Evolutionary history of the chili pepper *Capsicum baccatum* L. (Solanaceae): domestication in South America and natural diversification in the Seasonally Dry Tropical Forests

MARISEL A. SCALDAFERRO[1,2], GLORIA E. BARBOZA[1,3] and M. CRISTINA ACOSTA[1,2]*

[1]*Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET-Universidad Nacional de Córdoba, Argentina*
[2]*Facultad de Ciencias Exactas, Físicas y Naturales (Universidad Nacional de Córdoba), Argentina*
[3]*Facultad de Ciencias Químicas (Universidad Nacional de Córdoba), Argentina*

Dedicated to the Memory of Eduardo A. Moscone

Plant domestication genetics has been of interest not only for evolutionary biologists but also for anthropologists and breeders, because of its past and present role in human cultural evolution. *Capsicum baccatum*, which grows in the Seasonally Dry Tropical Forests of South America (SDTF), was domesticated by native peoples and used as a spice. The centre of domestication has been discussed, and vast territories of Peru and Bolivia have been proposed. The aims of this research were to elucidate the geographical origin of cultivated *C. baccatum* and to infer geological and climate events that have influenced the distribution of its genetic variability. Samples were collected from 25 localities across the entire range of wild *C. baccatum* and analysed using DNA sequences and phylogeographical approaches; they were then compared with 20 samples of cultivated forms obtained from different South American markets. We found a main centre of *C. baccatum* domestication spanning Bolivian Amazonia and the inter-Andean valleys. We also inferred an ancient cultivation site between Bolivia and Argentina. Finally, we found two lineages of wild populations distributed in nuclei of SDTFs which have been fragmented during glacial periods, when aridity increased and expansion of Chaco vegetation fragmented the region.

ADDITIONAL KEYWORDS: *Capsicum baccatum* – chloroplast DNA sequences – crop origin – haplotype network – molecular dating – Solanaceae – South America.

INTRODUCTION

The domestication of plants is one of the most significant cultural and evolutionary transitions in human history. Investigating when, where and how domestication took place is of interest to crop evolutionists and of practical importance for plant breeders and conservationists. Research into the domestication and spatio-temporal origins of this process has increased significantly over the past decade through archaeological research, advances in DNA/RNA sequencing techniques, and methods used to recover and formally identify those changes in interactions among plants and animals that lead to domestication (Larson *et al.*, 2014). Genetic studies that examine the geographical origins of a crop typically rely on genome-wide neutral markers, which are used to assess allele frequencies in a crop compared with populations of its wild relatives. Due to the strong genetic bottlenecks that occur during the domestication process, allelic diversity in the crop is expected to be a subset of that found in the wild population(s) from which it was derived. Thus, if populations of the wild progenitor are extant in the geographical location(s) where domestication occurred, the geographical origin(s) of domestication can potentially be pinpointed to a particular population or region (Gross & Olsen, 2010).

---

*Corresponding author. E-mail: mcacosta@imbiv.unc.edu.ar

466

One of the most successful applications of phylogeographical analysis has been the investigation of sites of origin of domesticated crops (Londo *et al.*, 2007). Phylogeographical studies draw inferences about the history of population divergence based on associations between the geographical distribution of alleles and their genealogical relationships (Avise, 2000). The number of domestication events and regions in cultivated rice and peanuts was determined using a phylogeographical approach based on DNA sequence variation (Londo *et al.*, 2007; Grabiele *et al.*, 2012). In addition, such approaches have made a significant contribution to the interpretation of species in response to climate and geological changes, such as the glaciation events of the Neogene (Turchetto-Zolet *et al.*, 2013; Daneck, Fér & Marhold, 2016; Zhang, Zhang & Sanderson, 2016).

Capsicum baccatum L. is one of the most economically important species of the genus, with three varieties, the wild *C. baccatum* var. *baccatum*, and two [var. *pendulum* (Willd.) Eshbaugh and var. *umbilicatum* (Vell.) Hunz. and Barboza] varieties, cultivated commercially mainly in South America (Pickersgill, 1997; Djian-Caporalino *et al.*, 2007; Peralta *et al.*, 2008). The wild form of the species typically bears small, erect, deciduous fruits (Fig. 1A; D'Arcy & Eshbaugh, 1974; Barboza, 2013), whereas the domesticated forms usually bear larger, pendant or bell-shaped, persistent fruits (Fig. 1B–D). The fruits are used as a spice in regional cuisine, although they are also of interest for a wide range of other applications, such as a natural colouring agent, an ingredient in pharmaceuticals (as a rubefacient and stomachic), and a highly pungent extract used in self-defence sprays, animal repellents and insecticides (Bosland, 1996). The wild form of the species is distributed in South America, from Colombia, Peru, Bolivia, Brazil and Paraguay to north-west and north-east Argentina (Moscone *et al.*, 2007; Barboza, 2013), and is recognized under various common names: 'cumbari', 'ají quitucho', 'quitucho' and 'ají del monte'.

Starch microfossils of *C. baccatum* in Peru date from 4000 years BP to European contact, indicating an early use of domesticated peppers in ancient America (Perry *et al.*, 2007). Based on combinations of archaeological evidence and genetic analyses, researchers suggested that the centre of domestication of *C. baccatum* was the lowlands of Bolivia (Pickersgill, 1984; Eshbaugh, 1993; Perry *et al.*, 2007). However, based on a Bayesian spatial clustering methodology with AFLP markers, Albrecht *et al.* (2012) found that the cultivated *C. baccatum* was domesticated independently at two sites, one in the Andes highlands (Peru and Bolivia) and the other in the lowland of Paraguay. More recently, however, based on ethnobotanical studies and bibliographic revisions, Clement *et al.* (2016) located the centre of domestication in the area of the upper Madeira River basin, in the Amazonia of Bolivia and Brazil.

Capsicum has been the focus of taxonomic, cytogenetic and phylogenetic studies (Scaldaferro *et al.*, 2006; Scaldaferro, Grabiele & Moscone, 2013; Barboza *et al.*, 2011; Carrizo García *et al.*, 2016). However, few studies have analysed the intraspecific genetic diversity across the total geographical distribution in order to reconstruct the natural evolutionary history of their species, and even fewer have attempted to elucidate the domestication origin of their cultivated species. Most genetic diversity studies are focused on germplasm bank characterization or on individuals cultivated in home gardens, always with an emphasis on the characterization of domesticated forms (Ibiza *et al.*, 2011; Albrecht *et al.*, 2012). In *C. baccatum*, Albrecht *et al.* (2012) using AFLP markers reported two principal genetic groups (western and eastern) in
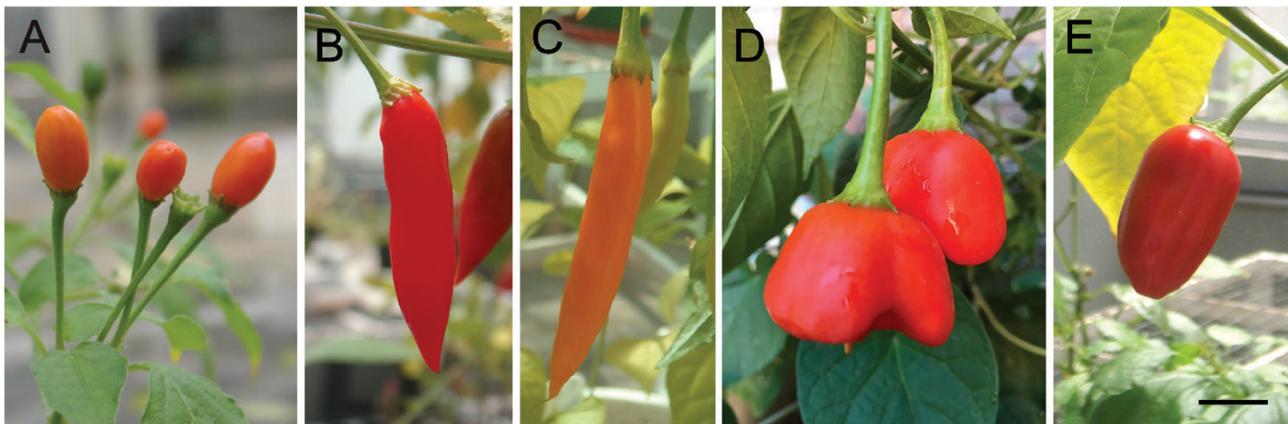
Figure 1. Fruits of *Capsicum baccatum*. A, *C. baccatum* var. *baccatum* SVL. B, *C. baccatum* var. *pendulum* 642. C, *C. baccatum* var. *pendulum* 215. D, *C. baccatum* var. *umbilicatum* 063. E, *C. baccatum* 'became wild' PRF. Scale bar = 2 cm.

accessions from a germplasm repository; most of the specimens studied were cultivated forms and only 13 accessions came from wild localities, although they did not include wild specimens of the extremes of the natural distribution range in Argentina.

The distribution of wild *C. baccatum* var. *baccatum* in South America is coincident with four nuclei of the Seasonally Dry Tropical Forests (SDTF). The SDTF is a wide and fragmented biome, extending from Mexico to Argentina and throughout the Caribbean (Prado & Gibbs, 1993; Pennington, Prado & Pendry, 2000). For much of the 20th century, the SDTF of Central and South America remained conceptually included within other neighbouring vegetation units of the continent, such as the Chaco, the Yungas or the Amazonian Dominions (e.g. Cabrera, 1971; Morrone, 2014); however, it has also been proposed as a new phytogeographical dominion (Mogni *et al.*, 2015). Prado & Gibbs (1993) and Pennington *et al.* (2000) proposed that the dry forest nuclei are remnants (i.e. refugia) of a much larger, single formation (the Pleistocene dry forest arc) that spanned these disjunct dry forest regions during drier periods of the Pleistocene, e.g. the Last Glacial Maximum (LGM, *c.* 20 000 years BP), extending into regions now covered by Amazon rainforest and Cerrado savannas.

In this study we used a phylogeographical approach based on DNA sequence variation from two chloroplast gene regions to elucidate the evolutionary history of *C. baccatum* and to determine the effect of biome past fragmentation on its genetic variation. Our objectives were to (1) examine the DNA sequence variation in cultivated and wild *C. baccatum*, (2) explore the geographical pattern of cultivated and wild haplotypes, (3) determine the number of potential sites of domestication within South America, and (4) infer geological events and climate changes that have influenced the distribution of genetic variability of the species. We hypothesized that cultivated forms of *C. baccatum* would have multiple sites of domestication, from where they expanded to the rest of South America. In addition, we expected that haplotype divergences would be concomitant with fragmentation of the STDF.

MATERIALS AND METHODS

SAMPLING OF NATURAL AND CULTIVATED SPECIMENS

Leaf material was collected from 38 individuals belonging to *C. baccatum* var. *baccatum* at 25 locations distributed throughout its entire geographical range in southern South America (Table 1). Some individuals presented transitional forms with moderately sized semipendulous fruits (eight); these were considered 'became wild', i.e. originated in the crop but growing in the wild (represented by squares in the map and by symbols with dots in the haplotype network; Figs 1E,

2). In addition, fruits from different localities of South America were purchased from the main markets of the cities (13 locations), and then cultivated in the glasshouse for further species identification. Thus, a total of 20 individuals belonging to *C. baccatum* var. *pendulum* and var. *umbilicatum* were selected and introduced in our study (Table 1). Voucher specimens were deposited in the herbarium of the Botanical Museum of Cordoba, Argentina (CORD). Note that the low number of individuals analysed per population is because the wild form of the species is very rare and only one or two individuals were found at most of the sites. In addition, many of the collected sites registered in herbarium data have already been degraded by anthropogenic activity. Thus, only DNA extracted from herbarium specimens in several localities was available to us. Here, a total of 25 wild locations were sampled throughout its entire range of distribution, for the first time, covering the whole range of the species distribution.

DNA EXTRACTION, AMPLIFICATION AND SEQUENCING

Fresh tissue was dehydrated in silica gel. Total DNA was extracted with a DNeasy plant mini kit (Qiagen, Hilden, Germany). Two non-coding regions of the chloroplast genome were amplified using primer pairs *rpl32-trnL* and *trnQ-5'rps16* (Shaw *et al.*, 2007). We analysed only chloroplast genomes to avoid the detection of signals of hybridization or introgression of cultivated individuals into wild *C. baccatum* nuclear genomes. In many traditional communities throughout Latin America, wild *Capsicum* species are often found growing close to cultivated peppers (Van Zonneveld *et al.*, 2015).

The PCR mix contained 2 µL of template DNA (10 ng), 0.625 U GoTaq DNA polymerase (Promega, Madison, WI, USA), 5 µL of Green GoTaq reaction buffer (Promega), 0.25 mM of each dNTP and 0.3 µM of each primer in a total volume of 25 µL. The PCR cycling conditions were DNA denaturation at 80 °C for 5 min followed by 30 cycles of denaturation at 95 °C for 1 min, primer annealing at 50 °C for 1 min, followed by a ramp of 0.3 °C/s to 65 °C, and primer extension at 65 °C for 4 min; and a final extension step of 5 min at 65 °C. All reactions were carried out on an Eppendorf Mastercycler thermocycler (Eppendorf, Hamburg, Germany). Amplification products were separated by electrophoresis on a 1% agarose gel, stained with Syber Safe (Invitrogen, Eugene, OR, USA), and visualized with a UV transilluminator. The PCR-amplified products were sequenced by Macrogen Inc. (Seoul, Korea). Sequences were deposited in GenBank (accession numbers: *rpl32-trnL*: MH160187–MH160204; *trnQ-5'rps16*: MH172185–MH172202). For all subsequent analyses, the two chloroplast regions were concatenated for each individual. Sequences

Table 1. Collection sites, geographical coordinates, altitude, code and voucher number, sample origin and haplotypes of the studied *Capsicum baccatum* specimens

Locality	Geographical coordinates (S, W)		Altitude (m)	Code and voucher number	Origin	Haplotype*
C. baccatum var. *baccatum*						
Peru, Cusco District, Convención Province, Santa Teresa.	13°00′47″	72°37′10″	1208	**CUZ** *GEB 3419*	wild	2 (1)
Bolivia, Santa Cruz Department, Andres Ibañez Province, Basilo.	18°05′26″	63°11′40″	499	**BAS** *AAC & ANS*	wild	2 (2)
Bolivia, Santa Cruz Department, Ñuflo de Chavez Province, San Ramón.	16°35′33″	62°29′34″	866	**SRA** *JGS et al. 3448*	wild	15 (1)
Bolivia, Santa Cruz Department, Velasco Province, Santa Rosa.	17°27′32″	63°24′29″	344	**VEL** *JGS et al. 3580*	wild	8 (1)
Bolivia, Santa Cruz Department, Florida Province, Bermejo.	18°07′48″	63°38′01″	807	**BER** *GEB & CC* 3646	wild	2 (2), 6 (1)
Bolivia, Chuquisaca Department, Luis Calvo Province, Ñancaroinza.	20°42′22″	63°17′42″	744	**NAN** *AAC & ANS 5488*	wild	7 (1)
Bolivia, La Paz Department, Sud Yungas Province, Coroico.	15°28′00″	67°10′00″	720	**COR** *LG et al. 1998*	wild	2 (3)
Brazil, São Paulo State, Mairinque.	23°32′45″	47°11′00″	859	**SPO** *MVR 8*	became wild	2 (3)
Brazil, São Paulo State, Campinas.	22°53′59″	47°56′16″	683	**JOA** *PTS 10*	wild	17 (1)
Brazil, Santa Catarina State, Rio do Sul.	27°28′01″	53°59′54″	280	**RGS** *LAM et al. 263*	wild	14 (1)
Brazil, Mato Grosso do Sul State, Corumbá.	19°00′33″	57°39′11″	226	**COU** *GAD 19126*	wild	18 (1)
Paraguay, Presidente Hayes Department, Nanawa.	25°16′32″	57°41′49″	63	**PRN** *MAS 40*	became wild	2 (1)
Paraguay, Presidente Hayes Department, Puerto Falcón.	25°15′23″	57°42′55″	66	**PRF** *MAS 38*	became wild	2 (1)
Paraguay, Paraguarí Department, Acahay.	25°52′56″	57°09′28″	280	**ACA** *RSW et al. 129*	wild	16 (1)
Argentina, Jujuy Province, El Carmen Department, Dique Las Maderas.	24°26′56″	65°15′01″	1200	**JUJ** *AAC & ANS 5262*	wild	1 (3)
Argentina, Jujuy Province, Ledesma Department, Agua de las Cañas.	23°45′09″	64°51′20″	718	**CAN** *GEB 1921*	wild	5 (1)
Argentina, Salta Province, Oran Department, Aguas Blancas.	22°43′30″	64°21′54″	417	**BLA** *AAC & ANS 5469*	wild	9 (2), 10 (1)
Argentina, Salta Province, La Caldera Department, La Caldera.	24°36′07″	65°22′51″	1380	**SAL** *AAC & ANS 5258*	wild	1 (1)
Argentina, Salta Province, Capital Department, Sierra de Vélez.	24°45′33″	65°22′38″	1350	**SVL** *GEB 1805*	wild	1 (2), 4 (1)
Argentina, Formosa Province, Pirané Department, Pirané.	25°37′08″	59°10′45″	87	**PIR** *MMC 823*	wild	12 (1)
Argentina, Misiones Province, Santa Ana Department, Santa Ana.	27°20′10″	55°35′15″	102	**ANA** *MAS 47*	wild	13 (1)
Argentina, Misiones Province, Gral. San Martín Department, Puerto Oro Verde.	26°49′40″	54°59′39″	171	**SMA** *GEB et al. 1038*	became wild	2 (1)
Argentina, Misiones Province, Santa Ana, Department, Santa Ana.	27°20′40″	55°35′59″	124	**SAN** *MAS 48*	became wild	3 (1)

Table 1. *Continued*

Locality	Geographical coordinates (S, W)		Altitude (m)	Code and voucher number	Origin	Haplotype*
Argentina, Misiones Province, San Ignacio Department, Teyucuaré.	27°16′12″	55°32′57″	181	**TEY** *MAS 49*	became wild	2 (1)
Argentina, Corrientes Province, San Martín Department, Guaviraví.	29°23′05″	56°47′04″	62	**GUA** *MAS 50*	wild	16 (1)
C. baccatum var. *pendulum*						
Ecuador, Manabí Province, Puerto López, bought at market place.	1°33′33″	80°48′39″	12	**061** *MAS 61*	cultivated	2 (2)
Peru, Piura Department, Piura, bought at market place.	5°11′13″	80°38′54″	43	**019** *MAS 19*	cultivated	2 (2)
Peru, Cusco District, Cusco, bought at market place.	13°31′06″	71°58′41″	1208	**573** *MAS & EK 70*	cultivated	2 (1)
Bolivia, Cochabamba Department, Pairumani Farmhouse, cultivated at Fundación Simón Patiño, Granja Modelo.	17°21′40″	66°19′28″	2632	**320** *MAS 43*	cultivated	2 (2)
Bolivia, Santa Cruz Department, Santa Cruz de La Sierra, bought at market place.	17°47′39″	63°10′57″	423	**642** *GEB & CC 3642*	cultivated	2 (1)
Bolivia, Santa Cruz Department, Valle Grande, bought at market place.	18°29′29″	64°06′26″	2050	**062** *MAS 62*	cultivated	2 (2)
Bolivia, Tarija Department, Villa Montes, bought at market place.	21°15′57″	63°28′37″	392	**037** *MAS 37*	cultivated	2 (1)
Brazil, Minas Gerais State, Belo Horizonte, bought at market place.	19°55′12″	43°56′40″	918	**798** *GEB et al. 798*	cultivated	2 (1)
Brazil, São Paulo State, Santo Antonio de Posse, bought at market place.	23°34′09″	46°36′41″	673	**056** *MAS 56*	cultivated	2 (2)
Argentina, Salta Province, Salta, bought at market place.	24°47′16″	65°24′38″	1196	**247** *EAM 247*	cultivated	2 (1)
Argentina, Salta Province, cultivated at INTA Cerrillos.	24°53′03″	65°28′05″	1242	**215** *EAM & RN 215*	cultivated	11 (2)
C. baccatum var. *umbilicatum*						
Brazil, Rio de Janeiro State, Petrópolis, bought at market place.	22°30′49″	43°11′17″	732	**057** *MAS 57*	cultivated	2 (1)
Argentina, Cordoba Province, Santa Rosa de Calamuchita, cultivated at a private home.	32°04′24″	64°32′13″	620	**063** *MAS 63*	cultivated	2 (2)

* Sample sizes are given in parentheses.
Collector's names: GEB, G. Barboza; CC, C. Carrizo; MMC, M. Cerana; AAC, A. Cocucci; GAD, G. Damasceno Junior; LG, L. Guzmán; EK, E. Kowaljow; LAM, L. A. Mentz; EAM, E. Moscone; RN, R. Neumann; MVR, V. Romero; MAS, M. Scaldaferro; JGS, G. Seijo; ANS, A. Sérsic; PTS, P. Souza, RSW, R. S. Williams.

were aligned using MEGA 6.0 (Tamura *et al.*, 2013), with manual adjustments made as needed. Gaps were coded following the 'simple indel coding' method (Simmons & Ochoterena, 2000).

DATA ANALYSES

To examine genetic relationships among the haplotypes defined by DNAsp v. 6.11.01 (Rozas *et al.*, 2003), a median-joining network was constructed using NETWORK v. 5.0.1.0 (Bandelt, Forster & Röhl, 1999). Ambiguous connections (loops) in the network were resolved using predictions from coalescent theory (Crandall & Templeton, 1993). Molecular diversity indices, i.e. number of haplotypes, percentage CG content, number of polymorphic sites, transitions, transversions and indels, and haplotype and nucleotide diversity (Table 2) for the species and

Table 2. Molecular diversity indices calculated for *Capsicum baccatum* var. *baccatum* and each group of populations identified ('became wild' accessions were excluded)

Grouping	P	N	Size	Range	H	CG%	Ps	ti	tv	I	h	π
C. baccatum var. *baccatum*	19	30	2463	2306–2423	16	26.91	22	5	6	11	0.8966 (0.0398)	0.0140 (0.0070)
West Group	12	23	2432	2306–2423	10	26.91	13	3	2	8	0.8261 (0.0590)	0.0103 (0.0052)
East Group	7	7	2428	2338–2403	6	26.90	9	2	4	3	0.9524 (0.0955)	0.0161 (0.0092)

Number of populations (P) and individuals (N), alignment size and range, number of haplotypes (H), percentage of CG content, number of polymorphic sites (Ps), transitions (ti), transversions (tv), indels (I), and haplotype (h) and nucleotide (π) diversity are shown. Alignment size without outgroups and range are indicated in base pairs (bp). Standard errors are indicated in parentheses.
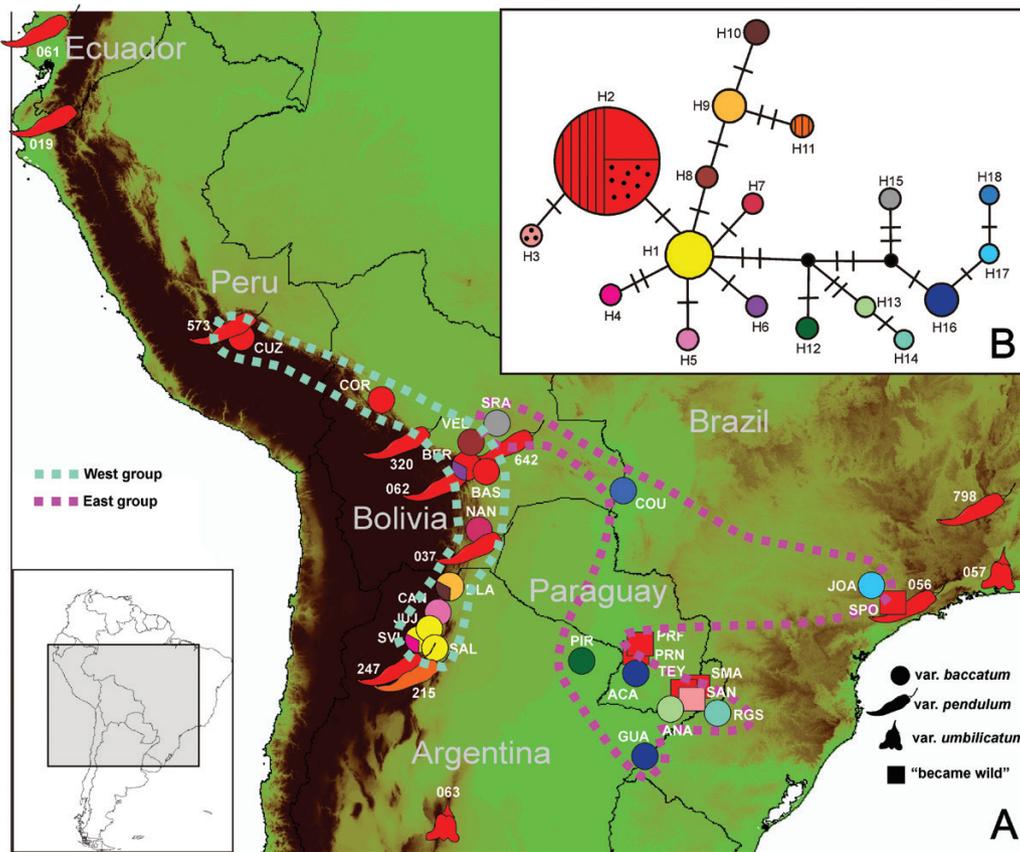
Figure 2. Geographical distribution and genealogical relationships of the chloroplast DNA (cpDNA) haplotypes found in *Capsicum baccatum*. A, partial map of South America showing the distribution of cpDNA haplotypes within the studied *C. baccatum* populations. Groups of populations and symbols representing the different varieties of the species are specified in the figure. A complete map of South America is shown in the left box. B, median-joining network showing relationships among the cpDNA haplotypes found. The size of the symbols is proportional to the haplotype frequencies. *Capsicum baccatum* var. *pendulum* is represented by striped symbols. *Capsicum baccatum* 'became wild' is represented by dotted symbols.

groups of populations were calculated in ARLEQUIN v. 3.5 (Excoffier & Lischer, 2010).

The phylogenetic tree was constructed and divergence times were estimated using BEAST version 1.6.2 (Drummond & Rambaut, 2007). The input file was prepared in BEAUti version 1.6.2 (provided in the BEAST package). To estimate haplotype divergences we obtained sequences of the *rpl* region

in *C. chacoense* (GEB et al. 1793 and MAS 44) using the same methodology as in *C. baccatum*. In addition, sequences of *Lycium* L. and *Solanum* L. species were taken from GenBank. The split of *Lycium* from the genera *Solanum* and *Capsicum* was considered the root of the tree and was set to 21 Mya. Additional time constraints were set for the split between *Solanum* and *Capsicum* at 19 Mya, between *S. tuberosum* and

S. lycopersicum at 8 Mya, and between *C. baccatum* and *C. chacoense* at 2.68 Mya; all dates were taken from Särkinen *et al.* (2013). The substitution model was GTR with a Gamma site heterogeneity model with four categories, following the results of MrModeltest version 2.2 (Nylander, 2004); the clock was set as an uncorrelated lognormal relaxed model, and the Yule process was selected as a prior for the distribution of divergence dates. The Monte Carlo Markov chain was set to run for 1.5×10^7 generations, sampling every 1000 cycles.

RESULTS

Sequence analysis of 38 individuals sampled from 25 natural locations yielded 17 distinct chloroplast DNA (cpDNA) haplotypes (Table 1, Fig. 2, Supporting Information Tables S1, S2). In addition, almost all cultivated specimens (18 individuals from 12 locations) showed the same haplotype (H2), except for one accession (two individuals) from INTA Cerrillos, Salta province, Argentina (Fig. 1C), which exhibited haplotype H11. Individuals here considered 'became wild' also displayed haplotype H2, except for specimens from Santa Ana (SAN), in Misiones province, Argentina, which exhibited haplotype H3. Most haplotypes (15) were unique to only one accession. The length of the *rpl32-trnL* sequence varied from 1071 bp in haplotype H5 to 1080 bp in haplotypes H9–H11, whereas the intergenic spacer *trnQ-5'rps16* was more variable, ranging between 1234 bp (H4) and 1351 bp (H6). Alignment of the final matrix included 2463 bp and required the introduction of 11 gaps, ranging from 1 to 148 bp in length. Haplotypes H9–H11 had one duplication of 8 bp in *rpl32-trnL*, and haplotypes H1–H7, H12–H14 and H16-H18 had one duplication of 20 bp in the *trnQ-5'rps16* spacer. In addition, a hypervariable region with T and A nucleobases was found in the latter.

The configuration of the haplotype network and the phylogenetic tree showed two main lineages that coincide with the geographical distribution (Figs 2B, 3). Group 1 consisted of the westernmost localities on the slopes of the Andes (hereafter West Group) and Group 2 included the eastern populations (hereafter East Group) (Fig. 2A). Median vectors (i.e. non-sampled or extinct haplotypes in the network) connect the haplotypes in the East Group (H12–H18) and with the haplotypes from the West Group (H1–H11). Within the West Group, H2 was the most frequent and widespread haplotype (N = 33), and H1 was the haplotype with the highest number of mutational connections, from which other related haplotypes were derived (H2, H4–H8). Furthermore, the East Group did not include any

widespread haplotype. Thus, H1 could be considered the ancestral haplotype of the West Group while the East Group would have originated from a non-sampled or extinct haplotype ancestor (represented here by a median vector). The highest haplotype and nucleotide diversity were found in the East Group (Table 2). Secondary contact zones between divergent lineages were found in SRA and VEL. Finally, the phylogenetic tree (Fig. 3) from cpDNA sequences of *C. baccatum* showed that the West Group and the East Group diverged at approximately 2.3 Mya (95% confidence and highest posterior density intervals = 28–62 Mya). Other more recent divergences occurred at *c.* 210 and 140 kya.

DISCUSSION

Our results show two principal genetic groups in *C. baccatum* according to the phylogenetic tree (West and East Groups), agreeing with the genetic groups found by Albrecht *et al.* (2012) using AFLP markers. Only one of these lineages (West Group) gave origin to most of the cultivated varieties. As indicated by Albrecht *et al.* (2012), genetic diversity in the wild *C. baccatum* was greater than in the domesticated form of the species. In the most commercially cultivated *C. annuum*, studies using RFLPs, isozyme loci, RAPDs, AFLPs and even cytogenetic methods showed the lack of variability in the domesticated types (Lefebvre, Palloix & Rives, 1993; Paran, Aftergoot & Shifriss, 1998; McLeod *et al.*, 1983; Scaldaferro *et al.*, 2013). Thus, in *C. baccatum*, the domestication process involved a genetic bottleneck associated with a founder effect event due to the selection processes.

ORIGIN OF *C. BACCATUM* CULTIVATED FORMS IN SOUTH AMERICA

We found two different haplotypes (H2 and H11) in the cultivated forms, with the red haplotype (H2) the most common, present in 12 cultivated accessions of the 13 studied. The geographical distribution of this haplotype in the wild suggests that lowlands of Bolivia and inter-Andean valleys of Peru have been the most important centres of *C. baccatum* domestication (Fig. 4A). Thus, phylogeographical methods support the early hypothesis of these regions as sites of domestication of this species, occurring by 4000 years BP at the latest, probably by Arawak people inhabiting the Llanos de Mojos (i.e. Moxeños), Guarani populations, both in Bolivia, or pre-Incas in Peru (see Fig. 4A) (Perry *et al.*, 2007; Piperno, 2011; Brown *et al.*, 2013; Clement *et al.*, 2016). The upper Madeira River basin inhabited by Moxeño people was a probable region
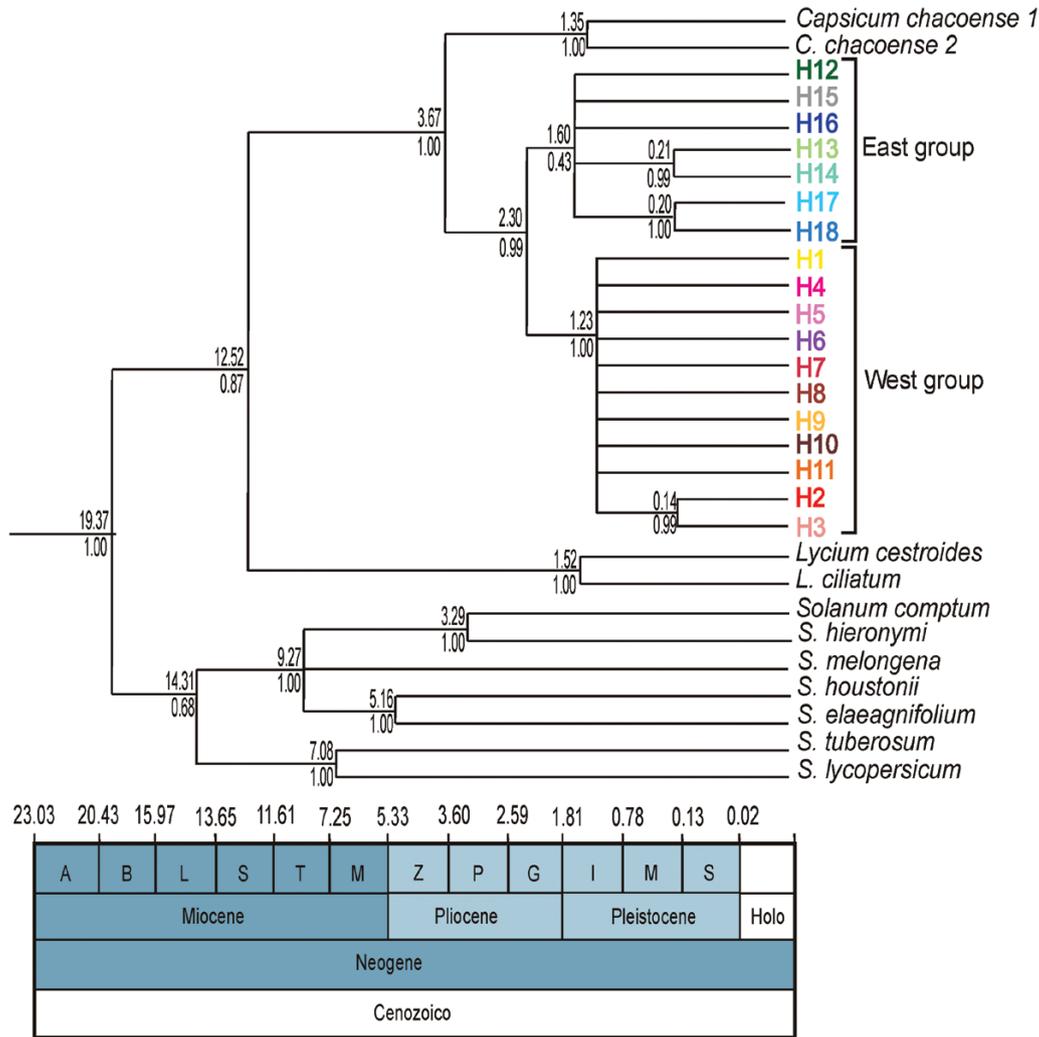
Figure 3. Chronogram showing relations among chloroplast haplotypes (H) and lineages found within *Capsicum baccatum*. Numbers above branches indicate the ages of nodes (Mya) from the molecular dating analysis. Numbers below branches indicate posterior probabilities from Bayesian inference analysis (BPP).

of origin of other crops, such as other chili peppers (*C. pubescens*), three legume seeds (*Arachis hypogaea* L., *Canavalia plagiosperma* Piper, *Phaseolus vulgaris* L.) and cocoyam [*Xanthosoma sagittifolium* (L.) Schott], important drugs, such as coca (*Erythroxylum coca* Lam.), and tobacco (*Nicotiana tabacum* L.), and the major root crop of Amazonia, manioc (*Manihot esculenta* Crantz) and peach palm (*Bactris gasipaes* Kunth) (Grabiele *et al.*, 2012; Clement *et al.*, 2016, and references herein). *Capsicum baccatum* var. *umbilicatum* originated from the west clade, as did var. *pendulum*, although this result is in disagreement with those reported by Albrecht *et al.* (2012). The only *C. baccatum* var. *umbilicatum* accession included in their analysis was collected from Brazil (São Paulo); it was deeply nested within the 'eastern' Group and

closely related to other cultivars from Brazil, primarily those from Minas Gerais.

Our results indicate that various accessions grow in nature, whose haplotypes or genetic variants do not correspond to the wild-type expected lineage according to phylogenetic and phylogeographical analyses. All of these wild specimens had the red haplotype (H2; PRF, PRN, SMA, TEY, SPO), and one derived from it (H3; SAN) in only one case. Fruit size is generally intermediate between that of cultivated and wild plants (Fig. 1E). One possible explanation could be hybridization with wild parental species. In many traditional communities throughout Latin America, wild and semi-wild *Capsicum* species are often found growing close to cultivated peppers, in home gardens or at the edges of cultivated fields, where introgressive
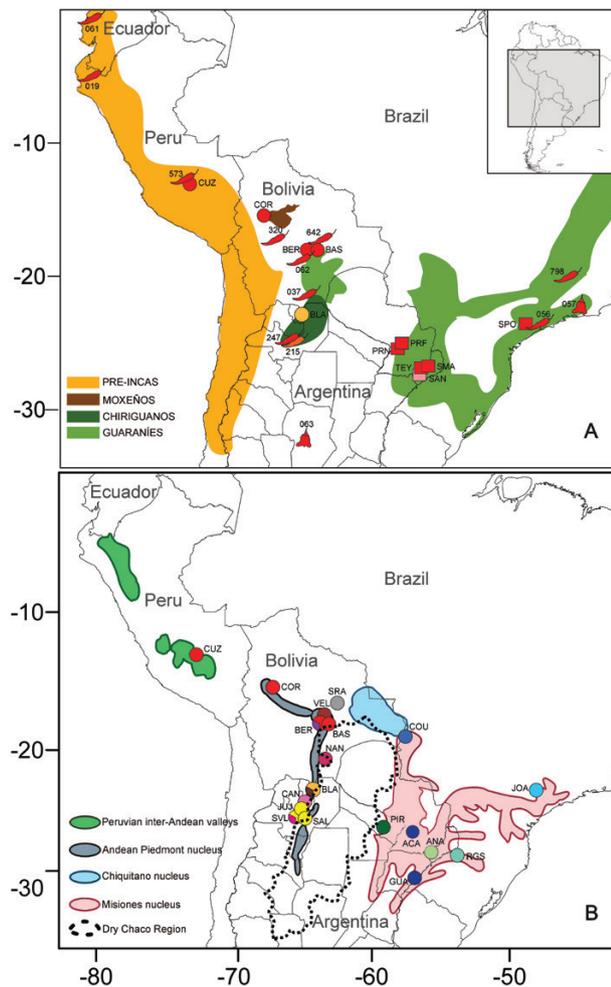
Figure 4. Maps of South America showing the geographical distribution of haplotypes along with different addressed aspects. A, distribution of aboriginal populations in South America before the Spanish conquest and geographical distribution of wild and cultivated samples of *Capsicum baccatum*. B, schematic distributions of the Seasonally Dry Tropical Forests (SDTF; adapted from Prado, 2000; Caetano *et al.*, 2008) and their relations with haplotype structure in the studied *C. baccatum* var. *baccatum* populations.

hybridization takes place between the cultigens and closely related wild species or sub-species, spontaneously giving rise to novel or intermediate forms (Van Zonneveld *et al.*, 2015). Notably, accessions here named 'became wild' have been found only in the east of the species distribution. The region comprising Paraguay, Brazil and Misiones province in Argentina was inhabited by Guarani peoples before arrival of Europeans (Fig. 4A). They were traditionally farmers and cultivated maize (acquired by interethnic contact), cassava (probably domesticated by them), peanut and

sweet potato. Occupation of their lands dedicated to agriculture together with cultural erosion, among other factors, would have caused the abandonment of this intensive practice and many of the crops may have been neglected (Keller, 2012). Hence, we propose that these agricultural peoples once cultivated chili peppers (domesticated in the west), probably obtained through commercial practices. After European conquest, this activity declined, which might have caused the de-selection in attributes of the cultivated fruits (e.g. from large to small size fruit) (F. Zamudio IMBIV, pers. comm.); since then, plants with smaller fruits may have continued to grow in the wild and 'become wild'. The fact that these specimens always carried the cultivated haplotype type (H2) reinforces the idea of their origin as abandoned crops. Chiou *et al.* (2014) noted the reduction of seed size after the Spanish conquest (after 1532 CE), suggesting a loss of selective pressure for seed size in this particular *Capsicum* species.

The cultivated accession EAM & RN 215 from INTA Cerrillos has a different haplotype (H11), which could indicate a secondary origin of domestication for *C. baccatum* var. *pendulum* or could reflect an ancient site of cultivation of the species where hybridization between cultivated and local samples may have occurred. Multiple domestications of different populations of the same species are rare, but there are some records, such as the best known examples of common bean (*P. vulgaris*) and Lima bean (*P. lunatus* L.) (Chacón, Pickersgill & Debouck, 2005; Serrano-Serrano *et al.*, 2010). According to the haplotype network, the haplotype found in this accession is derived from a wild haplotype found in Aguas Blancas, at the frontier between Argentina and Bolivia. Ethnobotanical studies in Baritu National Park (Salta, Argentina) indicate that this region was occupied in pre-Hispanic times by a group of 'Chiriguanos' belonging to the Guarani group, who originated from the north-east of the Bermejo River basin (Fig. 4A). These avid farmers had commercial relationships with people from Tarija and Padcaya in Bolivia, and Santa Victoria in Argentina (Hilgert 1999) and they probably cultivated and commercialized *C. baccatum*.

PHYLOGEOGRAPHICAL IMPLICATIONS

Two principal genetic groups were identified in the wild *C. baccatum* accessions in relation to their geographical distribution in South America. One group included accessions from the western territories of the species distribution (Peru, Bolivia and north-western Argentina) and corresponds to SDTF nuclei of the Andean Piedmont and Peruvian inter-Andean valleys. The second group included samples from the eastern regions (Brazil, Paraguay and eastern Argentina)

and corresponds to SDTF nuclei from Chiquitano and Misiones. *Capsicum baccatum* is in fact not considered in the species database of these forests (: database http://www.dryflor.info/data), and consequently, it was not considered in a recent study of these flora (Banda *et al.*, 2016). Nonetheless the species could effectively be integrated into this database, highlighting the importance of conservation of this biome as a reserve of wild relatives from cultivated species.

Our results supported by molecular dating indicate that a large split between the western and eastern groups of wild *C. baccatum* occurred at 2.3 Mya. A Late Pliocene to Early Pleistocene cooling and drying pulse was detected worldwide at about 2.5 Mya, in times of the Great Patagonian Glaciation (Rabassa, 2008). Ice covered part of the Andean region in Patagonia and part of the extra-Andean region, where extensive areas of frozen soils developed. The result would have been a displacement of the Dry Chaco region (Fig. 4B) northward, where the Chiquitano STDF nucleus exists today. Phylogeographical studies developed in *Astronium urundeuva* (Allemão) Engl. and *Anadenanthera colubrina* (Vell.) Brennan showed west and east groupings of populations located at sites that are comparable to our study sites (see fig. 2 in Caetano *et al.*, 2008 and fig. 2 in Barrandeguy *et al.*, 2016).

The Pleistocene Arc Theory (PAT) suggests that present-day disjunct fragments of dry forests in central tropical South America provide evidence of a previously more continuous distribution during the Pleistocene that has been disrupted by dry-cold vs. humid-warm climatic cycles. Prado & Gibbs (1993) hypothesized that the increasing temperatures and precipitation over the last glacial–Holocene transition caused contraction and fragmentation of the Pleistocene dry forest arc into the present-day disjunct distribution of dry forest nuclei. We propose that, in contrast to expectation, the SDTFs in their southernmost distribution have been fragmented during glaciation events due to advance of the Chaco region from the south. This agrees with Pennington *et al.* (2004), who postulate that Chiquitano nucleus in Bolivia would have moved north rather than expanded during periods of glaciation.

Population genetics studies remain very scarce and cover a relatively small part of the total distribution range of SDTF. Ecological niche modelling developed by Werneck *et al.* (2011) and Turchetto-Zolet *et al.* (2016) provides evidence that the SDTFs were even more disjunct during the LGM. Finally, areas of presumed long-term stability are identified and confirmed (Caatinga, Misiones, Piedmont and Chiquitano region), and possibly acted as current and historical refugia (Werneck *et al.*, 2011). By contrast,

the demographic history of *Tabebuia impetiginosa* (Mart. ex DC.) Standl. showed that the pattern of diversity found for the species is in agreement with distribution expansion during the LGM (Collevatti *et al.*, 2012). This species is distributed in the northern relics of SDTF in South America and Mesoamerica.

The results obtained here contribute to the discussion regarding the PAT. Thus, the response of the SDTF to glaciation events could have depended on the STDF nucleus: northern nuclei have expanded over the Amazon forest, whereas the southernmost nuclei have undergone fragmentation because of increased area of drier regions, and moved their range north. Further studies in representative species from this biome are needed to evaluate the response to climatic change and to elaborate upon conservation actions. These forests are priority areas for conservation because they are the home of wild ancestors of many important crop species.

REFERENCES

Albrecht E, Zhang D, Saftner RA, Stommel JR. 2012. Genetic diversity and population structure of *Capsicum baccatum* genetic resources. *Genetic Resources and Crop Evolution* **59:** 517–538.

Avise JC. 2000. *Phylogeography: the history and formation of species.* Cambridge: Harvard University Press.

Banda K, Delgado-Salinas A, Dexter KG, Linares-Palomino R, Oliveira-Filho A, Prado D, Pullan M, Quintana C, Riina R, Rodríguez GM, Weintritt J, Acevedo-Rodríguez P, Adarve J, Álvarez E, Aranguren A, Arteaga JC, Aymard G, Castaño A, Ceballos-Mago N, Cogollo A, Cuadros H, Delgado F, Devia W, Dueñas H, Fajardo L, Fernández A, Fernández MA, Franklin J,

Freid EH, Galetti LA, Gonto R, González R, Graveson R, Helmer EH, Idárraga A, López R, Marcano-Vega H, Martínez OG, Maturo HM, McDonald M, McLaren K, Melo O, Mijares F, Mogni V, Molina D, Moreno N, Nassar JM, Neves DM, Oakley LJ, Oatham M, Olvera-Luna AR, Pezzini FF, Reyes Dominguez OJ, Ríos ME, Rivera O, Rodríguez N, Rojas A, Särkinen T, Sánchez R, Smith M, Vargas C, Villanueva B, Pennington T. 2016. Plant diversity patterns in neotropical dry forests and their conservation implications. *Science* **353:** 1383–1387.

Bandelt HJ, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16:** 37–48.

Barboza GE, Agra MF, Romero MV, Scaldaferro MA, Moscone EA. 2011. New endemic species of *Capsicum* (Solanaceae) from the Brazilian caatinga: comparison with the re-circumscribed *C. parvifolium*. *Systematic Botany* **36:** 768–781.

Barboza GE. 2013. *Solanaceae*. IBODA-IMBIV, ed. Buenos Aires: CONICET.

Barrandeguy ME, Prado DE, Goncalves AL, García MV. 2016. Demografía histórica de *Anadenanthera colubrina* var. *cebil* (Leguminosae) en Argentina. *Boletín de la Sociedad Argentina de Botánica* **51:** 689–703.

Bosland PW. 1996. *Capsicums*: innovative uses of an ancient crop. In: Janick J. (ed.) *Progress in new crops*. Arlington: ASHS Press, 479–487.

Brown CH, Clement CR, Epps P, Luedeling E, Wichmann S. 2013. The paleobiolinguistics of domesticated chili pepper (*Capsicum* spp.). *Ethnobiology Letters* **4:** 1–11.

Cabrera AL. 1971. Fitogeografía de la República Argentina. *Boletín de la Sociedad Argentina de Botánica* **14:** 1–2.

Caetano S, Prado D, Pennington RT, Beck S, Oliveira-Filho A, Spichiger R, Naciri Y. 2008. The history of Seasonally Dry Tropical Forests in eastern South America: inferences from the genetic structure of the tree *Astronium urundeuva* (Anacardiaceae). *Molecular Ecology* **17:** 3147–3159.

Carrizo García C, Barfuss MHJ, Sehr EM, Barboza GE, Samuel R, Moscone EA, Ehrendorfer F. 2016. Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae). *Annals of Botany* **118:** 35–51.

Chacón MI, Pickersgill SB, Debouck DG. 2005. Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theoretical and Applied Genetics* **110:** 432–444.

Chiou KL, Hastorf CA, Bonavia D, Dillehay TD. 2014. Documenting cultural selection ressure changes on chile pepper (*Capsicum baccatum* L.) seed size through time in coastal Peru (7600 B.P.–present). *Economic Botany* **68:** 190–202.

Clement CR, Rodrigues DP, Alves-Pereira A, Santos Mühlen G, de Cristo-Araújo M, Ambrósio Moreira P, Lins J, Maciel Reis V. 2016. Crop domestication in the upper Madeira River basin. *Ciências Humanas, Belém* **11:** 193–205.

Collevatti RG, Terribile LC, Lima-Ribeiro MS, Nabout JOC, Oliveira G, Rangel TF, Rabelo SG, Diniz-Filho JAF. 2012. A coupled phylogeographical and species distribution modelling approach recovers the demographical history of a Neotropical seasonally dry forest tree species. *Molecular Ecology* **21:** 5845–5863.

Crandall KA, Templeton AR. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* **134:** 959–969.

Daneck H, Fér T, Marhold K. 2016. Glacial survival in northern refugia? Phylogeography of the temperate shrub *Rosa pendulina* L. (Rosaceae): AFLP vs. chloroplast DNA variation. *Biological Journal of the Linnean Society* **119:** 704–718.

D'Arcy WG, Eshbaugh WH. 1974. New World peppers (*Capsicum*–Solanaceae) north of Colombia: a resume. *Baileya* **19:** 93–105.

Djian-Caporalino C, Lefebvre V, Sage-Daubeze AM, Palloix A. 2007. *Capsicum*. In: Singh R. J., ed. *Genetic resources, chromosome engineering, and crop improvement: vegetable crops*, *Vol.* **3**. Boca Raton: CRC Press, 185–243.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

Eshbaugh WH. 1993. History and exploitation of a serendipitous new crop discovery. In: Janick J., Simon J. E., eds. *New crops*. New York: Wiley, 132–139.

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10:** 564–567.

Grabiele M, Chalup L, Robledo G, Seijo G. 2012. Genetic and geographic origin of domesticated peanut as evidenced by 5S rDNA and chloroplast DNA sequences. *Plant Systematics and Evolution* **298:** 1151–1165.

Gross BL, Olsen KM. 2010. Genetic perspectives on crop domestication. *Trends in Plant Science* **15:** 529–537.

Hilgert NI. 1999. Las plantas comestibles en un sector de las yungas meridionales (Argentina). *Anales de los Jardines Botánicos* **57:** 117–138.

Ibiza VP, Blanca J, Cañizares J, Nuez F. 2011. Taxonomy and genetic diversity of domesticated *Capsicum* species in the Andean region. *Genetics Resources and Crop Evolution* **59:** 1077–1088.

Keller HA. 2012. El origen y la decadencia de los cultivos guaraníes, un relato mítico de los avá chiripá de Misiones, Argentina. *Bonplandia* **21:** 27–44.

Larson G, Piperno DR, Allaby RG, Purugganan MD, Andersson L, Arroyo-Kalin M, Barton L, Climer Vigueira C, Denham T, Dobney K, Doust AN, Gepts P, Gilbert MTP, Gremillion KJ, Lucas L, Lukens L, Marshall FB, Olsen KM, Pires JC, Richerson PJ, De Casas RR, Sanjur OI, Thomas MG, Fuller DQ. 2014. Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences* **111:** 6139–6146.

Lefebvre V, Palloix A, Rives M. 1993. Nuclear RFLP between pepper cultivars (*Capsicum annuum* L.). *Euphytica* **71:** 189–199.

Londo JP, Chiang Y-C, Hung K-H, Chiang T-Y, Schaa BA. 2007. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated

rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences USA* **103:** 9578–9583.

McLeod MJ, Guttman SI, Eshbaugh WH, Rayle LE. 1983. An electrophoretic study of evolution in *Capsicum* (Solanaceae). *Evolution* **37:** 562–574.

Mogni VY, Oakley LJ, Maturo HM, Galetti LA, Prado DE. 2015. Biogeografía y florística de los Bosques Secos Estacionales Neotropicales (BSEN). *OKARA: Geografia em debate* **9:** 275–296.

Morrone JJ. 2014. Biogeographical regionalisation of the Neotropical region. *Zootaxa* **3782:** 1–110.

Moscone EA, Scaldaferro MA, Grabiele M, Cecchini NM, Sanchez García Y, Jarret R, Daviña JR, Ducasse DA, Barboza GE, Ehrendorfer F. 2007. The evolution of chili peppers (*Capsicum*–Solanaceae): a cytogenetic perspective. Acta Horticulturae **745:** 137–169.

Nylander JA. 2004. *MrModelTest*. Program distributed by the author. Uppsala: Evolutionary Biology Centre, Uppsala University.

Paran I, Aftergoot E, Shifriss C. 1998. Variation in *Capsicum annuum* revealed by RAPD and AFLP markers. *Euphytica* **99:** 167–173.

Pennington RT, Prado DA, Pendry C. 2000. Neotropical seasonally dry forests and Pleistocene vegetation changes. *Journal of Biogeography* **27:** 261–273.

Pennington RT, Lavin M, Prado DE, Pendry CA, Pell SK, Butterworth CA. 2004. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. *Philosophical Transactions of the Royal Society of London B* **359:** 515–537.

Peralta IE, Makuch M, García Lampasona S, Occhiuto PN, Asprelli PD, Lorello IM, Togno L. 2008. *Catálogo de poblaciones criollas de pimiento, tomate y zapallo colectadas en valles andinos de la Argentina*. Editorial INTA, Mendoza, Argentina.

Perry DR, Zarrillo S, Holst I, Pearsall DM, Piperno DR, Berman MJ, Cooke RG, Rademaker K, Ranere AJ, Raymond JS, Sandweiss DH, Scaramelli F, Tarble K, Zeidler JA. 2007. Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. *Science* **315:** 986–988.

Pickersgill B. 1984. Migration of chili peppers, *Capsicum* spp, in the Americas. In: Store P, ed. *Pré-Columbian plant migration*. Papers of the Peabody Museum of Archaelogy and Ethnology **76**. Cambridge: Harvard University Press, 105–123.

Pickersgill B. 1997. Genetic resources and breeding of *Capsicum* spp. *Euphytica* **96:** 129–133.

Piperno DR. 2011. The origins of plant cultivation and domestication in the New World tropics: patterns, process, and new developments. *Current Anthropology* **52:** S453–S470.

Prado DE. 2000. Seasonally dry forests of tropical south America: from forgotten ecosystems to a new phytogeographic unit. *Edinburgh Journal of Botany* **57:** 437–461.

Prado DE, Gibbs PE. 1993. Patterns of species distributions in the dry seasonal forests of South America. *Annals of the Missouri Botanical Garden* **80:** 902–927.

Rabassa J. 2008. Late Cenozoic glaciations in Patagonia and Tierra del Fuego. In: Rabassa J. ed. *Developments in Quaternary Sciences*. Amsterdam: Elsevier B.V. 151–204.

Rozas J, Sánchez-Del Barrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19:** 2496–2497.

Särkinen T, Bohs L, Olmstead RG, Knapp S. 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC Evolutionary Biology* **13:** 214.

Scaldaferro MA, Seijo GJ, Acosta MC, Barboza GE, Ducasse DA, Moscone EA. 2006. Genomic characterization of the germplasm in peppers (*Capsicum*–Solanaceae) by fluorescent *in situ* hybridization. *Plant Science* **43:** 291–297.

Scaldaferro MA, Grabiele M, Moscone EA. 2013. Heterochromatin type, amount and distribution in wild species of chili peppers (*Capsicum*-Solanaceae). *Genetics Resources and Crop Evolution* **60:** 693–709.

Serrano-Serrano ML, Hernández-Torres J, Castillo-Villamizar G, Debouck DG, Chacón Sánchez MI. 2010. Gene pools in wild Lima bean (*Phaseolus lunatus* L.) from the Americas: evidences for an Andean origin and past migrations. *Molecular Phylogenetics and Evolution* **54:** 76–87.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94:** 275–288.

Simmons MP, Ochoterena H. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49:** 369–381.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725–2729.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C. 2013. Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology* **22:** 1193–1213.

Turchetto-Zolet AC, Salgueiro F, Turchetto C, Cruz F, Veto NM, Barros MJF, Segatto ALA, Freitas LB, Margis R. 2016. Phylogeography and ecological niche modelling in *Eugenia uniflora* (Myrtaceae) suggest distinct vegetational responses to climate change between the southern and the northern Atlantic Forest. *Botanical Journal of the Linnean Society* **182:** 670–688.

Van Zonneveld M, Ramirez M, Williams DE, Petz M, Meckelmann S, Avila T, Bejarano C, Ríos L, Peña K, Jäger M, Libreros D, Amaya K, Scheldeman X. 2015. Screening genetic resources of *Capsicum* peppers in their primary center of diversity in Bolivia and Peru. *PLoS ONE* **10:** e0134663.

Werneck FP, Costa GC, Colli GR, Prado DE, Sites Jr JW. 2011. Revisiting the historical distribution of Seasonally Dry Tropical Forests: new insights based on palaeodistribution modelling and palynological evidence. *Global Ecology and Biogeography* **20:** 272–288.

Zhang HX, Zhang ML, Sanderson SC. 2016. Spatial genetic structure of forest and xerophytic plant species in arid Eastern Central Asia: insights from comparative phylogeography and ecological niche modelling. *Biological Journal of the Linnean Society* **120:** 612–625.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Table S1. Variable sites of the aligned chloroplast DNA sequences from the *rpl32F-trnL* region for the 18 haplotypes in *Capsicum baccatum*.
Table S2. Variable sites of the aligned chloroplast DNA sequences from the *trnQ-rps16* region for the 18 haplotypes in *Capsicum baccatum*.
Database S1. Sequence dataset used in the phylogeographical analysis.