ollerton, & Parra-Tabla, 2012). Such climatic variation is associated with different vegetation types ranging from spiny low deciduous forest in the north to medium-height subdeciduous forest in the south (Flores, & Espejel-Carvajal, 1994).

This species is self-compatible and exhibits varying low levels of pollen limitation, but its outcrossing rate and other mating system traits are unknown (Abdala-Roberts, Marrufo-Zapata, Arceo-Gómez, & Parra-Tabla, 2012). Additionally, populations of *R. nudiflora* in the Yucatán Peninsula display drastic spatial variation in floral traits such as degree of herkogamy (Marrufo, 2011). This, together with the presence of cleistogamous flowers that obligately self-pollinate, contributes to the high reproductive success of this species (a very common pattern in invasive plants; Campbell, Quinn, Cheplick, & Bell, 1983) and may also be responsible for spatial variations in outcrossing rate and genetic diversity.

All the aforementioned traits contribute to make *R. nudiflora* a species with a high potential to colonize anthropologically disturbed sites, although the location and date of its introduction to the Yucatán Peninsula are so far unknown.

The present study aimed to describe genetic structure and diversity as well as the outcrossing rate along a climatic gradient, in populations of the ruderal plant *R. nudiflora*. Since increased environmental variability on a small spatial scale favors genetic differentiation influenced by differential local adaptation (Lee, & Mitchell-Olds, 2011; Leiss, & Müller-Schäer, 2001), it seemed desirable to test the hypothesis that genetic variability and outcrossing rate are structured along a climatic gradient in which genetic differences between nearby populations are greater if these populations belong to different environments, and smaller if they are in similar environments. In particular, the following questions were addressed: (i) are populations of *R. nudiflora* in the Yucatán Peninsula genetically structured along a climatic gradient? and if so, what is the relative contribution of isolation by distance and isolation by environment on the genetic pattern? (ii) Is there evidence of historical demographic differences among these populations that suggest genetic bottlenecks and/or population expansion events? (iii) Is there a relationship between phenotypic variation (measured in levels of herkogamy) and genetic variation?

Materials and methods

Study species and characterization of study sites

Ruellia nudiflora is a 20–30-cm tall perennial plant that shows few ramifications (3–6) and produces chasmogamous (CH) and cleistogamous flowers (CL) (Long, 1971). Fruit
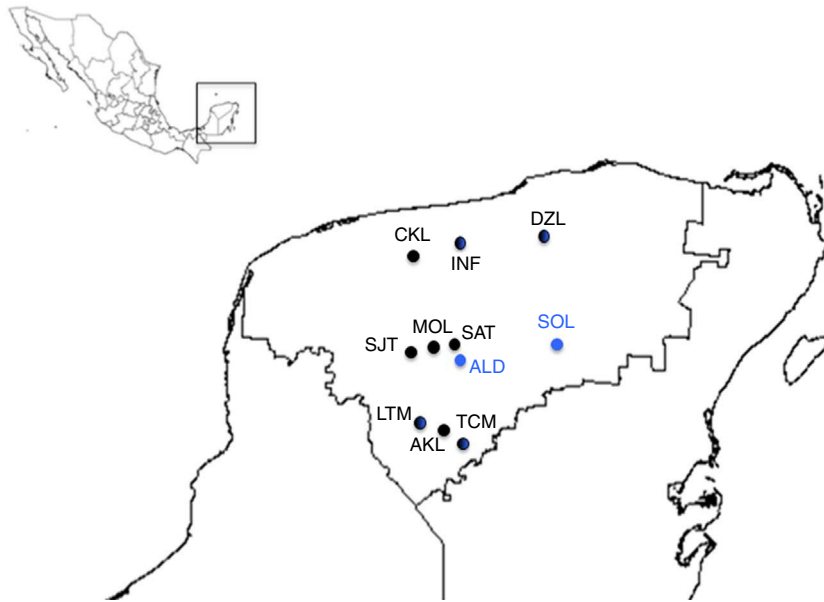
Figure 1. Location map of *Ruellia nudiflora* populations sampled in the Yucatán Peninsula, Mexico. Conkal (CKL), Inifap (INF), Dzemul (DZL), San José Tzal (SJT), Molas (MOL), San Antonio Tehitz (SAT), Lol-Tum (LTM), Akil (AKL), and Ticum (TCM) were sampled in 2007 (black circles) to assess genetic structure and diversity and demographic history, as well as in 2009 to determine outcrossing rates, while Aldana (ALD) and Solorio (SOL) were sampled only in 2009 to determine outcrossing rates (blue circles). (For interpretation of the references to color in this text, the reader is referred to the web version of the article.)

production is similar in both flower types (approximately 70% fruit set success) (Munguía-Rosas et al., 2012). The proportion of CH flowers in this species throughout the season is reported to be 0.43–0.47 (Munguía-Rosas, Campos-Navarrete, & Parra-Tabla, 2013). During the flowering peak in the area under study, individual plants produce 6–12 CH flowers every day, which are visited by at least 6 bee species, such as *Apis mellifera* Linnaeus (Apidae) and *Trigona fulviventris* Guérin-Méneville (Apidae), and 5 butterfly species, the most common being *Microtia elva* Bates (Nymphalidae) (Abdala-Roberts et al., 2012). CH flowers are self-compatible and of tubular shape, the 20–25-mm long corolla opening for only 1 day. Fruits from both flower types are dry and dehiscent, measure 1–2 cm in length, and produce 6–15 seeds each. Seeds from both flower types remain viable for 3–4 months. This plant rarely grows clonally; Abdala-Roberts et al. (2012) found that virtually all recruitment occurs via seeds. Species in the Acanthaceae family disperse their seeds by capsule explosion once they are mature, in both flower types (Witztuma, & Schulgasserb, 1995).

Population sampling

To estimate genetic structure and diversity, and to evaluate demographic history, nine populations were sampled: Conkal (CKL), Inifap (INF), Dzemul (DZL), San José Tzal (SJT), Molas (MOL), San Antonio Tehitz (SAT), Lol-Tum (LTM), Akil (AKL), and Ticum (TCM), as shown in Fig. 1. A total of 76 adult flowering plants (5–23 individuals from each population; Tables 1 and 2) were sampled. Since *R. nudiflora* is considered a weed, it is frequently removed by people from growing areas, rendering it impossible to obtain many individuals at some sites, and to thereby achieve an equally large sample for all populations.

Table 1

Geographic coordinates of *Ruellia nudiflora* study populations, and sampling site vegetation, mean annual temperature in °C (TEMP), mean annual precipitation in mm (PP), and mean herkogamy in mm (HER). SDF = spiny deciduous forest; DEF = deciduous forest; STF = subdeciduous tropical forest; MEF = medium-height evergreen forest.

Population	Long	Lat	Vegetation	Temp	Pp	Her
Conkal	21.0404	89.3119	DEF	25.8	851	1.167
Inifap	21.0623	89.2622	DEF	25.8	825	0.510
Dzemul	21.1204	89.1903	SDF	25.8	798	0.320
S. J. Tzal	20.4919	89.3927	DEF	25.9	946	1.116
Molas	20.4902	89.3828	DEF	25.9	946	2.181
San A. Tehitz	20.5338	89.2359	DEF	26.0	955	1.248
Lol-Tun	20.1509	89.2725	STF	25.6	933	4.685
Akil	20.1526	89.2016	STF	25.8	920	2.203
Ticum	20.0858	89.1302	STF	25.7	920	1.922
Added populations						
Aldana	20.4742	89.3742	DEF	25.8	978	1.460
Solorio	20.5746	89.3661	MEF	25.9	1060	1.560

To define the climatic gradient along which these *R. nudiflora* populations that are found, mean annual temperature and mean annual precipitation in these same 9 sites were obtained from WorldClim v1.4 (Hijmans, Cameron, Parra, Jones, & Jarvis, 2005). This database uses historical records from 1950 to 2000 (Table 1) and provides weather surfaces from which global climate data with a spatial resolution of approximately 1 km^2 can be extracted. Mean temperature and precipitation were selected since these variables have been reported to influence genetic diversity and structure in ruderal plants (Hamasha et al., 2013). For each population, the surrounding vegetation type was characterized according to Duch-Gary (1988) (Table 1).

In 2007 during June and July (the months when the flowering peak typically occurs), the same nine populations were

Table 2

Genetic diversity in nine populations of *Ruellia nudiflora* from the Yucatán Peninsula. n = sample size; n_e = effective number of alleles; h = Nei's genetic diversity index; I = Shannon's index of diversity; polymorphic loci (%); DW = frequency-down-weighted-marker value.

Population	n	n_e	h	I	% polymorphic loci	DW
Conkal	5	1.271 ± 0.289	0.177 ± 0.164	0.281 ± 0.242	76.82	87.51
Inifap	23	1.259 ± 0.251	0.178 ± 0.141	0.297 ± 0.199	59.75	91.04
Dzemul	4	1.231 ± 0.316	0.143 ± 0.177	0.221 ± 0.262	40.24	60.24
San José Tzal	4	1.268 ± 0.326	0.166 ± 0.180	0.257 ± 0.264	47.56	103.33
Molas	8	1.253 ± 0.252	0.173 ± 0.147	0.283 ± 0.218	78.04	109.55
San A. Tehuitz	13	1.275 ± 0.283	0.182 ± 0.155	0.296 ± 0.221	65.85	77.98
Lol-Tum	5	1.233 ± 0.278	0.153 ± 0.165	0.241 ± 0.249	62.19	87.20
Akil	6	1.184 ± 0.265	0.122 ± 0.154	0.197 ± 0.232	42.68	46.11
Ticum	8	1.234 ± 0.239	0.163 ± 0.139	0.273 ± 0.204	67.07	75.05

visited and 5–23 plants were selected from each for sampling. Selected plants were similar in size and floral display, free of folivory, and at a minimum distance of 2 m from each other. Three chasmogamous flowers were randomly selected from each of these plants and herkogamy was determined as the height of the lowest stigma lobe minus the height of the highest anther, measured from the apical tip of the ovary using a digital caliper (Fisher; Grass Valley, CA; accuracy ±0.02 mm/±0.001 mm). Measurements were made between 8:00 and 10:00 h.

Evaluation of the outcrossing rate in *R. nudiflora* requires an array of genotypes of siblings from one maternal plant (Ritland, 2002). Therefore, in June and July 2009, fruits of CH flowers were collected from 22 plants in six of the populations (3–5 fruits per maternal plant); not enough seeds were found in some populations for the outcrossing rate analysis and therefore some sites could not be adequately sampled. Four of these populations (INF, DZL, LTM, TCM) had been sampled previously to estimate genetic diversity and two more were added: Aldana (ALD) and Solorio (SOL). Seeds taken from these fruits were germinated in plastic growing trays. After the seeds had sprouted (approximately 90% germination), seedlings were transplanted to 2-L pots filled with commercial substrate. All seeds from one maternal plant were considered as a genetic family. After 6 months of growth, leaf tissue was collected per genetic family, from a total of 108 individuals and 22 families (see column 2, Table 5).

It should be noted that the nine populations selected to estimate herkogamy, genetic diversity, and demographic history adequately represent the bioclimatic gradient and along which *R. nudiflora* occur in the Yucatán Peninsula (i.e., three populations per each of three bioclimatic regions); this particular bioclimatic gradient in Yucatán has been described previously and was used to explain the distribution of genetic diversity and structure in other plant species in the region (Martínez-Natarén, Parra-Tabla, Ferrer-Ortega, & Calvo-Irabién, 2013 and references therein). Similarly, the six populations selected in order to evaluate the mating system represent the intermediate and two extreme values of variation in herkogamy found in this species (Marrufo, 2011). In particular, herkogamy in *R. nudiflora* is reported to follow a pattern of reduced anther-stigma separation in the north to greater separation in the south (Marrufo, 2011). The two selected populations therefore represent the "degree" of herkogamy along this gradient: DZL

and INF in the north have reduced anther-stigma separation (0.32 ± 0.48 mm and 0.51 ± 0.5 mm, mean ± SD, respectively), the two central region populations an intermediate separation (ALD 1.46 ± 1.00 mm and SOL 1.56 ± 0.51 mm), and the two southern populations a greater separation (TCM 2.99 ± 0.51 mm and LTM 3.69 ± 1.5 mm) (Marrufo, 2011).

DNA extraction and AFLP protocol

Total genomic DNA of adult plants and seedlings was isolated from 200 mg lyophilized leaf material by the mini-prep CTAB method (Doyle, & Doyle, 1990). Isolated DNA was purified using the Wizard® DNA Clean-Up kit (Promega, Madison, WI) according to manufacturer instructions. The quality of the genomic DNA was verified in 1× TBE agarose gels and quantification of each sample was made in an Implen Nanophotometer (Thermo Fisher Scientific; Toronto, Canada).

Table 3

(A) Deviance information criterion (DIC) values for each test model of *Ruellia nudiflora* in the Yucatán Peninsula. (B) Bayesian analysis of genetic structure. θ^{II} and θ^{III} represent scaled allele frequency variances, both reflect only differentiation among contemporaneous populations; $G_{ST} – B$ is the Bayesian version of G_{ST}. Values are the mean and 2.5% and 97.5% confidence intervals for all populations (full model only). (C) Population genetic structure inferred by analysis of molecular variance (Amova).

(A)

	Model			
	Full model	$f = 0$	$\theta = 0$	Free model
DIC	2402.29	2405.19	2427.52	2482.98

(B)

	Mean	(2.5–97.5%)
θ^{II}	0.010	(0.001–0.022)
θ^{III}	0.008	(0.001–0.019)
G_{ST} - B	0.009	(0.001–0.019)

(C)

Source of variation	DF	Sum of squares	Variance components	% Variation	Φ-Statistic
Among populations	8	114.081	0.218	1.62%	$\Phi_{ST} = 0.016$
Within populations	67	942.912	13.866	98.38%	$P < 0.05$

Amplified fragment length polymorphism (AFLP) analysis was performed using a modified version of the protocol proposed by Vos et al. (1995): 250 ng genomic DNA (1.4–13.4 µL) was digested in a 20-µL reaction with 0.2 and 0.25 µL of the restriction enzymes *Mse*I and *Eco*RI respectively, 0.2 µL BSA, 2 µL NE Buffer 2 (both reagents from New England Biolabs; Ipswich, MA), and water to attain the final volume. The reaction was maintained at a constant temperature of 36 °C for 2 h. Adaptor ligation was performed in a 10-µL reaction with 1.08 µL of each of the 100 µM *Eco*RI and *Mse*I adapters, 1 µL T4 DNA ligase, 5 µL ligase buffer (both reagents from New England Biolabs), and 3.86 µL water. Ligation occurred at 16 °C for 1 h, and the ligase reaction was inactivated at 65 °C. Samples were then stored at 4 °C.

Pre-amplification was performed in 20-µL reactions using the restriction-ligation products and the primers *Eco*RI + A and *Mse*I + C. The temperature profile for amplification was 30 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 2 min. A total of 18 selective primer combinations were used to amplify DNA from adult plants and seedlings, based on which the following primer combinations were selected: *Eco*RI + ACC with *Mse*I + CTT and *Eco*RI + ACT with *Mse*I + CGT, for which a larger number of loci and a clearer band pattern were obtained. Adapters and oligonucleotides used were obtained from MWG Biotech (Huntsville, AL).

Bands of different molecular weight were identified in a transilluminator (Ultra-Violet Products; Upland, CA) using as a reference the band patterns of a marker of known molecular weight (PROMEGA® Lambda DNA/*Hind*III Markers, Promega). A presence/absence (1/0) matrix was constructed for use in these analyses.

To find the scoring error rate and genotype repeatability, eight individuals were selected at random from among all populations, and their band patterns were evaluated twice. The proportion of unshared bands in these two assays averaged 20%. Since this is a high value, bands <100 bp and >1,000 bp (i.e., those with a higher probability of evidencing homoplasy) were eliminated, thus reducing the scoring error rate to 5.54% on average.

To identify outlier loci among the above-mentioned markers, a Bayesian analysis to estimate the posterior probability of selection was performed with BAYESCAN v2.01 (Foll & Gaggiotti, 2008). This program calculates a Bayes factor (BF), which indicates that there is conclusive evidence in favor of selection when BF > 2, while with BF values equal to or close to 1 the null hypothesis of neutrality cannot be rejected (Foll & Gaggiotti, 2008). None of the 104 loci analyzed in *R. nudiflora* showed conclusive evidence of selection, since all BF values were between 0.8 and 1.2. Therefore, all genetic diversity and structure analyses were run on the full set of 104 loci.

Data analysis

Overall mean proportion of polymorphic loci (as a percentage), effective number of alleles (n_e), Nei's (1978) genetic diversity index (h), and Shannon's index of diversity (I) were estimated in the nine populations, using POPGENE v1.32 (Yeh & Boyle, 1997). To quantify the distribution of rare alleles, we

calculated the frequency-down-weighted-marker value (*DW*); a high value of this index for a given population indicates a high level of genetic diversity distinctiveness (Schönswetter, & Tribsch, 2005). Furthermore, to complement *DW* data, the degree of genetic divergence among populations and gene flow (M = *Nm*) in each population pair were estimated with ARLEQUIN v3.5.1.2 software (Excoffier, Laval, & Schneider, 2005).

Genetic structure was estimated using a Bayesian approach implemented with HICKORY v1.1 (Holsinger, & Lewis, 2002), which permits evaluation of different models and selection of the one with the best fit (full model, $f = 0$, $\theta = 0$ and free model), i.e., the one with the lowest deviance information criterion (DIC) (Holsinger, & Lewis, 2002). HICKORY v1.1 was also used to estimate the parameters G_{ST}-B, θ^{II} and θ^{III}, which are Bayesian analogs of F_{ST}, to enable contemporary population differentiation rather than historical differentiation to be tested (Holsinger, & Lewis, 2002). These analyses were performed using the following conditions for all four models: burn-in = 5,000, number of iterations = 100,000, thinnin = 20. Estimates were made using all sampled populations. Analysis of molecular variance (Amova) was also performed with ARLEQUIN v3.5.1.2 software (Excoffier et al., 2005); additionally, paired F_{ST} values and gene flow were estimated for all possible combinations.

To assess if each plant could be genetically assigned to a given population, STRUCTURE v2.3 (Earl, 2011; Pritchard, Stephens, & Donnelly, 2010) was used. Similarly, to determine the number of homogeneous groups that sampled populations can cluster into, a Bayesian search was performed using this same program with 20 replicates, considering $K = 1$–9 as the number of possible groups. Additionally, a neighbor-joining (NJ) tree was constructed with the matrix of Nei's (1978) genetic distance, and principal coordinates analysis (PCoA) was carried out using the genetic frequencies of each population; both were done with PAST v.2.15 software (Hammer, Harper, & Ryan, 2001).

To determine if populations were undergoing demographic or spatial expansion, the following were estimated: *1)* distribution of paired differences between individuals (mismatch distribution), and *2)* number of segregating sites within a population and changes in population size (Rogers, & Harpending, 1992). Populations that have had sustained rapid demographic expansion are expected to have Poisson-type distributions of paired differences while multimodal distributions are expected in populations that are in demographic equilibrium (Rogers, & Harpending, 1992). Distribution was analyzed per population by estimating the sum of squared deviations observed over that expected based on a demographic expansion model (*SSD*), as well as by Harpending's raggedness index (spatial expansion model) using the program ARLEQUIN v3.01 (Excoffier et al., 2005). The standard error was estimated based on 1000 resampling replicates. The peak of mismatch distribution provides an estimate of τ, initial time of expansion. Mismatch distribution also provided θ_0 and θ_1 values, which are approximations of the initial and final effective population size, respectively (Excoffier et al., 2005). The standard error and 95% confidence interval were estimated based on 5,000 resampling replicates.

To evaluate potential correlations between neutral genetic differentiation (paired F_{ST}) and geographic distance, environmental factors and phenotypic divergence, several Mantel and partial Mantel tests were performed (Crispo, Bentzen, Reznicl, Kinnison, & Hendry, 2006). The partial Mantel test evaluates the correlation of matrices X and Y, controlling for similarities given in a third matrix Z (Mantel, & Valand, 1970). The following matrices were compared: (i) geographic distance yielded a spatial genetic pattern (IBD model) when paired F_{ST} and paired distance in kilometers were compared (mean temperature and mean precipitation were included in the partial Mantel test; Appendix, matrices A and B); (ii) environmental variables (mean temperature and precipitation; Appendix, matrices D and E) and genetic distance (IBE); geographic distance was included in the partial Mantel, and (iii) phenotypic differences (herkogamy; Appendix, matrix F) and genetic pattern; geographic distance was included in the partial Mantel (Sexton, Hangartner, & Hoffmann, 2013). To obtain matrices of environmental, phenotypic, and geographic variables, pairwise Euclidean distance between populations was estimated using PAST v2.15 (Hammer et al., 2001). The Mantel and partial Mantel tests were carried out with XLSTAT v2014.6.01 (Microsoft Excel); the *p*-value was estimated by using *r* (AB) to create a distribution from 10,000 permutations.

AFLP markers have been used to estimate outcrossing rates in numerous plant species regardless of allelic dominance (Gaiotto, Bramucci, & Grattapaglia, 1997; see references in Meudt, & Clarke, 2007). Values in the seedling presence/absence matrix were recoded to classify dominant (1) and recessive (2) genotypes in each band and individual, for use with MLRT v3.4 (Ritland, 2002). Based on the correlated-mating model proposed by Ritland (2002) assuming a mixed mating system model, the following mating system indices were obtained: single-locus outcrossing rate (t_s), multilocus outcrossing rate (t_m), biparental inbreeding ($t_m - t_s$), multilocus correlation of outcrossed paternity (r_p), and correlation of selfing (r_s). The latter 2 indices are based on a sib-pair arrangement, which was used as the observation unit. In this case, r_p estimates the probability that a given pair of plants are full-sibs due to an outcrossing event, and r_s the probability that they are full-sibs due to a selfing event (Ritland, 2002). Hence, a high r_s value has been previously correlated with pollen limitation (Takebayashi, Wolf, & Delph, 2006). Confidence intervals were estimated by bootstrap resampling (1,000 iterations) of progeny within each family.

Results

Genetic diversity

The overall mean proportion of polymorphic loci was $60.73 \pm 13.88\%$ (mean $\pm$ SD). At the population level, the percentage of polymorphic loci fluctuated between 40.24% and 78.04% (Table 2). The overall mean effective number of alleles (n_e) was 1.245 ± 0.029 (mean $\pm$ SD). The overall mean for Nei's genetic diversity index (h) was 0.162 ± 0.02 (mean $\pm$ SD), the highest value occurring in SAT and the lowest in AKL. For Shannon's index (I), the overall mean was 0.261 ± 0.035

(mean $\pm$ SD), with the same two populations having the highest (SAT) and lowest (AKL) values. As regards *DW*, the highest values occurred in MOL, SJT, and INF (Table 2), while the lowest was found in AKL and DZL. Mean values for this index indicate that the latter two populations and SAT may have the longest divergent times while CKL, MOL, TCM, SJT, and LTM have recent divergence times (Table 2). All *Nm* estimates were >1.0 individuals; CKL and LTM had the lowest gene flow values ($Nm = 1.585$). In 22 comparisons between populations, *Nm* had an infinite value while in the remaining 13, $Nm = 1.945–60.625$ individuals (Appendix, matrix C).

Genetic structure

Bayesian analyses indicate differences among populations (Table 3A), as evidenced by the full model (the model with the lowest DIC) in which θ^{II}, θ^{III} and $G_{ST} - B$ were very low, albeit significantly different from zero (Table 3). Similarly, the Amova indicates genetic differences among populations ($\Phi_{ST} = 0.016$, $p < 0.05$), which account for 1.62% of the total variation. However, within-population differences explain most of the variation (98.38%). The paired F_{ST} matrix showed low levels of differentiation in almost all comparisons, suggesting a high rate of gene flow in the Yucatán Peninsula, as stated above (Appendix, matrices B and C).

The analysis performed with STRUCTURE v2.3.3 obtained an optimum number of clusters at $\Delta K = 3$. Also, as shown in Fig. 2, populations were not observed to be genetically homogeneous groups of individuals; in other words, individuals from the same predefined population had partial membership in multiple populations. NJ and PCoA showed no clear geographic clustering of sampled populations or of individual plants (Fig. 2). This suggests that there are individuals that have ancestral relationships with plants from populations other than the one in which they were collected, which in turn provides indirect evidence of some degree of genetic connectivity.

Demographic history

Mismatch distribution (*SSD* index) and Harpending's raggedness index showed that all populations have undergone demographic and spatial expansion (Table 4) except DZL ($SSD = 0.202$, $p = 0.04$). Also, the peak of the mismatch distribution ranged from 19.2 (*IC* 10.6–30.6) in AKL to 40.7 (*IC* 18.9–178.3) in DZL, although the most common value was 27, indicating that population growth is not recent. Overall, these results show a drastic expansion of all populations, each population starting with a few dozen individuals (θ_0 ranged from 0.7 to 11.0; Table 4) and now comprising thousands (θ_1 ranged from 14,456.3 to 91,862.4) as shown in Table 4. These values are approximate, but such gross estimation suggests large demographic expansion along the climatic gradient.

Mantel and partial Mantel tests

Results of the 3 Mantel and partial Mantel tests were as follows: (i) IBD model: a moderate and significant correlation

Table 4
Mean values of tau (τ), theta 0 (θ_0) and theta 1 (θ_1) in populations of *Ruellia nudiflora*. The 95% confidence interval is shown in parentheses. *SSD* = sum of squared deviations.

Population	τ	θ_0	θ_1	*SSD*	Raggedness index
Conkal	27.2 (16.5–40.7)	4.0 (0.0–30.8)	77,431.6 (129.0–99,999.0)	0.119*ns*	0.300*ns*
Inifap	27.3 (20.6–33.2)	0.7 (0.0–6.34)	14,456.3 (88.0–99,999.0)	0.004*ns*	0.009*ns*
Dzemul	40.7 (18.9–178.3)	1.4 (0.0–13.2)	91,862.4 (179.3–99,999.0)	0.202*	0.278*ns*
San Jose Tzal	27.3 (17.7–37.9)	2.0 (0.0–14.1)	88,672.8 (122.7–99,999.0)	0.122*ns*	0.222*ns*
Molas	25.6 (16.5–38.2)	4.3 (0.0–24.5)	65,554.0 (111.9–99,999.0)	0.021*ns*	0.041*ns*
San A. Tehuitz	27.0 (18.4–37.0)	3.9 (0.0–23.7)	26,108.4 (87.1–99,999.0)	0.014*ns*	0.033*ns*
Lol-Tum	24.5 (13.2–39.2)	6.1 (0.0–36.1)	66,536.0 (93.0–99,999.0)	0.085*ns*	0.160*ns*
Akil	19.2 (10.6–30.6)	4.6 (0.0–29.2)	70,758.8 (78.8–99,999.0)	0.040*ns*	0.098*ns*
Ticum	23.9 (12.1–43.1)	11.0 (0.0–46.0)	36,829.3 (81.3–99,999.0)	0.032*ns*	0.064*ns*

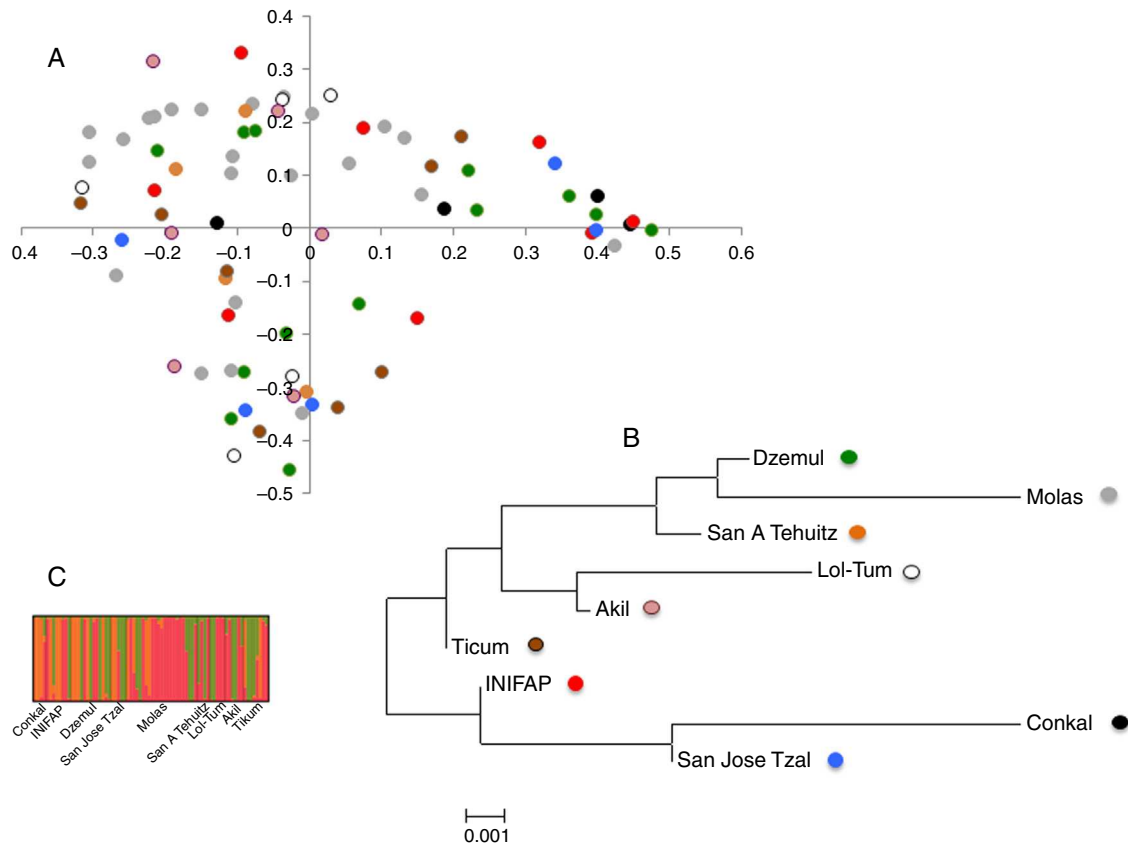* $p < 0.05$, *ns* = not significant.

Figure 2. (A) Plot of first 2 axes of Principal Coordinates Analysis (PCoA) of sampled individuals; different populations are represented by different colors; (B) neighbor-joining tree of nine populations sampled, the same colors are used as in the PCoA, and (C) results from STRUCTURE analysis with $K = 3$, each individual is represented by a thin vertical line partitioned into different colored segments that represent membership in a given population, in the north (orange), central (red) and south (green) of the Yucatán Peninsula. (For interpretation of the references to color in this text, the reader is referred to the web version of the article.)

Table 5
Mating system indices in *Ruellia nudiflora* populations. (n_{fam}) family number, (n_{seed}) seedling number, (t_m) multilocus outcrossing rate, (t_s) single-locus outcrossing rate, ($t_m - t_s$) biparental inbreeding, (r_p) multilocus correlation of outcrossed paternity, (r_s) correlation of selfing. Values in parentheses are the standard error.

| Population	n_{fam}/n_{seed}	t_m	t_s	$t_m - t_s$	r_p	r_s
Dzemul	3/14	0.998 (0.226)	0.899 (0.256)	0.100 (0.215)	0.090 (0.351)	0.110 (0.189)
Inifap	4/18	0.998 (0.130)	0.899 (0.157)	0.100 (0.109)	0.029 (0.218)	0.105 (0.063)
Aldana	3/17	0.999 (0.243)	0.913 (0.251)	0.086 (0.118)	0.029 (0.260)	0.109 (0.261)
Solorio	5/31	0.999 (0.076)	0.913 (0.102)	0.086 (0.074)	0.080 (0.469)	0.116 (0.063)
Lol-Tum	3/14	0.998 (0.229)	0.907 (0.277)	0.092 (0.120)	0.036 (0.173)	0.106 (0.170)
Ticum	4/14	0.998 (0.146)	0.903 (0.128)	0.095 (0.077)	0.010 (0.135)	0.107 (0.121)
Total	22/108	0.988 (0.168)	0.973 (0.096)	0.015 (0.086)	0.041 (0.151)	0.109 (0.010)
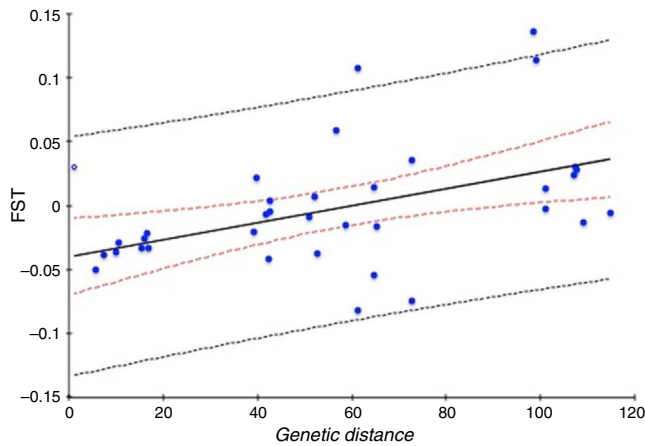
Figure 3. Relationship between geographic distance (in km) and paired F_{ST} estimates. The 95% confidence interval (dotted line) and 99% confidence interval (solid line) are shown ($r = 0.439$, $p = 0.009$).

was found between F_{ST} and geographic distance; when environmental variables were included in the partial correlation ($r = 0.439$, $p = 0.009$), geographic distance explains up to 19% of variance, and when these variables were removed ($r = 0.416$, $p = 0.012$), explained up to 17% of the variance (Fig. 3); (ii) effect of environmental variable matrices: mean temperature ($r = 0.010$, $p = 0.944$) and mean precipitation ($r = -0.224$, $p = 0.182$) had no significant IBE effect on the genetic matrix both when geographic distance was included in the partial correlation and when it was removed (mean temperature $r = 0.015$, $p = 0.896$; and mean precipitation $r = 0.162$, $p = 0.346$); (iii) Phenotypic differences (herkogamy) were not correlated with genetic distance when geographic distance was included ($r = 0.208$, $p = 0.234$) and a weak, almost significant correlation was found when geographic distance was removed ($r = 0.308$, $p = 0.085$).

Mating system

Elevated single-locus and multilocus outcrossing rates were obtained in *R. nudiflora* ($t_m = 0.998 \pm 0.009$; $t_s = 0.983 \pm 0.012$; mean $\pm$ SE) indicating that nearly 99% of progeny were produced by noninbreeding events. The values obtained for both multilocus correlation of outcrossed paternity (r_p) and correlation of selfing (r_s) were low, suggesting that in most cases seeds from one family do not share the same father ($r_p = 0.041 \pm 0.013$; mean $\pm$ SE) and that pollen limitation is low in this species ($r_s = 0.108 \pm 0.003$; mean $\pm$ SE) (Table 5).

Discussion

Colonization of new areas usually entails loss of genetic diversity in plant species, since such events are initiated by a small number of individuals. Thus, a strong founder effect, which is related to genetic bottlenecks in recently formed populations, is to be expected (Austerlitz, Jung-Muller, Godelle, & Gouyon, 1997; Bossdorf et al., 2005). Despite this prediction, the genetic diversity observed in ruderal and invasive plant species

is usually not low (Bossdorf et al., 2005; Hamasha et al., 2013; Mandák, Zákravský, Korínková, Dostál, & Placková, 2009). This agrees with our findings in *R. nudiflora* which show no evidence of genetic erosion, although θ_0 values showed populations starting with a few dozen individuals, large demographic expansion occurred in a short period of time. This may explain why no indication was found of sampled individuals being genetically similar, highly homozygous, or clones, and all sampled plants had a distinct haplotype.

There are some possible explanations for the high genetic diversity observed in *R. nudiflora* populations. First, the number of loci used may have been high, enabling detection of a great part of the variation within each population (Sunnucks, 2000). This seems to be a common result when AFLP markers are used to measure genetic variation in plants (Lihová, Kudoh, & Marhold, 2010; Nybom, 2004; Steinger, Haldimann, Leiss, & Müller-Schärer, 2002) in contrast to findings using isozyme markers. This is a clear pattern in Cactaceae (Cornejo-Romero, Vargas-Mendoza, Valverde, & Rendón-Aguilar, 2013) and Agavaceae (Eguiarte et al., 2013) or in other ruderal plants with the same reproductive system as *R. nudiflora* (Culley, & Wolfe 2001; Imaizumi, Wang, & Tominaga, 2008). There is no similar review of genetic data in the Acanthaceae, but it is assumed that the same pattern holds true.

Secondly, high pollinator visitation rates (*A. mellifera*, *T. fulviventris*, and several butterfly species; Abdala-Roberts et al., 2012) in *R. nudiflora* may result in high pollen movement among plants. Populations in which pollen movement is high usually exhibit few inbreeding events and high levels of genetic variation (Barrett, 2003). In fact, if there are open pollination events among apomictic or clonal individuals, the level of genetic variation present produces high heterozygosity (Maynard-Smith, 1982). Likewise, *R. nudiflora* is characterized by low genetic differentiation among populations (as shown by the paired F_{ST} matrix, NJ and PCoA) and this is frequently characteristic of plants with a high genetic connectivity, i.e., low population differentiation is consistent with high-vagility species (Linhart, & Grant, 1996; Slatkin, 1993), which seems to be the case of *R. nudiflora* (Abdala-Roberts et al., 2012).

However, a low genetic differentiation was statistically significant with 1.62% of genetic differences among populations grouping in three regions. Possible explanations of this low but significant differentiation are recent origin of populations, large effective population sizes, and separate, temporally persistent, local populations (Knutsen et al., 2011). In *R. nudiflora*, persistent local populations and a recent origin may be the more plausible explanations as shown by *DW* index as well as demographic and STRUCTURE analysis results. It is not common for ruderal or invasive plants to show low inter-population differentiation (Leiss, & Müller-Schäer, 2001; Poulin et al., 2005). Thus, some potentially confounding factors such as nonrepresentative sampling cannot be discarded (Knutsen et al., 2011). Further research with a larger sample size and/or other markers may contribute to understanding this pattern.

Also, a weak but significant pattern in which genetic differentiation increased with geographic distance (IBD) was revealed by the partial Mantel test (Sexton et al., 2013; Slatkin, 1993). A similar situation has been observed in the ruderal invader plant *Microstegium vimineum*, where IBD was not very consistent throughout its distribution range (Baker, & Dyer, 2011). Under a strict IBD model, a constraint is placed by the inherent dispersal limitations of the organism and the size of its populations, but not by the nature or structure of environmental variation. In *R. nudiflora*, low similarity across distances (IBD) indicates that species expansion has been a slow diffuse spread leading to a patchy environment (Lee, & Mitchell-Olds, 2011; Leiss, & Müller-Schäer, 2001) that is the result of multiple invasions and colonization/extinction events in the Yucatán Peninsula. A similar finding has been reported in North America for the invasive plant *Silene vulgaris* (McCauley, Smith, Lisenbay, & Hsieh, 2003). In this case, *S. vulgaris* had undergone several colonization events associated with a demographic expansion process occurring after its introduction from Europe (McCauley et al., 2003). On the other hand, spatial genetic structure is expected to be stronger for maternally inherited genes (mitochondrial and chloroplast), which are dispersed through seed flow, than for nuclear genes which are dispersed by both pollen and seed flow (Ennos, 2001); in the present study only nuclear markers (AFLP) were used.

Demographic and range expansion in *R. nudiflora* may be explained by human-induced disturbances that have prevailed in southeastern Mexico since the late 19th and early 20th centuries. One effect of these anthropogenic impacts has been an increase in the surface area covered by agricultural fields (Mizrahi, Ramos-Prado, & Jiménez-Osomio, 1997). These vast abandoned areas constitute suitable habitats for colonization by ruderal species such as *R. nudiflora*. The demographic and range expansion of ruderal annual species in the Yucatán Peninsula has been suggested to coincide with human-induced disturbances beginning in the late 19th century, and even much earlier during the pre-Hispanic era (Mizrahi et al., 1997). This is a common finding in ruderal species that colonize new environments (Bossdorf et al., 2005), since they are able to persist in a given site during the stabilization phase, following which they undergo a phase of population growth that allows subsequent colonization events (Richardson, Pysek, Rejmánek, Barbour, Panetta, & West, 2000). This type of pattern has been reported in other species that invade different environments, such as *Barbarea vulgaris* (Buschmann, Edwards, & Dietz, 2005), *Eschscholzia californica* (Leger, & Rice, 2003), and *Senecio jacobaea* (Stastny, Schaffner, & Elle, 2005). Although virtually no studies have been conducted on invasive herb species with CH and CL flowers, there is evidence of other species with a similar reproductive system that follow the same pattern of "resistance" (Cheplick, 2005).

The t_m estimates obtained in the present study seem unexpectedly high, but are consistent with the high Nm values reported; this result could be explained by multiple factors such as floral traits variation and asynchrony (Barrett, 2003; Boshier, Chase, & Bawa, 1995), pollen movement patterns (Murawski, & Hamrick, 1992), genetic differences at the self-fertilization level (Warwick, Bain, Wheatcroft, & Thompson, 1989), and pollen competition inside pollen tubes (Morgan, & Barrett, 1990), although high outcrossing rates have been observed in herbs with a similar reproductive system (Culley, & Wolfe 2001; Imaizumi et al., 2008). Regarding pollen movement, a previous study indicates that *R. nudiflora* is visited by bees such as *A. mellifera* and *T. fulviventris* as well as several butterfly species (Abdala-Roberts et al., 2012). Species of the genus *Trigona* may travel up to 1.5 km when foraging (Roubik, & Aluja, 1983) and *A. mellifera* up to 13.5 km (especially in disturbed habitats). Such foraging patterns by pollinator insects partly explain the occurrence of high t_m values, high gene flow, and low genetic structure across populations (Linhart, & Grant, 1996), as observed in *R. nudiflora* in the present study. Additionally, recent data suggest that self-pollinating CH flowers are twice as likely to be aborted than outcrossing CH flowers (Abdala-Roberts et al., 2012). Furthermore, a better performance has been observed in progeny produced by outcrossing than in those derived by selfing (Munguía-Rosas et al., 2013). Thus, in *R. nudiflora*, the addition of new individuals to the population is apparently generated mostly through outcrossing events, which may help explain genetic diversity levels and the slight genetic structure observed, as occurs in the cleistogamus species *Viola pubescens* (Culley, & Wolfe 2001). On the other hand, spatial variation in key floral traits such as herkogamy may affect the mating system as well as patterns of genetic diversity and structure; however, in the case of *R. nudiflora* no clear correlation was found between herkogamy and F_{ST} (Barrett, 2003; Takebayashi et al., 2006).

The hypothesis was put forth initially that genetic differences between nearby populations of *R. nudiflora* will be great if they belong to different environments and small if they are in similar environments, as has been described in other species (Lee, & Mitchell-Olds, 2011; Leiss, & Müller-Schäer, 2001). However, our results do not support this assumption. A possible explanation of the pattern described in *R. nudiflora* may be the existence of high gene flow. An increase in the number of individuals in a population, multiple invasions, and a corresponding increase in the distribution range of a species may cause increment in connectivity between populations, thereby shaping its genetic structure within a region (Austerlitz et al., 1997; Crispo et al., 2006; Sexton et al., 2013). IBE may arise through nonrandom mating, which is the result of adaptation to different environments (Sexton et al., 2013). However, gene flow in *R. nudiflora* may run counter to patterns of environmental variation or local adaptation, being higher across dissimilar environments. Under this scenario, adaptation may be difficult under stable conditions, but can be faster under rapid environmental change if strong gene flow across different environments maintains high genetic variation (Lee, & Mitchell-Olds, 2011; Leiss, & Müller-Schäer, 2001; Sexton et al., 2013). As demonstrated by Ortegón-Campos et al. (2012), *R. nudiflora* is not adapting rapidly to the local soil environment.

Finally, the lack of a clear pattern in the present data suggests that this plant has recently and successively colonized the Yucatán Peninsula area. To have a clearer idea of this situation, a phylogeographic study that might clarify the evolutionary history of *R. nudiflora* is required.

Acknowledgments

Appendix.

Matrices used in Mantel and partial Mantel test. Names of populations are: Conkal (CKL), Inifap (INF), Dzemul (DZL), San José Tzal (SJT), Molas (MOL), San Antonio Tehitz (SAT), Lol-Tum (LTM), Akil (AKL) and Ticum (TCM).

(A) Geographic distance in kilometers.

	CKL	INF	DZL	SJT	MOL	SAT	LTM	AKL	TCM
CKL	0	5.71	15.44	61.33	61.38	56.66	98.59	98.99	107.40
INF	5.71	0	9.86	64.62	64.59	58.59	100.94	100.94	109.01
DZL	15.44	9.86	0	72.73	72.62	65.14	107.71	107.18	114.74
SJT	61.33	64.62	72.73	0	1.05	17.01	39.80	42.55	52.68
MOL	61.38	64.59	72.62	1.05	0	16.07	39.30	41.91	51.98
SAT	56.66	58.59	65.14	17.01	16.07	0	42.57	42.36	50.82
LTM	98.59	100.94	107.71	39.80	39.30	42.57	0	7.42	16.54
AKL	98.99	100.94	107.18	42.55	41.91	42.36	7.42	0	10.51
TCM	107.40	109.01	114.74	52.68	51.98	50.82	16.54	10.51	0

(B) Matrix of F_{ST} values.

	CKL	INF	DZL	SJT	MOL	SAT	LTM	AKL	TCM
CKL	0	−0.0502	−0.0337	−0.0816	0.1077	0.0592	0.1362	0.1138	0.0304
INF	−0.0502	0	−0.0365	−0.0541	0.0141	−0.0154	−0.0023	0.0132	−0.0136
DZL	−0.0337	−0.0365	0	−0.0743	0.0356	−0.0166	0.0279	0.0237	−0.0059
SJT	−0.0816	−0.0541	−0.0743	0	0.0303	−0.0328	0.0220	−0.0044	−0.0373
MOL	0.1077	0.0141	0.0356	0.0303	0	−0.0258	−0.0206	−0.0071	0.0074
SAT	0.0592	−0.0154	−0.0166	−0.0328	−0.0258	0	0.0041	−0.0413	−0.0091
LTM	0.1362	−0.0023	0.0279	0.0220	−0.0206	0.0041	0	−0.0387	−0.0213
AKL	0.1138	0.0132	0.0237	−0.0044	−0.0071	−0.0413	−0.0387	0	−0.0289
TCM	0.0304	−0.0136	−0.0059	−0.0373	0.0074	−0.0091	−0.0213	−0.0289	0

(C) Gene flow ($Nm = [1 − F_{ST}]/[4*F_{ST}]$) between populations.

	CKL	INF	DZL	SJT	MOL	SAT	LTM	AKL	TCM
CKL	–	inf	inf	inf	2.070	3.975	1.585	1.945	7.980
INF	inf	–	inf	inf	17.485	inf	inf	18.650	inf
DZL	inf	inf	–	inf	6.770	inf	8.695	10.315	inf
SJT	inf	inf	inf	–	8.010	inf	11.100	inf	inf
MOL	2.070	17.485	6.770	8.010	–	inf	inf	inf	33.525
SAT	3.975	inf	inf	inf	inf	–	60.625	inf	inf
LTM	1.585	inf	8.695	11.100	inf	60.625	–	inf	inf
AKL	1.945	18.650	10.315	inf	inf	inf	inf	–	inf
TCM	7.980	inf	inf	inf	33.525	inf	inf	inf	–

(D) Temperature Euclidean distance among populations.

	CKL	INF	DZL	SJT	MOL	SAT	LTM	AKL	TCM
CKL	0	0	0	0.1	0.1	0.2	0.2	0	0.1
INF	0	0	0	0.1	0.1	0.2	0.2	0	0.1
DZL	0	0	0	0.1	0.1	0.2	0.2	0	0.1
SJT	0.1	0.1	0.1	0	0	0.1	0.3	0.1	0.2
MOL	0.1	0.1	0.1	0	0	0.1	0.3	0.1	0.2
SAT	0.2	0.2	0.2	0.1	0.1	0	0.4	0.2	0.3
LTM	0.2	0.2	0.2	0.3	0.3	0.4	0	0.2	0.1
AKL	0	0	0	0.1	0.1	0.2	0.2	0	0.1
TCM	0.1	0.1	0.1	0.2	0.2	0.3	0.1	0.1	0

(E) Precipitation Euclidean distance among populations.

	CKL	INF	DZL	SJT	MOL	SAT	LTM	AKL	TCM
CKL	0	26	53	95	95	104	82	69	69
INF	26	0	27	121	121	130	108	95	95
DZL	53	27	0	148	148	157	135	122	122
SJT	95	121	148	0	0	9	13	26	26
MOL	95	121	148	0	0	9	13	26	26
SAT	104	130	157	9	9	0	22	35	35
LTM	82	108	135	13	13	22	0	13	13
AKL	69	95	122	26	26	35	13	0	0
TCM	69	95	122	26	26	35	13	0	0

(F) Herkogamy Euclidean distance among populations.

	CKL	INF	DZL	SJT	MOL	SAT	LTM	AKL	TCM
CKL	0	0.657	0.847	0.051	1.014	0.081	3.518	1.036	0.755
INF	0.657	0	0.19	0.606	1.671	0.738	4.175	1.693	1.412
DZL	0.847	0.19	0	0.796	1.861	0.928	4.365	1.883	1.602
SJT	0.051	0.606	0.796	0	1.065	0.132	3.569	1.087	0.806
MOL	1.014	1.671	1.861	1.065	0	0.933	2.504	0.022	0.259
SAT	0.081	0.738	0.928	0.132	0.933	0	3.437	0.955	0.674
LTM	3.518	4.175	4.365	3.569	2.504	3.437	0	2.482	2.763
AKL	1.036	1.693	1.883	1.087	0.022	0.955	2.482	0	0.281
TCM	0.755	1.412	1.602	0.806	0.259	0.674	2.763	0.281	0

References

Abdala-Roberts, L., Marrufo-Zapata, D., Arceo-Gómez, G., & Parra-Tabla, V. (2012). Pollen limitation, fruit abortion, and autonomous selfing in three populations of the perennial herb *Ruellia nudiflora*. *Plant Species Biology*, http://dx.doi.org/10.1111/j.1442-1984.2012.00392.x

Alexander, J. M., Van Kleunen, M., Ghezzi, R., & Edwards, P. J. (2012). Different genetic clines in response to temperature across the native and introduced ranges of a global plant invader. *Journal of Ecology*, *100*, 771–781.

Austerlitz, F., Jung-Muller, B., Godelle, B., & Gouyon, P. H. (1997). Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Populations Biology*, *51*, 148–164.

Baker, S. A., & Dyer, R. J. (2011). Invasion genetics of *Microstegium vimineum* (Poaceae) within the James River Basin of Virginia, USA. *Conservation Genetics*, *12*, 793–803.

Barrett, S. C. H. (2003). Mating strategies in flowering plants: The outcrossing-selfing paradigm and beyond. *Philosophical Transactions of the Royal Society Series B*, *358*, 991–1004.

Boshier, D. H., Chase, M. R., & Bawa, K. S. (1995). Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 3. Gene flow, neighborhood, and population substructure. *American Journal of Botany*, *82*, 484–490.

Bossdorf, O., Auge, H., Lafuma, L., Rogers, W. E., Siemann, E., & Prati, D. (2005). Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia*, *144*, 1–11.

Buschmann, H., Edwards, P. J., & Dietz, H. (2005). Variation in growth pattern and response to slug damage among native and invasive provenances of four perennial *Brassicaceae* species. *Journal of Ecology*, *93*, 322–334.

Campbell, C. S., Quinn, J. A., Cheplick, J. P., & Bell, T. J. (1983). Cleistogamy in grasses. *Annual Review of Ecology and Systematics*, *14*, 411–441.

Cervera, J. C., & Parra-Tabla, V. (2009). Seed germination and seedling survival traits of invasive and non-invasive congeneric *Ruellia* species (Acanthaceae) in Yucatan. Mexico. *Plant Ecology*, *205*, 285–293.

Cheplick, G. P. (2005). Biomass partitioning and reproductive allocation in the invasive, cleistogamous grass *Microstegium vimineum*: Influence of the light environment. *Journal of the Torrey Botanical Society*, *132*, 214–224.

Cornejo-Romero, A., Vargas-Mendoza, C. F., Valverde, P., & Rendón-Aguilar, B. (2013). Estructura genética y filogeografía en cactáceas. *Cactáceas y Suculentas Mexicanas*, *58*, 4–28.

Crispo, E., Bentzen, P., Reznicl, D. N., Kinnison, M. T., & Hendry, A. P. (2006). The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, *15*, 49–62.

Culley, T. M., & Wolfe, A. D. (2001). Population genetic structure of the cleistogamous plant species *Viola pubescens* Aiton (Violaceae), as indicated by allozyme and ISSR molecular markers. *Heredity*, *86*, 545–556.

Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, *12*, 13–15.

Duch-Gary, J. (1988). *La conformación territorial del estado de Yucatán*. México D.F.: Universidad Autónoma Chapingo.

Earl, D. A. (2011). *Structure Harvester v0.6.7 [homepage on the Internet]*. Retrieved from http://users.soe.ucsc.edu/~dearl/software/structure Harvester/

Eguiarte, L. E., Aguirre-Planter, E., Aguirre, X., Colín, R., González, A., Rocha, M., et al. (2013). From isozymes to genomics: Population genetics and conservation of *Agave* in Mexico. *Botanical Review*, *79*, 483–506. http://dx.doi.org/10.1007/s12229-013-9123-x

Ennos, R. A. (2001). Inferences about spatial processes in plant populations from the analysis of molecular markers. In J. Silvertown, & J. Antonovics (Eds.), *Integrating ecology and evolution in a spatial contex* (pp. 45–71). Oxford: Blackwell Science Ltd.

Excoffier, L., Laval, G., & Schneider, S. (2005). ARLEQUIN (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, *1*, 47–50.

Flores, J. S., & Espejel-Carvajal, I. (1994). Tipos de vegetación de la península de Yucatán. In *Etnoflora Yucatanense, Fascicle 3*. Mérida: Universidad Autónoma de Yucatán.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*, 977–993.

Gaiotto, F. A., Bramucci, M., & Grattapaglia, D. (1997). Estimation of outcrossing rate in a breeding population of *Eucalyptus urophylla* with dominant RAPD and AFLP markers. *Theoretical and Applied Genetic*, *95*, 842–849.

Grime, J. P. (1979). *Plant strategies and vegetation processes*. Chichester, UK: John Wiley and Sons.

Hamasha, H. R., Schmidt-Lebuhn, A. N., Durka, W., Schleuning, M., & Hensen, I. (2013). Bioclimatic regions influence genetic structure of four Jordanian Stipa species. *Plant Biology*, *15*, 882–891.

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, *4*, 9. http://folk.uio.no/ohammer/past

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978.

Holsinger, K. E., & Lewis, P. O. (2002). A Bayesian approach to inferring population structure from dominant markers. *Molecular Ecology*, *11*, 1157–1164.

Imaizumi, T., Wang, G. X., & Tominaga, T. (2008). Pollination of chasmogamous flowers and the effects of light and emergence time on chasmogamy and cleistogamy in *Monochoria vaginalis*. *Weed Biology and Management*, *8*, 260–266.

Knutsen, H., Olsen, E. M., Jorden, P. E., Espeland, S. H., André, C., & Stenseth, N. C. (2011). Are low but statistically significant levels of genetic differentiation in marine fishes 'biologically meaningful'? A case study of coastal Atlantic cod. *Molecular Ecology*, *20*, 768–783.

Lee, C. R., & Mitchell-Olds, T. (2011). Quantifying effects of environmental and geographical factors on patterns of genetic differentiation. *Molecular Ecology*, *20*, 4631–4642.

Leger, E. A., & Rice, K. J. (2003). Invasive Californian poppies (*Eschscholzia californica* Cham.) grow larger than native individuals under reduced competition. *Ecology Letters*, *6*, 257–264.

Leiss, K. A., & Müller-Schäer, H. (2001). Adaptation of *Senecio vulgaris* (Asteraceae) to ruderal and agricultural habitats. *American Journal of Botany*, *88*, 1593–1599.

Lihová, J., Kudoh, H., & Marhold, K. (2010). Genetic structure and phylogeography of a temperate-boreal herb. *Cardamine scutata* (Brassicaceae), in northeast Asia inferred from AFLPS and cpDNA haplotypes. *American Journal of Botany*, *97*, 1058–1070.

Linhart, Y. B., & Grant, M. C. (1996). Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, *27*, 237–277.

Long, R. W. (1971). Genetic and morphological relationships of the southeastern coastal plain endemic *Ruellia noctiflora* (Acanthaceae). *Bulletin Torrey Botany Club*, *98*, 16–21.

Mandák, B., Zákravsky, P., Korínková, D., Dostál, P., & Placková, I. (2009). Low population differentiation and high genetic diversity in the invasive species *Carduus acanthoides* L. (Asteraceae) within its native range in the Czech Republic. *Biological Journal of the Linnean Society*, *98*, 596–607.

Mantel, N., & Valand, R. S. (1970). A technique of nonparametric multivariate analysis. *Biometrics*, *26*, 547–558.

Marrufo, D. Z. (2011). *Variación geográfica en la hercogamia y tasas de entrecruzamiento poblacional en Ruellia nudiflora (Acanthaceae) en la península de Yucatán (MS diss)*. Mérida: Universidad Autónoma de Yucatán.

Martínez-Natarén, D., Parra-Tabla, V., Ferrer-Ortega, M., & Calvo-Irabién, L. M. (2013). Genetic diversity and genetic structure in wild populations of Mexican oregano (*Lippia graveolens* HBK) and its relationship with the chemical composition of the essential oil. *Plant Systematics and Evolution*, *299* http://dx.doi.org/10.1007/s00606-013-0902-y

Maynard-Smith, J. (1982). *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.

McCauley, D. E., Smith, R. A., Lisenbay, J. D., & Hsieh, C. (2003). The hierarchical spatial distribution of chloroplast DNA polymorphism across the introduced range of *Silene vulgaris*. *Molecular Ecology*, *12*, 3227–3235.

Meudt, H. M., & Clarke, A. C. (2007). Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science*, *12*, 106–117.

Meyer, J. Y., & Lavergne, C. (2004). Beautés fatales: Acanthaceae species as invasive alien plants on tropical Indo-Pacific Islands. *Diversity and Distributions*, *10*, 333–347.

Mizrahi, A., Ramos-Prado, J. M., & Jiménez-Osornio, J. (1997). Composition, structure, and management potential of secondary dry tropical vegetation in two abandoned henequen plantations of Yucatan, Mexico. *Forest Ecology and Management*, *6*, 273–282.

Morgan, M. T., & Barrett, S. C. H. (1990). Outcrossing rates and correlated mating within a population of *Eichhornia paniculata* (Pontederiaceae). *Heredity*, *64*, 271–280.

Müller-Schäer, H., Schaffner, U., & Steinger, T. (2004). Evolution in invasive plants: implications for biological control. *Trends in Ecology and Evolution*, *19*, 417–422.

Munguía-Rosas, M. A., Ollerton, J., & Parra-Tabla, V. (2012). Environmental control of reproductive phenology and the effect of pollen supplementation on resource allocation in the cleistogamous weed. *Ruellia nudiflora* (Acanthaceae). *Annals of Botany*, *109*, 343–350.

Munguía-Rosas, M. A., Campos-Navarrete, M. J., & Parra-Tabla, V. (2013). The effect of pollen source vs. flower type on progeny performance and seed predation under contrasting light environments in a cleistogamous herb. *PLoS ONE*, *8*, e80934. http://dx.doi.org/10.1371/journal.pone.0080934

Murawski, D. A., & Hamrick, J. L. (1992). The mating system of *Cavanillesia platanifolia* under extremes of flowering tree density: A test of predictions. *Biotropica*, *24*, 99–101.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, *89*, 583–590.

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, *13*, 1143–1155.

Ortegón-Campos, I., Abdala-Roberts, L., Parra-Tabla, V., Cervera, C., Marrufo, D., & Herrera, C. M. (2012). Influence of multiple factors on plant local adaptation: Soil type and folivore effects in *Ruellia nudiflora* (Acanthaceae). *Evolutionary Ecology*, *26*, 545–558.

Pino, J., Font, X., de Cáceres, M., & Molowny-Horas, R. (2009). Floristic homogenization by native ruderal and alien plants in north-east Spain: The effect of environmental differences on a regional scale. *Global Ecology and Biogeography*, *18*, 563–574.

Poulin, J., Weller, S. G., & Skal, A. A. (2005). Genetic diversity does not affect the invasiveness of fountain grass (*Pennisetum setaceum*) in Arizona. California and Hawaii. *Diversity and Distributions*, *11*, 241–247. http://dx.doi.org/10.1111/j.1366-9516.2005.00136.x

Pritchard, J. K., Stephens, M., & Donnelly, P. (2010). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Richardson, D. M., Pysek, P., Rejmánek, M., Barbour, M. G., Panetta, F. D., & West, C. J. (2000). Naturalization and invasion of alien plants: Concepts and definitions. *Diversity and Distributions*, *6*, 93–107.

Ritland, K. (2002). Extensions of models for the estimation of mating systems using n independent loci. *Heredity*, *88*, 221–228.

Rogers, A. R., & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic divergences. *Molecular Biology and Evolution*, *9*, 552–569.

Roubik, D. W., & Aluja, M. (1983). Flight ranges of *Melipona* and *Trigona* in tropical forest. *Journal of the Kansas Entomological Society*, *56*, 217–222.

Schönswetter, P., & Tribsch, A. (2005). Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L. (Apiaceae). *Taxon*, *54*, 725–732.

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2013). Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, *68*, 1–15.

Slatkin, M. (1993). Isolation by distance in equilibrium and nonequilibrium populations. *Evolution*, *47*, 264–279.

Stastny, M., Schaffner, U., & Elle, E. (2005). Do vigour of introduced populations and escape from specialist herbivores contribute to invasiveness? *Journal of Ecology*, *93*, 27–37.

Steinger, T., Haldimann, P., Leiss, K. A., & Müller-Schärer, H. M. (2002). Does natural selection promote population divergence? A comparative analysis of population structure using amplified fragment length polymorphism markers and quantitative traits. *Molecular Ecology*, *11*, 2583–2590.

Sunnucks, P. (2000). Efficient genetic markers for population biology. *Trends in Ecology and Evolution*, *15*, 199–203.

Takebayashi, N., Wolf, D. E., & Delph, L. F. (2006). Effect of variation in herkogamy on outcrossing within a population of *Gilia achilleifolia*. *Heredity*, *96*, 159–165.

Villaseñor, J. L., & Espinoza, G. F. J. (1998). *Catálogo de malezas de México*. Mexico D.F.: UNAM-FCE.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., et al. (1995). AFLP: A new concept for DNA fingerprinting. *Nucleic Acids Research*, *23*, 4407–4414.

Warwick, S., Bain, J. F., Wheatcroft, R., & Thompson, B. K. (1989). Hybridization and introgression in *Carduus nutans* and *C. acanthoides* reexamined. *Systematic Botany*, *14*, 476–494.

Witztuma, A., & Schulgasserb, K. (1995). The mechanics of seed expulsion in Acanthaceae. *Journal of Theoretical Biology*, *176*, 531–542.

Yeh, F. C., & Boyle, T. (1997). *POPGENE, version 1.2 Microsoft Windows-based Software for population genetics analysis*. Alberta, Canada: University of Alberta and Center for International Forestry Research., retrieved from http://www.ualberta.ca/~fyeh/popgene_download.html