Entomologia
Experimentalis et Applicata

DOI: 10.1111/eea.12719

Molecular characterization of *Apis mellifera* colonies from Argentina: genotypic admixture associated with ecoclimatic regions and apicultural activities

Marcelo Nicolás Agra[1]*, Claudia Alejandra Conte[2], Pablo Marcelo Corva[1], Jorge Luis Cladera[2], Silvia Beatriz Lanzavecchia[2] & María Alejandra Palacio[1]

[1]*Unidad Integrada Facultad de Ciencias Agrarias, UNMdP – Estación Experimental INTA Balcarce, Ruta Nacional 226 km 73.5, Balcarce, Buenos Aires 1033, Argentina, and* [2]*Laboratorio de Genética de Insectos de Importancia Económica, Instituto de Genética Ewald A. Favret, CICVyA, INTA-Castelar, Aristizabal y El Ñandú s/n, Hurlingham, Buenos Aires 1033, Argentina*

Accepted: 25 April 2018

Key words: European honeybees, Africanized honeybees, PCR-RFLP, microsatellites, population structure, Hymenoptera, Apidae, invasion route

Abstract

The European honeybee, *Apis mellifera* L. (Hymenoptera: Apidae), is considered as a main contributor to pollination of important crops and to honey production. Originally, beekeeping in Argentina was performed in an extended area covering the north and central region of the country and involving *A. mellifera* of European origin. Later, honeybees of African origin entered South America through Brazil and hybridized with European genetic resources, giving rise to Africanized populations that are characterized by a more aggressive behavior among other unfavorable traits. In this study, a genetic characterization of 396 honeybee colonies from the most important apicultural region of Argentina is presented in order to provide an updated description of population structure and genetic diversity of commercial and feral colonies. Diversity was analyzed using mitochondrial (COI-COII region) and nuclear (eight microsatellites) markers. Three European (M4, C1, C2J) and three African (A1, A4, A30) haplotypes were detected. European haplotypes were mostly found in commercial apiaries, whereas African haplotypes were detected at high frequencies in feral colonies. Microsatellite data were analyzed to estimate population genetic variability at the province level and to evaluate genetic admixture. A high level of hybridization between Africanized and European honeybees was detected with a significant latitudinal cline from north to south. Extensive population admixture resulted in the definition of four clusters that included both feral and commercial colonies and that are explained not only by geographical distribution and degree of Africanization but also by human influence through beekeeping activities.

Introduction

The European honeybee, *Apis mellifera* L. (Hymenoptera: Apidae), is a worldwide distributed species and highly appreciated for honey production. This social insect is essential for crop pollination. One third of the total human diet is dependent on plants pollinated by insects, predominately honeybees (McGregor, 1976). In Argentina, apiculture is recognized as a highly valued resource for the development of small farmers and it has also positioned the country as one of the most important producers and exporters of honey in the world. Consequently, there are several ongoing programs that join official and private efforts to support the activity, including selection and breeding of honeybees.

Between 1830 and 1834, successive introductions of honeybees from various origins to Argentina have been well documented, mainly involving the European subspecies *A. m. mellifera*, *A. m. iberiensis*, and *A. m. carnica*. A fourth, *A. m. ligustica* (or Italian honeybee), was introduced in 1855 and has become the one preferred by most beekeepers (Bierzychudek, 1979). Some of these introduction events did not respond to well defined

*Correspondence: Marcelo Nicolás Agra, Ruta Nacional 226 km 73.5, Balcarce, Buenos Aires 1033, Argentina.
E-mail: agra.marcelo@inta.gob.ar

strategies for the improvement of honey production, but rather to other events such as the occurrence of successive waves of human immigration from various countries. Nevertheless, both planned and occasional introductions of *A. mellifera* subspecies for honey production have contributed to the current existence of a highly variable genetic pool in Argentina.

More recently, an unplanned honeybee introduction (the African *A. m. scutellata*) occurred from Brazil adding complexity to the genetic structure of local honeybee populations (Kerr, 1967). African honeybees were introduced into Brazil in 1956 from South Africa in order to improve honey production in tropical and subtropical areas. In 1957, these African honeybees were accidentally released and proved to be highly successful in their adaptation to the new environment. Since then, hybridization between African and European honeybees (EHB) has been occurring in the Americas, a process that is known as Africanization (Kerr, 1967; Taylor, 1977; Kerr et al., 1982; Fletcher, 1991).

Africanized honeybees (AHB) were first detected in swarms from the North of Argentina in 1965 (De Santis & Cornejo, 1968). Genetic studies performed by Sheppard et al. (1991) produced the first insight about the distribution of AHB in Argentina. The authors described the existence of an AHB-saturated area in the northeastern area of the country. A transition area with both AHB and EHB extended approximately between 28 and 35° S, whereas the region below 35° S was free of AHB. The location of the transition area was refined by Whitfield et al. (2006) to the interval between 31 and 33° S, based on a more precise genomic analysis with nuclear markers, that also included several *A. mellifera* subspecies.

Sheppard et al. (1999) conducted a detailed analysis using PCR-restriction fragment length polymorphism (RFLP) of a region of the mitochondrial DNA (mtDNA) that included samples from Argentina to other reference countries from Europe to Africa. They concluded that the extent of Africanization may have been overestimated, because some of the honeybees carrying haplotypes from lineage A (up to 25% according to their estimations) did not belong to *A. m. scutellata* but rather to other subspecies from northern Africa and the Iberian Peninsula such as *A. m. iberiensis* (Sheppard et al., 1999). Abrahamovich et al. (2007) performed a detailed genetic study of honeybee colonies from 71 localities in Buenos Aires Province. This study confirmed the existence of an ecological limit for AHB at ca. 35° S. Since then, no study of the population structure of honeybees in Argentina has been reported.

Microsatellite markers have been developed in *A. mellifera* and have enabled the analysis of the parental contribution (queen and drones) to honeybee population genetics (Estoup et al., 1995; Solignac et al., 2003). They have been extensively used in the analysis of genetic structure, introgression inferences, and to understand hybridization processes between African and European honeybees (Franck et al., 2001; Solignac et al., 2003; Muñoz et al., 2014a). Hybridization between European and African subspecies is currently considered a complex process that frequently results in asymmetrical levels of introgression in the nuclear and mitochondrial compartments (Clarke et al., 2002; Schneider et al., 2004). This asymmetric introgression is a consequence of different behavior (migration, reproduction) of the males and females from the various subspecies (Oleksa et al., 2013). Whereas, completely Africanized populations may harbor more than 97% African mitotypes, the introgression of African nuclear genes rarely seems to exceed 65–70% (Kraus et al., 2007; De la Rúa et al., 2009; Bouga et al., 2011; Meixner et al., 2013; Muñoz et al., 2014a,b).

Here, we describe the genetic variability of feral and commercial honeybee colonies from 10 provinces that correspond to the apicultural core region of Argentina. We hypothesized that the present population structure of honeybees, particularly the distribution of genes of African origin, has changed since previous studies due to the influence of not only geographical and ecological factors, but also by human intervention through commercial beekeeping activities. Using a PCR-RFLP approach to study the intergenic region of Cytochrome Oxidase I and II mitochondrial genes (COI-COII) and eight microsatellite loci we evaluated maternal and biparental contribution to the genetic diversity of the honeybee populations. The results obtained are analyzed and discussed considering potential source of introduction of honeybees in the country and main factors influencing the complex genetic structure of *A. mellifera* populations. These findings provide valuable information about the current status of Africanization and should be taken into account to preserve selected honeybee stocks for apicuture. In addition, the updated knowledge of the distribution of genetic diversity will be useful to enhance honeybee genetics by breeding programs in support of a well developed and sustainable beekeeping in each ecoclimatic region of Argentina.

Materials and methods

Sampling and DNA extraction

Adult honeybee workers were collected from colonies in the following provinces of Argentina: La Pampa (LP), Buenos Aires (BA), Entre Rios (ER), Santa Fe (SF), Córdoba (CB), Tucumán (TU), Santiago del Estero (SE), Chaco (CH), Corrientes (CO), and Formosa (FS) (Figure 1).

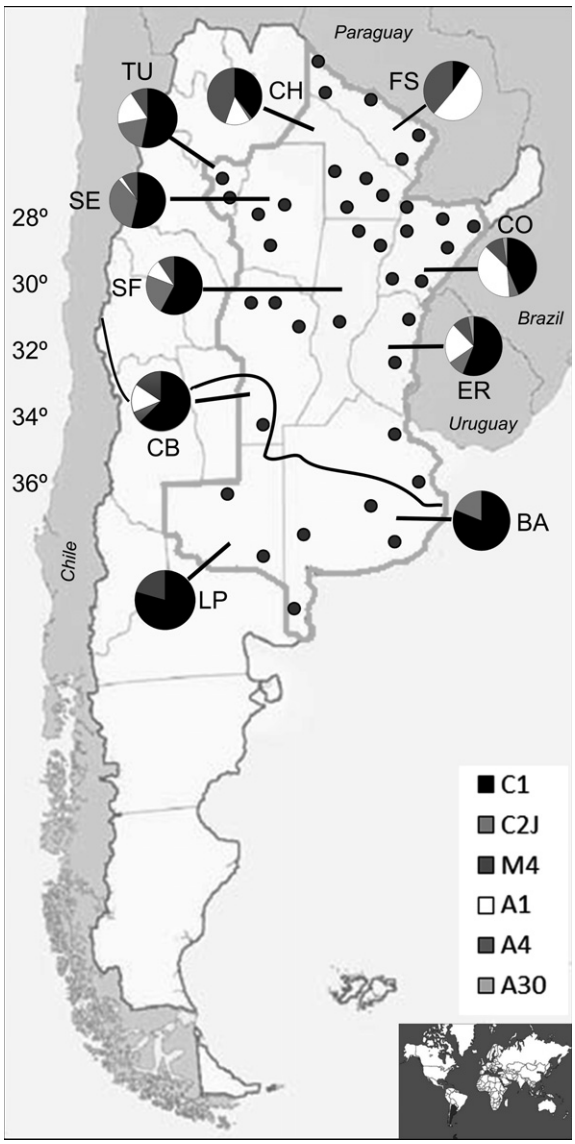

Figure 1 Geographical distribution of mtDNA haplotypes from Argentinian *Apis mellifera* populations. Dots indicate sampling locations (feral and commercial colonies). Pie charts in each province represent the frequencies of the six detected mitochondrial haplotypes (C1, C2J, M4, A1, A4, and A30). See Material and methods section for province abbreviations. The solid black line indicates the 16 °C annual isotherm.

These provinces were selected based on their importance in apicultural activities. The geographic area covered by the sampling includes commercial apiaries that contribute with more than 80% of the honey production of the country. Also, the sampling strategy ensured the inclusion of the area where AHB were present according to previous estimations (Whitfield et al., 2006; Abrahamovich et al., 2007). A representative sampling of each province was performed, resulting in 10–16 apiaries sampled per province. Three colonies were randomly sampled from each apiary. The sampling also included feral colonies that were located far away from known commercial apiaries, with the exception of BA and FS, where only individuals belonging to either commercial or feral colonies were collected, respectively (Table S1). Ten to 30 individuals were taken from the center of the colony and preserved in ethanol 96% (vol/vol). Total DNA was extracted from the thorax of one worker per commercial/feral colony following a high-salt protocol (Baruffi et al., 1995). Six colonies from Riberão Preto (Brazil) were used as positive controls of the AHB.

Mitochondrial DNA analysis

In total 396 DNA samples of honeybee workers were analyzed using a PCR-RFLP-based method. A fragment of either 847 or 1 101 bp from the mitochondrial COI-COII region was amplified by PCR using primers and conditions described by Hall & Smith (1991) and Lobo Segura (2000) with some modifications as it is described below. The 25-µl reaction mix consisted of 1 µM of each primer, 0.5 mM of PCR nucleotide mix (Genbiotech, Buenos Aires, Argentina), 1.5 mM $MgCl_2$ (InbioHighway, Tandil, Argentina), 1× reaction buffer (InbioHighway), 1 U *Taq* Polymerase (InbioHighway), and 5 µl of DNA template. The amplifications were conducted in a MJ PTC-100 thermal cycler (GMI, Ramsey, MN, USA) and the cycle consisted of an initial denaturation step of 2 min at 94 °C followed by 40 cycles of 30 s at 94 °C, 30 s at 55 °C, and 1 min at 72 °C, followed by a final extension step of 2 min at 72 °C. A 10-µl aliquot of each PCR product was digested with *Hinf* I (Promega, Madison, MN, USA) following the manufacturers recommendations. The restriction fragments were separated on 4% (wt/vol) agarose gels, stained with GelRed and photographed under UV light. The haplotypes detected in the restriction analysis using *Hinf* I were compared to the restriction patterns obtained with *Dra* I restriction enzyme (Promega) as it was previously described by Garnery et al. (1993). Haplotypes such as C1 and C2J (see Results) can be distinguished with *Hinf* I, but not *Dra* I. In order to confirm the identity of haplotypes detected in our samples, direct sequencing of PCR products was performed on an ABI3130XL Genetic Analyzer (Applied Biosystems, Carlsbad, CA, USA). The sequences were processed with the Staden Package (Staden, 1996) and aligned with Bioedit (Hall, 1999) and Clustal X 1.81 (Thompson et al., 1997). GenBank accessions EF0326335.1, EF033650.1, EF033654.1, JQ977699.1, JF723978.1, and EF033656.1 that correspond to A1, A4, A30, C1, C2J, and M4 haplotypes, respectively, were used as references (see more details of sequence comparisions in Table S2).

Microsatellite analysis

Eight highly polymorphic microsatellite markers (Ap043, Ap068, Ap085, Ap226, Ap256, Ac011, Ag005a, and A113) developed by Estoup et al. (1995) and Solignac et al. (2003) were screened in 238 DNA samples of high-quality DNA (unexpectedly, 158 samples from a total of 396 exhibited low-quality DNA and it was not possible to amplify the majority of microsatellite loci; these samples were not used in this analysis and therefore the data set was smaller). The forward primers were labeled at the 5′ end with the fluorescent dyes FAM and HEX. PCR reactions were performed in a final volume of 25 μl containing: 2 mM dNTPs, 1.5 mM $MgCl_2$, 1 U *Taq* DNA polymerase (Inbio Highway, Tandil, Argentina), 0.5 μM of each primer, and 40 ng of total DNA template. Amplification of single markers was conducted in a PTC-100 thermal cycler (MJ Research, St Bruno, Canada) with a denaturalization step of 3 min at 94 °C followed by 30 cycles of 15 s at 94 °C, 15 s at 50–60 °C (according to primer annealing temperature; Table S3) and 15 s at 72 °C, with a final elongation step of 10 min at 72 °C. PCR products were analyzed in a MegaBACE 1000 Scientific Capillary DNA Sequencer (Amersham Biosciences, Little Chalfont, UK). Fragment analysis and allele calling was performed with the MegaBACE Genetic Profiler Software Suite v.2.2 (Amersham Biosciences).

Data analysis

Various indicators of genetic diversity were estimated from the microsatellite information. Frequency of null alleles was estimated with Micro-Checker (Van Oosterhout et al., 2004). Expected and observed heterozygosity and number of alleles at each locus were estimated with ARLEQUIN v.3.11 (Excoffier et al., 2005). Deviation from the Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium after Bonferroni corrections were tested using GENEPOP v.3.4 (Rousset, 2008). The genetic differentiation among provinces was calculated using the F_{ST} (Weir & Cockerham, 1984) and the Fisher's method (exact G test) with GENEPOP v.3.4. The distribution of total variation was quantified by the analysis of molecular variance (AMOVA) with ARLEQUIN.

Population structure was analyzed with the Bayesian approach implemented in STRUCTURE v.3.2.1 (Pritchard et al., 2000). Analyzes were based on the admixture model with correlated allele frequencies, with the number of clusters (K) ranging from 2 to 18. Ten runs of 10 000 iterations after a burn-in of 10 000 iterations were performed for each K. The most appropriate K was identified by the ΔK method (Evanno et al., 2005) implemented in the Structure Harvester website (Earl & von Holdt, 2012). Population structure was also inferred by a discriminant analysis of principal components (DAPC) of the microsatellite data using the Adegenet package in the R software environment (Jombart & Collins, 2015). The most probable number of clusters was evaluated by the method of successive K-means (*find.clusters*), covering K from 1 to 10. The best K-value corresponds to the lowest Bayesian information criterion (BIC).

Regression analyzes were conducted to explain the distribution of AHB in the area under study in relation to bioclimatic variables. Mitotypes were coded as either African (1) or European (0) and a logistic regression model was fitted, including alternatively one of the following explanatory variables: latitude, maximum daily temperature in January (summer) and in July (winter), minimum daily temperature of January and July, and number of hours with temperature below 4 °C in July. The climate information was obtained from the agrometeorological stations closest to the sampling sites using SIGA v.1.0.5 (SIGA, 2018) developed by the Institute of Agricultural Technology of Argentina (INTA, Buenos Aires, Argentina). In all cases, the origin of the sample (commercial or feral colonies) was also included in the models as a fixed effect.

The population assignment probabilities generated with STRUCTURE were analyzed by linear regression, using the same explanatory variables described above. To define the dependent variable, we assumed that feral colonies with haplotype A4 from the North of the country had a comparatively larger contribution of African genetic material. Therefore, for each sample we chose the assignment probability corresponding to the population that comparatively made the largest contribution to the referred group and at the same time made the lowest contribution to samples from Buenos Aires Province, for K = 4 (see Results). Tests of parallelism and coincidence were included in the linear regression analyzes to evaluate the contribution of sample origin to the distribution of AHB. All the regression analyzes were conducted with the SAS University Edition software package (SAS Institute, Cary, NC, USA).

Results

Mitochondrial haplotype analysis

After the analysis of 396 honeybee colonies distributed in the main apicultural region of Argentina, six mitochondrial haplotypes were detected. Three haplotypes corresponded to evolutionary branch A (A1, A4, and A30), one corresponded to evolutionary branch M (M4, found only in feral colonies from LP), and two haplotypes corresponded to branch C (C1 and C2J) (Figure 1, Table 1). African haplotypes (A1 and A4) were detected both in feral and commercial colonies from the north and center of the country. Haplotype A1 tended to be more abundant than

Table 1 Mitochondrial haplotypes (C1, C2J, M4, A1, A4, and A30) detected in feral (F) and commercial (C) honeybee colonies from each sampled province (see Figure 1 for explanation of the codes)

		C1		C2J		M4		A1		A4		A30	No. mt
Province	n	F	C	F	C	F	C	F	C	F	C	C	haplotypes
BA	63	–	51	–	12	–	–	–	–	–	–	–	2
LP	39	–	31	–	0	8	–	–	–	–	–	–	2
CB	32	2	18	–	2	–	–	5	–	3	2	–	4
SF	31	4	14	–	7	–	–	–	3	3	–	–	4
ER	32	6	12	–	3	–	–	2	5	–	3	1	5
CO	39	7	10	–	2	–	–	11	4	1	3	1	5
CH	54	6	15	–	1	–	–	2	6	6	18	–	4
FS	31	3	–	–	0	–	–	16	–	12	–	–	3
SE	43	1	22	2	13	–	–	1	–	4	–	–	4
TU	32	–	15	–	6	–	–	3	3	2	3	–	4
Total	396	29	188	2	46	8	–	40	21	29	29	2	
Total (F + C)			217		48		8		61		58	2	

n = number of colonies sampled per province.

A4 in most sampling locations where African haplotypes were found, with the exception of CH (commercial and feral colonies) and feral colonies from TU and SE, in which A4 was at a higher frequency. All the six samples from Riberão Preto (Brazil) used as a control of our assays presented the A4 haplotype. Another African haplotype, A30, was found in single colonies from ER and CO. No African haplotypes were found in BA and LP, the southern provinces in the sampling area (Figure 1).

European haplotypes were detected mainly in commercial colonies; haplotype C1 was found in almost 80% of the samples and the rest corresponded to haplotype C2J. Another European haplotype (M4) was found only in feral colonies of LP, close to the southern limit of the sampling area (Figure 1, Table 1). Haplotype diversity among provinces showed that ER and CO (two provinces located to the East of the territory) have the highest values (with five haplotypes identified: C1, C2J, A1, A4, and A30), whereas BA showed the lowest value of diversity with only two haplotypes detected (Table 1).

If only commercial colonies are compared across provinces, significant differences in the proportion of African and European haplotypes were detected in TU, SE, CB, and SF; in all these cases European haplotypes were significantly more frequent (Table 1). In addition, it is worth noticing the absence of African haplotypes (A1, A4, and A30) in BA and LP, and the presence at low frequency (less than 0.05) of these haplotypes in CB, SF, ER, SE, and TU. Conversely, CO and CH showed a high percentage of African haplotypes (50 and 59%, respectively) both in feral and commercial colonies. FS represent an extreme of the presence of African haplotypes, with 90% of colonies showing these haplotypes (Table 1).

Microsatellite analysis

The number of alleles per locus detected in the analysis of microsatellite markers ranged from 5 (Agoo5a) to 16 (Ap043). The highest number of alleles and expected heterozygosity values were observed in SE, CH, and CO, whereas the lowest values were found in BA and LP with ER (Table 2). Linkage disequilibrium was not detected in any of the provinces/loci comparisons performed. MicroChecker analysis suggested that the presence of null alleles might explain some of the detected heterozygote deficiencies, and the following provinces (loci) were more likely in Hardy-Weinberg equilibrium: BA (Ap226), FS (Ap068, Ap256, A113), CB (Ap256, A113), CO (Ap256), CH (Ap043, Ap256, Ag005a), SF (Ap043), ER (Ac011, Ap043, Ap256), and TU (A113) (Table 2).

To get insight in the population structure of honeybees in Argentina, the first analyzes were conducted using the province as a reference for the grouping of samples. An AMOVA showed that most of the genetic variability was found within provinces (95.4%) with only 3.6% of variation among provinces (F_{ST} = 0.03573, P<0.01), indicating that this grouping strategy provided little information to justify genetic variation. Consistently, most F_{ST} values estimated among pairs of provinces showed low values (range 0.014–0.129) being higher F_{ST} values between BA and FS, TU, CO, and CH (Table 3).

The Evanno's ΔK method (Evanno et al., 2005) suggested that the genetic pool of honeybees from Argentina would most likely be the result of the admixture of two genetic origins (k = 2) (Figure 2). However, the comparison of results for K = 3 and K = 4 provided interesting information about population substructure. These analyzes displayed a group of samples with a homogeneous

Table 2 Microsatellite analysis of Argentinian honeybee populations (see Figure 1 for explanation of the codes): number of alleles per locus (N_a), observed heterozygosity (H_O), expected heterozygosity (H_E), and estimated null allele frequency (Null)

		ER		TU		FS		SE		CB		CO		CH		SF		BA		LP	
Locus	Allelic range	N_a	H_O H_E Null	N_a	H_O H_E Null	N_a	H_O H_E Null	N_a	H_O H_E Null	N_a	H_O H_E Null	N_a	H_O H_E Null	N_a	H_O H_E Null	N_a	H_O H_E Null	N_a	H_O H_E Null	N_a	H_O H_E Null
Ap043	117–171	16	0.71429 0.86542* 0.0885#	12	0.92857 0.90476 −0.0312	9	0.81818 0.85518 0.0106	9	0.83333 0.8587 −0.0063	13	0.76923 0.88084 0.058	11	0.84 0.85633 −0.0005	12	0.74286 0.88903* 0.0824#	12	0.71429 0.91534* 0.1054#	8	0.70833 0.6844 −0.0277	10	0.76471 0.90553 0.0695
Ap068	130–180	9	0.79412 0.86435 0.035	9	1.00 0.88923 −0.0781	8	0.59091 0.81078* 0.1456#	8	0.73333 0.77701 0.012	10	1.000 0.87273 −0.0769	10	0.69565 0.85314 0.0908	10	0.73529 0.83845 0.0582	8	0.66667 0.83448 0.095	9	0.73333 0.76158 0.0105	8	0.83333 0.85079 −0.0037
Ap085	176–214	15	0.88571 0.88199 −0.0093	9	0.85714 0.8836 −0.003	12	0.86364 0.87949 −0.0024	11	0.87714 0.86772 −0.012	11	0.96296 0.89518 −0.0458	13	0.88 0.88 −0.0101	15	0.91429 0.91014 −0.0095	7	0.93333 0.82529 −0.0783	9	0.9 0.7791 −0.0804	10	0.88235 0.81996 −0.0515
Ap226	227–259	9	0.81818 0.80746 −0.0142	7	0.92857 0.83598 −0.0706	9	0.84211 0.84637 −0.0108	10	1.000 0.88736 −0.0766	9	1.000 0.84067 −0.0958	10	0.86957 0.85507 −0.0194	13	0.91176 0.89113 −0.0188	8	0.66667 0.74713 0.04	9	0.5 0.76299* 0.1996#	11	0.76471 0.90553 0.0695
Ap256	150–190	12	0.71429 0.85921* 0.085#	11	0.85714 0.91005 0.0118	12	0.68182 0.88584* 0.1188#	13	0.8 0.91494 0.0501	12	0.59259 0.89797* 0.1959#	15	0.625 0.91312* 0.1771#	15	0.72222 0.89163* 0.098#	10	0.93333 0.88506 −0.0435	13	0.8 0.80565 −0.0049	10	0.83333 0.87143 0.0083
Ag005a	81–108	9	0.76471 0.85075 0.0459	7	0.57143 0.80688 0.1531	10	0.72727 0.80021 0.0363	9	0.86667 0.87356 −0.013	8	0.76923 0.76923 −0.0097	8	0.64 0.80163 0.1021	11	0.72222 0.88537* 0.0946#	5	0.8 0.73103 −0.0619	9	0.66667 0.6791 0.0008	11	0.76471 0.8984 0.0655
Ac011	83–145	13	0.68571 0.8617* 0.1066#	9	0.85714 0.84921 −0.0228	12	0.68182 0.82875 0.0859	10	0.73333 0.86897 0.0678	10	0.92857 0.85844 −0.0482	12	0.875 0.9078 0.0079	12	0.8 0.84969 0.0229	8	0.9 0.84211 −0.0588	8	0.68966 0.71385 0.0085	9	0.77778 0.87778 0.0464
A113	200–238	11	0.68571 0.81035 0.0762	8	0.75 0.8587 0.0464	8	0.5 0.79704* 0.2181#	9	0.93333 0.85057 −0.0633	10	0.55556 0.73445* 0.1295#	11	0.91667 0.87234 −0.0353	12	0.86111 0.87402 0.0004	8	0.6 0.8023 0.1276	11	0.82143 0.75779 −0.0493	10	0.55556 0.88095* 0.2131#

*Significantly different from expectation under Hardy–Weinberg equilibrium (after Bonferroni correction).

#Significant frequencies of null alleles (Fisher's combined probability test: $P<0.05$).

Table 3 Pairwise F_{ST} values among Argentinian honeybee populations (see Figure 1 for explanation of the codes). Nuclear differentiation based on eight microsatellites

	ER	TU	FS	SE	CB	CO	CH	SF	BA
TU	0.0302*								
FS	0.0369*	0.0372*							
SE	0.0223*	0.0331*	0.0333*						
CB	0.0048	0.0396*	0.0456*	0.0329*					
CO	0.0161*	0.0215*	0.0162*	0.0109	0.0226*				
CH	0.0231*	0.0251*	0.0358*	0.0112*	0.0194*	0.0057			
SF	0.0218*	0.0453*	0.0640*	0.0002	0.0253*	0.0311*	0.0264*		
BA	0.0469*	0.1111*	0.1296*	0.0666*	0.0493*	0.0847*	0.0704*	0.0354*	
LP	0.0174*	0.0467*	0.0505*	0.0200*	0.0167*	0.0277*	0.0144*	0.0160	0.0512*

*Significant differentiation between paired comparison ($P<0.05$).

genetic background across increasing K values. All those samples had C1 haplotype and corresponded to BA; therefore, we inferred that they were EHB. With K = 3, there was a genetic pool different from the one seen in honeybees from BA, that made a proportionally higher contribution to feral colonies, especially those with haplotype A4, and this result led us to conclude that it corresponded to an African origin. In addition, a high level of admixture was detected in most commercial samples with haplotype C1, others than those from BA. In turn, in the analysis with K = 4 it can be seen that honeybees with haplotype M4 from LP, where no AHB have been detected so far, have a genetic makeup that differs from other EHB (those from BA, for example). If these assumptions were correct, they would indicate a complex genetic structure and widespread genetic admixture, with a significant contribution of African genetic material to honeybees of all provinces other than BA and LP, and especially to those from feral colonies. In turn, the European germ plasm would be more heterogeneous than it would be expected based on the comparison of mitochondrial haplotypes alone, given the prevalence of haplotype C1.

The results obtained in the DAPC, performed with the Adegenet package, were consistent with those generated with STRUCTURE software and confirmed a complex population structure. For the DAPC, the first 80 PCs of PCA were used in the discriminant analysis and two discriminant functions were saved. A DAPC reports an estimated number of clusters in order to attain the largest between-group variance with the smallest within-group variance. In the present case, there was little variation among the results of the BIC from the DAPC for K = 3 and K = 4. Because the lowest BIC value corresponded to K = 4 (Figure S1), this number of clusters was used to complete the discriminant analyzes (Figure 3). The contribution of the various provinces to the four inferred clusters is presented in Figure S2.

The first discriminant function explained most of the genetic variability and was associated with the degree of Africanization; the eigenvalues for the first and second functions were 39.2 and 23.9, respectively. Four clusters, named 1–4, were identified with 60, 54, 57, and 67 samples, respectively (Table 4). All clusters were very heterogeneous in terms of sample origin. Particularly, in clusters 2 and 3 there were samples from the 10 provinces. Cluster 1 included samples from all provinces with the exception of BA, whereas cluster 4 did not include any samples from FS or TU and most individuals corresponded to colonies from BA. Most samples corresponding to feral colonies from FS were included either in clusters 1 or 3 but none of them was allocated to cluster 4. In cluster 4, most samples corresponded to colonies from BA. From five samples of LP feral colonies that had haplotype M4, four were placed in cluster 2 and one in cluster 3. In addition, in term of diversity parameters, high and similar values of genetic variability were estimated for each cluster (Table 5).

Interestingly, the distribution of clusters with respect to the first discriminant function was highly concordant with the percentage of feral colonies in each of them and also with the percentage of African haplotypes, particularly haplotype A4. Very few African haplotypes (feral colonies and commercial apiaries) were found among samples from cluster 4 (only 6%). In turn, 40% of the samples in cluster 1 were from feral colonies, being 36.7% African haplotypes with higher percentage of A4 (59.1%) (Table 4).

Honeybee distribution associated to bioclimatic variables

In all the logistic regression analyzes that linked mitotype probabilities (African or European) and several bioclimatic variables, we observed a significant effect ($P<0.01$) of the origin of the sample (commercial or feral colonies; Figures 4A and S3). The analysis including the latitude had the highest goodness of fit (Figure 4A). This result is

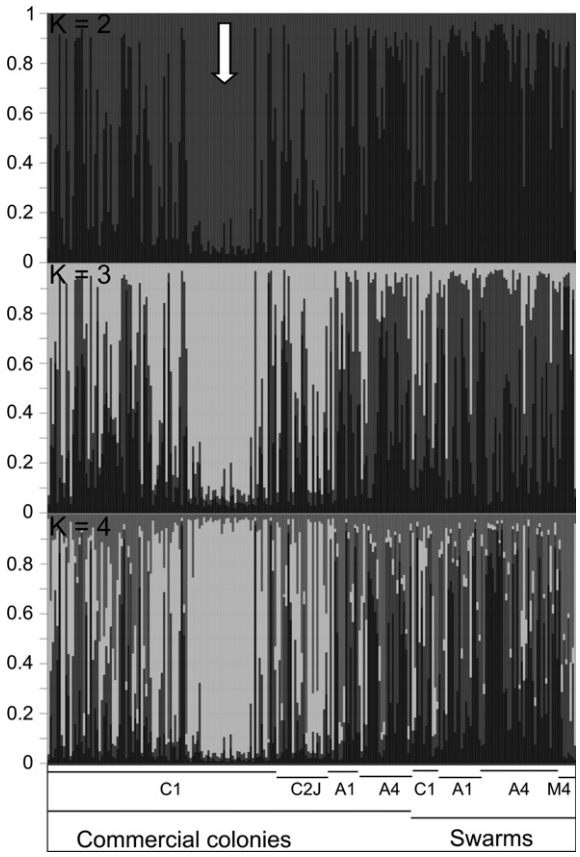

Figure 2 Population structure analysis. Individual honeybees were grouped by origin – commercial colonies and swarms (feral colonies) – and their corresponding mitochondrial haplotype (C1, C2J, A1, A4, and M4). The genetic ancestry of each individual is represented by a vertical bar partitioned into different shades of grey segments with the length of each segment representing the proportion of the individual's genome from K = 2–4 populations, that represent the admixture genetic ancestry in the corresponding genetic groups. The arrow indicates a group of samples from BA (Buenos Aires) that remained highly homogeneous through the analyses with different K values.

not surprising given that there is not a very significant variation in altitude or other geographical features among sampling locations and latitude is highly correlated with various climate variables. A significant latitudinal cline was detected and the probability of finding an African haplotype in a sample decreased from North to South. Among the climate variables, the model including the maximum daily temperature of January had the lowest goodness of fit, suggesting that temperature in summer is not a limiting factor for AHB.

The results of the linear regression analyzes were consistent with those based on haplotype information (Figures 4B and S3). All variables had a significant effect ($P<0.01$) on the putative degree of Africanization estimated in the STRUCTURE analysis, with the exception of maximum daily temperature of January ($P>0.09$). Also in this case, the analysis with latitude provided the best goodness of fit (Figure 4B). The coincidence test was significant for all the variables, whereas the parallelism test was not significant in any case. Taken together, these results imply that under similar conditions, the degree of Africanization estimated with nuclear markers was always higher in feral than in commercial colonies (Figures 4B and S3).

Discussion

We provide an updated description of the genetic variability found in *A. mellifera* colonies from the most important apicultural area of Argentina. The molecular markers have given valuable information to comprehend the complex genetic structure present in commercial and feral honeybee colonies. The results obtained highlight the importance of management strategies implemented by beekeepers in Argentina, consistent with the use of selected honeybee stocks for the production of high-quality honey and the existence of unmanaged genetic resources, which must be taken into account by both, honeybee breeding programs and queen breeders.

Mitochondrial haplotypes have been extensively used to understand the genetic variability and route of migration of *A. mellifera* subspecies through continents (Cornuet et al., 1991; Garnery et al., 1992; Franck et al., 2000). Four lineages of honeybees (A, O, C, and M) have been previously reported that had spreaded worldwide from their center of origin (Ruttner et al., 1978; Garnery et al., 1998a; Franck et al., 2001). African honeybees are mainly represented by A and O haplotypes, Western European honeybees by M haplotypes, and Eastern (Central) European honeybees by C (Franck et al., 2001; De la Rúa et al., 2004). The introduction of honeybees to the New World was mostly associated with human migration and beekeeping activities. In the USA, a high frequency of C haplotypes was detected in commercial apiaries (beekeepers) and also in unmanaged colonies (swarms and feral colonies) but haplotype C1 in particular seems to be highly frequent in commercial colonies only (Magnus et al., 2011, 2014). It is worth noting that DNA sequencing, as used by Magnus et al. (2011, 2014) allows for a better resolution of mitochondrial haplotypes than RFLP analysis. Previous reports (Sheppard et al., 1999; Abrahamovich et al., 2007) found a similar distribution of mtDNA haplotypes in Argentinian colonies, indicating that European genetic resources were extensively established in temperate regions of the country, whereas Africanized ones are found

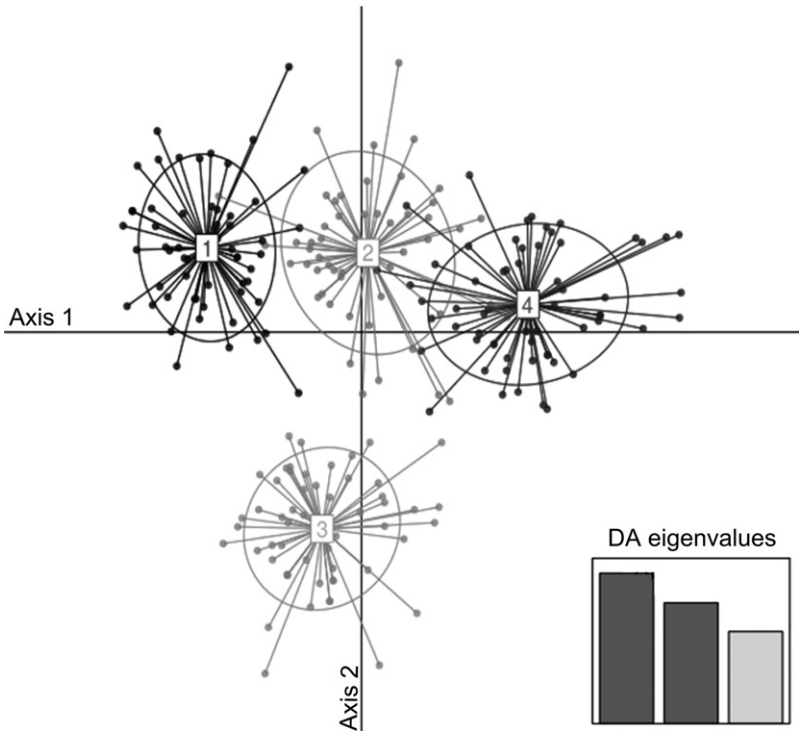

Figure 3 Scatterplot of honeybee individuals on the two principal components of the discriminant analysis (DA) of principal components based on the Bayesian information criterion. The dots represent the individuals and the inertia elipses the four groups in different shades of grey (see Table 4 for clusters identification). Eigenvalues of the analysis are displayed in the lower righthand corner. The two dark gray bars represent the variance explained by principal components (39.2 and 23.9%).

Table 4 Percentage of feral colonies (FC), African haplotypes (A1 and A4), and Africanized honeybees (AHB, from commercial apiaries or FC) in the clusters defined in the discriminant analysis of principal components (see Figure 3)

			African haplotypes		AHB	
Cluster[1]	n	FC	A1	A4	Apiaries	FC
3	57	52.6	12.3	28.0	19.3	40.3
1	60	40.0	15.0	21.7	23.3	35.0
2	54	27.7	5.5	9.2	20.4	14.8
4	67	4.5	0.0	1.5	4.5	1.5

[1]Ordered in decreasing % FC.
n = number of colonies grouped per cluster.

particularly in subtropical areas. Abrahamovich et al. (2007) reported a high frequency of the C1 haplotype in Buenos Aires province, which is characteristic of the *A. m. ligustica* subspecies. The broad distribution of this subspecies in this province described by Abrahamovich et al. (2007) and in other temperate regions of the country (Sheppard et al., 1999) is in line with the results obtained in the present paper, after the analysis of honeybee colonies belonging to the main apicultural region of the country. We describe here the presence of European haplotypes in commercial colonies with a high prevalence of C1 and, in lower frequency, C2J. This last haplotype is a variant of C2, characteristic of the *A. m. carnica* subspecies, original to the North Mediterranean C group (Ruttner et al., 1978; Nedić et al., 2009).

African-derived honeybee populations also have a great representation in the American continent. Africanized honeybees have been previously described in Mexico (Clarke et al., 2001; Kraus et al., 2007), Puerto Rico (Rivera-Marchand et al., 2012; Galindo-Cardona et al., 2013), Brazil (Collet et al., 2006), Perú (Quezada-Euán et al., 2003), and Uruguay (Branchiccela et al., 2014) in commercial and feral colonies of *A. mellifera*. Here, we describe the presence of three mitochondrial haplotypes from the A lineage. Within this group, the most frequent haplotypes were A4 and A1. A4 is characteristic of South African honeybees that belong to the subspecies *A. m. scutellata* (Schneider et al., 2004; Collet et al., 2006), and as such, it is typically associated with the Africanization process that started in Brazil in 1957. There

Table 5 Genetic diversity parameters (mean ± SE) for the four clusters defined in the discriminant analysis of principal components (see Figure 3): number of alleles (N_a), number of private alleles (pN_a), observed heterozygosity (H_O), expected heterozygosity (H_E), and unbiased expected heterozygosity (uH_E)

Cluster	n	N_a	pN_a	H_O	H_E	uH_E
1	60	15.10 ± 1.43	1.00 ± 0.42	0.70 ± 0.10	0.86 ± 0.02	0.86 ± 0.02
2	54	14.87 ± 1.66	1.37 ± 0.46	0.71 ± 0.10	0.81 ± 0.06	0.82 ± 0.06
3	57	15.50 ± 1.28	2.00 ± 0.60	0.71 ± 0.10	0.85 ± 0.01	0.86 ± 0.01
4	67	13.87 ± 1.42	1.12 ± 0.23	0.65 ± 0.10	0.80 ± 0.03	0.79 ± 0.03

n = number of colonies grouped per cluster.

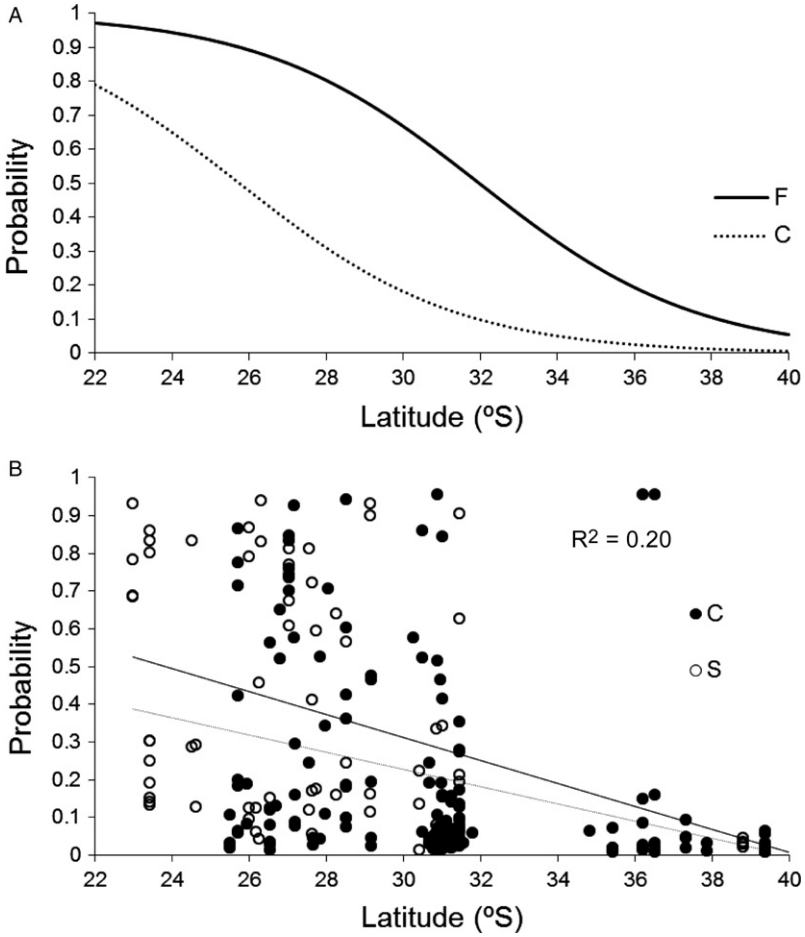

Figure 4 (A) Probability of detection of African honeybee haplotypes depending on the latitude, estimated by logistic regression (goodness of fit: 0.79). (B) Linear regression of the degree of Africanization on latitude, based on the analysis of nuclear markers with STRUCTURE (R^2 = 0.20). Assignment probabilities correspond to the population that made the largest genetic contribution to feral colonies of Formosa province, which was assumed to be from African origin (see Figure 1). Honeybees originated from feral (solid line, white dots) and commercial (dotted line, black dots) colonies.

are no other documented introductions of African honeybees to this country since that event, and therefore it is the most likely source of introduction of this haplotype in Argentina.

The information about possible routes of introduction of the A1 haplotype in the country provided by previous studies is inconclusive. It could either correspond to *A. m. intermissa* from the North of Africa trough Spain

(Cánovas et al., 2007) or to *A. m. iberiensis* from the South of the Iberian Peninsula (De la Rúa et al., 2004; Whitfield et al., 2006). In fact, Sheppard et al. (1999) stated that the degree of Africanization in Argentinian honeybee colonies could have been overestimated, given the occurrence of other sources of introduction of African haplotypes in the country. In Mexico and the south of the USA, A1 was reported in higher frequency than A4 (Kraus et al., 2007). In turn, in Uruguay and Brazil, A4 and A1 have been detected with a relative abundance of 68 and 26%, respectively, but A1 showed a cline from north to south, similar to the original distribution pattern described in Africa (Collet et al., 2006). Therefore, at the same latitude, A1 had a much lower frequency in Uruguay and the south of Brazil than in the region considered here. This result supports the hypothesis proposed by Sheppard et al. (1999) of other possible sources of African mitotypes in Argentina, such as Spanish immigrants who could have introduced honeybees with A1 to the country.

The presence of the A30 haplotype only in colonies from CO and ER is possibly associated with introductions of honeybees from the Iberian Peninsula where this haplotype has been previously detected (Garnery et al., 1995; Franck et al., 2001; Collet et al., 2006). In turn, the presence of haplotype M4 in feral colonies from a single location in LP is probably justified by documented introductions of *A. m. mellifera* from Germany between the end of the 19th century and the beginning of the 20th century (Bierzychudek, 1979).

The distribution of mtDNA haplotypes in Argentina reported here allows for the evaluation of the current 'Africanization' status of the country. The Africanization process in the Americas has been described as 'one of the most spectacular biological invasions yet documented' (Schneider et al., 2004; Pinto et al., 2005). In agreement with that statement, the first local evaluations of Africanization using molecular tools, reported the presence of a saturated area with AHB above 28° S latitude (Sheppard et al., 1991). However, our results demonstrated the presence of EHB haplotypes in northern provinces (above this latitude). Although we show that there are no saturated areas as it was originally proposed by Sheppard et al. (1991), there is still a strong presence of AHB in the north of the country, particularly in feral colonies.

The Africanization process has been well studied between genetic and behavioral points of view. AHB populations have demonstrated some advantageous behavioral characteristics that allowed the successful colonization of new habitats. As an example, Africanized drones outperform their European counterparts at the moment of queen mating, in addition to other physiological traits such as the faster development displayed by AHB queens, that increases their chances to lead a colony (Pinto et al., 2004; Schneider et al., 2004; Harrison et al., 2006; Rivera-Marchand et al., 2012) making a significant contribution to population admixture in commercial colonies. In spite of the displacement of EHB by AHB described in the New World, our results are consistent with the idea that beekeeping in Argentina is mainly performed using EHB and beekeepers have maintained the use of this selected honeybee genetics by the implementation of appropriated management strategies (queen breeders, private cooperatives, honeybee breeding programs) (Palacio et al., 2000, 2010; Bedascarrasbure, 2011).

The development of nuclear markers greatly improved the study of the genetic complexity found in hybrid honeybee populations, not only due to their high degree of polymorphism but also to the possibility of monitoring the genetic contribution of both parents. Extensive studies have been performed applying this kind of markers to analyze the structure of honeybee populations from their center of origin and derived hybrids (Garnery et al., 1998b; Franck et al., 2000, 2001; Clarke et al., 2002; Kraus et al., 2007). In our study, mtDNA markers revealed the existence of an important reservoir of African germ plasm in feral colonies. The analysis of microsatellite markers allowed a deeper knowledge of the hybridization process that occurred in Argentina. Introgression of African genotypes was evident in commercial honeybee stocks and feral colonies, showing as mixed patterns in the STRUCTURE analysis. This complex population structure was also evident by DAPC where a mixed contribution of individuals from different honeybee populations is present in the four clusters identified. Taken together, the results from STRUCTURE and Adegenet would indicate widespread genetic admixture, with a significant contribution of African genetic material to honeybees of all provinces except for BA and LP, and especially to those from feral colonies. In turn, the European germ plasm appears more heterogeneous than expected based on the comparison of mitochondrial haplotypes alone, given the prevalence of haplotype C1.

Climatic conditions may affect the distribution of honeybees. Carneiro Vital et al. (2012) modeled the potential distribution of AHB in America based on nine climate-related environmental variables that defined the original distribution of seven subspecies of *A. mellifera* in Africa. They concluded that the current distribution of AHB in America reflects niche-specific characteristics of the subspecies from which they originated, particularly in the case of *A. m. scutellata*. Interestingly, their models predict that a suitable area that includes most of the Buenos Aires province is not currently occupied by AHB. Based on results of Kerr et al. (1982), Taylor & Spivak (1984)

proposed a limit for permanent colonization of AHB at 32° S. A further limit of temporary colonization at ca. 35° S was defined by the ability of AHB to survive in the growing season but not to overwinter. Carneiro Vital et al. (2012) hypothesized that the suitable but unoccupied area defined by their models could overlap with the area of temporary colonization of Taylor & Spivak (1984), but the existence of limiting factors for colonization others than climate itself cannot be ruled out. In the present study, no African mitotypes were found beyond 32° S. It must be taken into account that any ecological limits to AHB would be confounded by the effect of selection against this genetic group in a region that concentrates more than 50% of beekeeping in the country. Combined effects of unfavorable climate conditions and increasingly professionalized beekeeping practises have contributed to the current preponderance of EHB in this temperate area. The results of the regression analyzes partially explain the present distribution of honeybees, as the degree of Africanization and genetic variability are associated to bioclimatic variables, mainly latitudinal clines, as was previously reported by others (Taylor & Spivak, 1984; Sheppard et al., 1991; Whitfield et al., 2006).

Our results are consistent with those of Abrahamovich et al. (2007) who analyzed the mitotypes of 300 colonies in the Buenos Aires province; only 3.35% of them had haplotypes A1 or A4 and they were all restricted to the area above 35° S. Southwick et al. (1990) predicted that AHB would be limited by the isocline of 120 consecutive days not exceeding 10 °C in winter. In the Buenos Aires province the latitude of 35° S approximately agrees with the 16 °C annual mean isotherm and the 9 °C mean isotherm for July (winter in the Southern Hemisphere; see Figure 1), which could potentially set a limit for the spread of AHB. Interestingly, Kono & Kohn (2015) established a putative limit at 35.5° N for AHB at the West Coast of USA (California) corresponding to the 10 °C mean isotherm in winter. Taking into account the current climate changes and the global warming process, further studies should include higher number of apiaries and feral colonies established in temperate regions in order to monitor the distribution of AHB in Argentina and its potential influence on apiculture.

Different factors other than the Africanization process can justify the present complex genetic structure of honeybees in Argentina. Gene flow among honeybee populations is facilitated by the absence of significant natural barriers, not only within the country but also with bordering countries. ER and CO were the provinces with the highest haplotype diversity. It is worth noting that CO has common borders with Brazil and Paraguay, whereas ER shares border with Uruguay, where close urban settlements and rural beekeepers could exchange genetic material, an unregistered process that potentially contributes to the detected variability of honeybee populations in Argentina.

Another important factor contributing to gene flow in honeybees from Argentina is transhumance. Every year, beekeepers from temperate regions (mainly from BA) move colonies from south (temperate climate) to north (subtropical climate) of the territory searching for better conditions for colony multiplication and to have advantages at initial stages of the honey production season. Additionally, transhumance is also a habitual activity related to the needs of crop pollination.

In summary, our results indicate that despite the successive planned and unplanned introductions of the most varied genetic materials, there is currently in commercial colonies a high frequency of the Italian honeybee *A. m. ligustica*, a subspecies highly appreciated by beekeepers and queen breeders and widely recognized by its docility, productivity, and prolificacy. Nevertheless, there is a strong presence of AHB in the north of the country, mainly in feral colonies, being a permanent source of Africanized germ plasm. We have no evidence that the southern limit of the area where AHB are present has changed since the first estimations, ca. 37 years ago.

The information provided here can be of help for strengthening the strategy of honeybee breeding programs and queen breeders in order to maintain the high-quality of honeybee genetic resources adapted to each ecoclimatic region of the country for a sustainable beekeeping. Moreover, the evaluation of the genetic structure of Argentinian honeybees should be kept updated in order to be available to honeybee management programs.

Acknowledgements

The authors thank private beekeepers (especially Lisandro Balbuena, Raúl Daniele, Juan José Mange, and Miguel Pelliza), queen breeders of MVC Association (Antonio Fabbro, Daniel Avena, Carlos Villavicencio, and Lorena Casanova), and INTA and governmental institutions personnel (Andrea Aignasse, María Belén Camacho, Emilio Figini, Analía Martinez, Pablo Chipulina, Miguel Busatto, Clemencia Barberena, and Natalia Messina) who provided the samples for this work. The support of Dr. S. Feingold and the technical assistance of the staff of the Agrobiotechnology Laboratory at INTA-Balcarce are greatly appreciated. We also thank Martín Castellote for his assistance in population structure analyses. We thank the two anonymous reviewers for their constructive comments, which helped us to improve the manuscript. This work was financially supported by INTA through projects PNAPI

123022 2009-2012 and PNAPI 1112042 2013-2019, and by the Universidad Nacional de Mar del Plata (Argentina) through grant AGR 395-12 to A.P. This work is part of a PhD dissertation submitted by M.N. Agra towards fulfilling the requirement of the Doctoral degree in Agricultural Science, for the Postgraduate Program of Balcarce Integrated Unit (INTA-UNMdP).

References

Abrahamovich AH, Atela OE, De La Rúa P & Galián J (2007) Assessment of the mitochondrial origin of honeybees from Argentina. Journal of Apicultural Research 6: 191–194.

Baruffi L, Damián G, Guglielmino C, Malacrida A & Gasperi G (1995) Polymorphism within and between populations of *Ceratitis capitata*. Comparison between RAPD and multilocus enzyme electrophoresis data. Heredity 74: 425–437.

Bedascarrasbure E (2011) Consolidando la Apicultura como herramienta de desarrollo. Ediciones INTA, pp. 25–39.

Bierzychudek A (1979) Historia de la Apicultura Argentina. H. J. Mattone, Buenos Aires, Argentina.

Bouga M, Alaux C, Bienkowska M, Büchler R, Carreck NL et al. (2011) A review of methods for discrimination of honeybee populations as applied to European beekeeping. Journal of Apicultural Research 50: 51–84.

Branchiccela B, Aguirre C, Parra G, Estay P, Zunino P et al. (2014) Genetic changes in *Apis mellifera* after 40 years of Africanization. Apidologie 45: 752–756.

Cánovas F, De La Rúa P, Serrano J & Galián J (2007) Geographical patterns of mitochondrial DNA variation in *Apis mellifera iberiensis* (Hymenoptera: Apidae). Journal of Zoological Systematics and Evolutionary Research 46: 24–30.

Carneiro Vital V, Hepburn R, Radloff S & Fuchs S (2012) Geographic distribution of Africanized honeybees (*Apis mellifera*) reflects niche characteristics of ancestral African subspecies. Natureza & Conservação 10: 184–190.

Clarke KE, Oldroyd B, Javier J, Quezada-Euán G & Rinderer T (2001) Origin of honeybees (*Apis mellifera* L.) from the Yucatan peninsula inferred from mitochondrial DNA analysis. Molecular Ecology 10: 1347–1355.

Clarke KE, Rinderer TE, Franck P, Quezada-Euán JG & Oldroyd BP (2002) The Africanization of Honeybee (*Apis mellifera* L.) of the Yucatan: a study of a massive hybridization event across time. Evolution 56: 1462–1474.

Collet T, Ferreira KM, Arias MC, Soares AEE & Del Lama MA (2006) Genetic structure of Africanized honeybee populations (*Apis mellifera* L.) from Brazil and Uruguay viewed through mitochondrial DNA COI–COII patterns. Heredity 97: 329–335.

Cornuet JM, Garnery L & Solignac M (1991) Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA. Genetics 128: 393–403.

De la Rúa PR, Hernández-García R, Pedersen BV, Galián J & Serrano J (2004) Molecular diversity of Honeybee *Apis mellifera iberica* L. (Hymenoptera: Apidae) from western Andalusia. Archivos de Zootecnia 53: 195–203.

De la Rúa P, Jaffé R, Dall'Olio R, Muñoz I & Serrano J (2009) Biodiversity, conservation and current threats to European Honeybees. Apidologie 40: 263–284.

De Santis L & Cornejo LG (1968) La abeja africana *Apis mellifera adansonii* en América del Sur. Revista de la Facultad de Agronomía 44: 17–35.

Earl DA & von Holdt BM (2012) Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359–361.

Estoup A, Garnery L, Solignac M & Cornuett JM (1995) Microsatellite variation in honeybee (*Apis mellifera* L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. Genetics 140: 679–695.

Evanno G, Regnaut S & Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611–2620.

Excoffier L, Laval G & Schneider S (2005) ARLEQUIN ver 3.0: an integrated software package for population genetic data analysis. Evolutionary Bioinformatics Online 1: 47–50.

Fletcher DJC (1991) Interdependence of genetics and ecology in a solution to the African bee problem. The 'African' Honey Bee (ed. by M Spivak, DJC Fletcher & MD Breed), pp. 77–94. Westview Press, Boulder, CO, USA.

Franck P, Garnery L, Solignac M & Cornuet JM (2000) Molecular confirmation of a fourth lineage in honeybees from the Near East. Apidologie 31: 167–180.

Franck P, Garnery L, Loiseau A, Oldroyd BP, Hepburn HR et al. (2001) Genetic diversity of the honeybee in Africa: microsatellite and mitochondrial data. Heredity 86: 420–430.

Galindo-Cardona A, Acevedo-Gonzalez JP, Rivera-Marchand B & Giray T (2013) Genetic structure of the gentle Africanized honeybee population (gAHB) in Puerto Rico. BMC Genetics 14: 65.

Garnery L, Cornuet JM & Solignac M (1992) Evolutionary history of the honeybee (*Apis mellifera* L.) inferred from mitochondrial DNA analysis. Molecular Ecology 1: 145–154.

Garnery L, Solignac M, Celebrano G & Cornuet JM (1993) A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic strucuture of *Apis mellifera*. Experientia 49: 1016–1021.

Garnery L, Mosshine EH, Oldroyd BP & Cornuet JM (1995) Mitochondrial DNA variation in Moroccan and Spanish honeybee populations. Molecular Ecology 4: 465–471.

Garnery L, Franck P, Baudry E, Vautrin D, Cournet JM et al. (1998a) Genetic biodiversity of the West European honeybee (*Apis mellifera mellifera* and *Apis mellifera iberica*). I. Mitochondrial DNA. Genetics Selection and Evolution 30: 31–47.

Garnery L, Franck P, Baudry E, Vautrin D, Cournet JM et al. (1998b) Genetic biodiversity of the West European honeybee (*Apis mellifera mellifera* and *Apis mellifera iberica*). II. Microsatellite loci. Genetics, Selection and Evolution 30: 49–74.

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95–98.

Hall HG & Smith DR (1991) Distinguishing African and European honeybee matrilines using amplified mitocondrial DNA. Proceedings of the National Academy of Sciences of the USA 88: 4548–4552.

Harrison JF, Fewell JH, Anderson KE & Loper GM (2006) Environmental physiology of the invasion of the Americas by Africanized honeybees. Society for Integrative and Comparative Biology 46: 1110–1122.

Jombart T & Collins C (2015) A Tutorial for Discriminant Analysis of Principal Components (DAPC) Using Adegenet 2.0.0. MRC Centre for Outbreak Analysis and Modelling, Imperial College, London, UK.

Kerr WE (1967) The history of the introduction of African bees to Brazil. South African Bee Journal 39: 3–5.

Kerr WE, Leon D, del Río S & Barrionuevo M (1982) The southern limits of the distribution of the Africanized honeybee in South America. American Bee Journal 122: 196–198.

Kono Y & Kohn JR (2015) Range and frequency of Africanized honey bees in California (USA). PLoS ONE 10: e0137407.

Kraus FB, Franck P & Vandame R (2007) Asymmetric introgression of African genes in honeybee populations (*Apis mellifera* L.) in Central Mexico. Heredity 99: 233–240.

Lobo Segura JA (2000) Highly polymorphic DNA markers in an Africanized honeybee population in Costa Rica. Genetics and Molecular Biology 23: 317–322.

Magnus RM, Tripodi AD & Szalanski AL (2011) Mitochondrial DNA diversity of honeybees, *Apis mellifera* L. (Hymenoptera: Apidae) from queen breeders in the United States. Journal of Apicultural Science 55: 5–14.

Magnus RM, Tripodi AD & Szalanski AL (2014) Mitochondrial DNA diversity of honeybees (*Apis mellifera*) from unmanaged colonies and swarms in the United States. Biochemical Genetics 52: 245–257.

McGregor SE (1976) Insect pollination of cultivated crop plants, Agriculture Handbook, pp. 1–9. United States Department of Agriculture, Washington, DC, USA.

Meixner MD, Pinto MA, Bouga M, Kryger P, Ivanova E et al. (2013) In standard methods for characterizing subspecies and ecotypes of *Apis mellifera*. Journal of Apicultural Research 52: 1–27.

Muñoz I, Dall'Olio R, Lodesani M & De la Rúa P (2014a) Estimating introgression in *Apis mellifera* siciliana populations: are the conservation islands really effective? Insect Conservation and Diversity 7: 563–571.

Muñoz I, Pinto MA & De la Rúa P (2014b) Effects of queen importation on the genetic diversity of island honeybee populations (*Apis mellifera* Linneaus 1758). Journal of Apicultural Research 53: 296–302.

Nedić N, Stanisavljević L, Mladenović M & Stanisavljević E (2009) Molecular characterization of the honeybee *Apis mellifera carnica* in Serbia. Advances of Biological Sciences 61: 587–598.

Oleksa A, Wilde J, Tofilski A & Chybicki IJ (2013) Partial reproductive isolation between European subspecies of honeybees. Apidologie 44: 611–619.

Palacio MA, Figini E, Rodriguez E, Ruffinengo S, Bedascarrasbure E et al. (2000) Changes in a population of *Apis mellifera* selected for its hygienic behaviour. Apidologie 31: 471–478.

Palacio MA, Rodriguez E, Gonçalves L, Bedascarrasbure E & Spivak M (2010) Hygienic behaviors of honeybees in response to brood experimentally pin-killed or infected with *Ascosphaera apis*. Apidologie 41: 602–612.

Pinto MA, Rubink WL, Coulson RN, Patton JC & Johnston JS (2004) Temporal pattern of Africanization in a feral honeybee population from Texas inferred from mitochondrial DNA. Evolution 58: 1047–1055.

Pinto MA, Rubink WL, Patton JC, Coulson RN & Johnston JS (2005) Africanization in the United States: replacement of feral european honeybees (*Apis mellifera* L.) by an African hybrid swarm. Genetics 170: 1653–1665.

Pritchard JK, Stephens M & Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Quezada-Euán J, Pérez-Castro EE & May-Itzá WJ (2003) Hybridization between European and African-derived honeybee populations (*Apis mellifera*) at different altitudes in Perú. Apidologie 34: 217–225.

Rivera-Marchand B, Oskay D & Giray T (2012) Gentle Africanized bees on an oceanic island. Evolutionary Applications 5: 746–756.

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Molecular Ecology Resources 8: 103–106.

Ruttner F, Tassencourt L & Louveaux J (1978) Biometrical-statistical analysis of the geographic variability of *Apis mellifera* L. I. Material and methods. Apidologie 9: 363–381.

Schneider SS, De Grandi-Hoffman G & Smith DR (2004) The African honeybee: factors contributing to a successful biological invasion. Annual Review of Entomology 49: 351–376.

Sheppard WS, Rinderer T, Mazzoli J, Stelzer J & Shimanuki H (1991) Gene flow between African- and European-derived honeybee populations in Argentina. Nature 349: 782–784.

Sheppard WS, Rinderer TE, Garnery L & Shimanuki H (1999) Analysis of Africanized honeybee mitochondrial DNA reveals further diversity of origin. Genetics and Molecular Biology 22: 73–75.

SIGA (2018) Sistema de Información y Gestión Agrometeorológica – Instituto Nacional de Tecnología Agropecuaria, INTA, v.1.0.5. Available at: www.siga2.inta.gov.ar (accessed 1 March 2018).

Solignac M, Vautrin D, Loiseau A, Mougel F, Baudry E et al. (2003) Five hundred and fifty microsatellite markers for the study of the honeybee (*Apis mellifera* L.) genome. Molecular Ecology Notes 3: 307–311.

Southwick EE, Roubik DW & Williams JM (1990) Comparative energy balance in groups of Africanized and European honey bees: ecological implications. Comparative Biochemistry and Physiology 97: 1–7.

Staden R (1996) The Staden sequence analysis package. Molecular Biotechnology 5: 233–241.

Taylor OR (1977) The past and possible future spread of Africanized honeybees in the Americas. Bee World 58: 19–30.

Taylor OR & Spivak M (1984) Climatic limits of tropical African honeybees in the Americas. Bee World 65: 38–47.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG (1997) The clustal x windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25: 4876–4882.

Van Oosterhout C, Hutchinson WF, Wills DPM & Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538.

Weir BS & Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Whitfield CW, Behura SK, Berlocher SH, Clark AG, Johnston JS et al. (2006) Thrice Out of Africa: ancient and recent expansions of the honeybee, *Apis mellifera*. Science 314: 642–645.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Number of Bayesian inferred criterion.

Figure S2. Number of samples per province included in each cluster after applying Bayesian inferred criterion for discriminant analysis of principal components (see Material and methods section for explanation of the province codes).

Figure S3. (A–E) Linear regression of the degree of Africanization on various climate-related variables: maximum daily temperature in (A) January and (B) July, minimum daily temperature in (C) January and (D) July, (E) number of hours below 4 °C in July. Assignment probabilities are based on the analysis of nuclear markers with STRUCTURE and correspond to the population that made the largest genetic contribution to feral colonies of Formosa province (see Figures 1 and 2). (F–J) Probability of detection of African haplotypes estimated by logistic regression: maximum daily temperature in (F) January and (G) July, minimum daily temperature in (H) January and (I) July, (J) number of hours below 4 °C in July. Honeybees were collected from feral (F) and commercial (C) colonies. Goodness of fit (GF) in logistic regressions corresponds to the ratio deviance/degrees of freedom.

Table S1. Number of feral and commercial Argentinian honeybee colonies per province (see Material and methods section for explanation of the codes).

Table S2. Comparison among nucleotide sequences obtained in the present work (JQ582432-40) and the Genbank database. Reference sequences were established taking into account the best hit in Blast (NCBI) with annotated haplotype and the most recent date of publication.

Table S3. Basic characteristics and PCR amplification conditions of eight *Apis mellifera* microsatellite loci used.