4-4-2017

Jujuy Province (NW Argentina): STR Markers Unveil Microgeographic Differentiation Over a Steep Mountainous Landscape

Luis Gómez-Pérez
Universidad del Pais Vasco (UPV/EHU), luisgp1982@gmail.com

Miguel A. Alfonso-Sánchez
Universidad del Pais Vasco (UPV/EHU), miguel.alfonso@ehu.es

José Edgardo Dipierri
Instituto de Biología de la Altura, Universidad Nacional de Jujuy, Argentina, dipierri@inbial.unju.edu.ar

Jose A. Pena
Universidad del Pais Vasco (UPV/EHU), joseangel.pena@ehu.es

Recommended Citation
Gómez-Pérez, Luis; Alfonso-Sánchez, Miguel A.; Dipierri, José Edgardo; and Pena, Jose A., "Jujuy Province (NW Argentina): STR Markers Unveil Microgeographic Differentiation Over a Steep Mountainous Landscape" (2017). *Human Biology Open Access Pre-Prints*. 99.
http://digitalcommons.wayne.edu/humbiol_preprints/99

Jujuy province (NW Argentina): STR markers unveil microgeographic differentiation over a steep mountainous landscape

Luis Gómez-Pérez [1#], Miguel A. Alfonso-Sánchez [1#], José E. Dipierri [2], José A. Peña [1]*

1. Departamento de Genética y Antropología Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco (UPV/EHU), 48080 Bilbao, Spain

2. Instituto de Biología de la Altura, Universidad Nacional de Jujuy, 4600 San Salvador de Jujuy, Argentina

 # These authors contributed equally to this work

*** Address for correspondence**:

Dr. José A. Peña
Departamento de Genética, Antropología Física y Fisiología Animal
Universidad del País Vasco (UPV/EHU)
Apartado 644, 48080 Bilbao, Spain
e-mail: joseangel.pena@ehu.es

Issue: 88.3

Key words: microevolution, landscape genetics, genetic drift, gene flow, Jujuy, STRs

Abstract

This study explores potential signals of microdifferentiation on the gene pool of three high-altitude populations from Jujuy province (NW Argentina) using highly polymorphic markers. These human communities are characterized by extreme living conditions and very low population densities owing to considerable height above sea level and steep orography. A set of autosomal STRs located at chromosome 6 (6p21.3) was typed in samples from Quebrada Baja (~2,500 m), Quebrada Alta (~ 3,300 m), and Puna (> 3,500 m). Genetic diversity was estimated through the observed (Ho) and expected (He) heterozygosities, and the haplotype diversity. Analyses of the molecular variance (AMOVA) and population differentiation tests based on allele and haplotype frequencies were performed to assess genetic heterogeneity among subgroups. No deviation from HWE expectations was detected for each separate subpopulation; yet, significant departures were detected in the analysis considering the whole area (D6S2792 and D6S105 loci). Overall, genetic diversity showed a decreasing trend as the altitude increases. Thus, allele and haplotype frequencies showed the most significant differences between Puna and Quebrada Baja, which are the populations sited at the edges of the altitude range. The trend to the reduction of the heterozygosity with altitude proves to be compatible with historical patterns of colonization, interregional migration trends, population density, and genetic admixture. The main consequence of the complex mountainous landscape of Jujuy would be an imbalance in the interplay gene flow-genetic drift favoring the latter. The combined effect of restricted gene flow with intense genetic drift would have promoted local genetic differentiation between Jujuy highlands' subpopulations, leading to spatial patterning of the allele frequencies not entirely attributable to geographic distance. Our findings corroborate the effectiveness of STRs to identify microevolutionary changes.

Genomic markers of quite different nature have been used to identify microevolutionary changes in human populations. Among them, polymorphisms from noncoding DNA regions are assumed to be selectively neutral, so their evolution in populations can be mostly explained by the interplay between gene flow and genetic drift. Because of their hypervariable nature, microsatellite or short tandem repeat (STR) polymorphisms have been extensively analyzed in various fields of population genetics. Microsatellites feature comparatively high mutation rates that range between 3.3×10^{-4} per locus per generation (Forster et *al.*, 2000) and 15.2×10^{-4} per locus per generation (Zhivotovsky et *al.*, 2004), considering generations of 25 years. The high mutation rate of human STRs may thus provide valuable genetic information for the elucidation of human population history and for studying genetic substructuring among local populations (Agrawal and Khan, 2005; Shepard et *al.*, 2005; Gómez-Pérez et *al.*, 2011a, b; Capocasa et *al.*, 2014).

The timing of early human dispersal in South America is a key issue in studies of human evolution. The oldest documented human presence in the province of Jujuy (NW Argentina) predates the Neolithic revolution and has been estimated in at least 10.500 years (Prates et *al*., 2013). Some evidence suggests the existence of preincaic ethnic groups in the territory, probably related to other Andean communities from the genetic and cultural viewpoint (Lorandi, 1992). The annexation of these ethnic groups to the Inca empire was relatively pacific (Krapovickas, 1981-82), contrasting with what happened in regions located south of Jujuy province (Lorandi, 1988). Some centuries later, the arrival of the Spanish colonists to Jujuy provoked no great conflict nor massive population displacements in the region, whereas the conquest of southern neighboring provinces took more than one hundred years (Lorandi, 1992). There were patent differences in the settlement process of the Spanish colonists and their endowments of African slaves along the territory of Jujuy mainly during the 16th century, with a significant presence in the lowest regions and virtually negligible in the high-altitude zones (Dipierri et *al*., 2000). More detailed information about geographic, demographic and historical features of Jujuy can be found in earlier papers by these authors (Dipierri *et al*., 2000; Gómez-Pérez *et al*., 2011a).

The province of Jujuy has proved to be a suitable model for exploring the impact of the interaction between the different evolutionary forces on the genetic diversity of the human communities inhabiting the region. Located in a mountainous region, Jujuy is predominantly characterized by striking variations in height above sea level among its population nuclei (Fig. 1), whose main demographic consequence would be the emergence of many partially isolated villages. Such geomorphological features would have propitiated a reduction of the homogenizing effects of the gene flow (Dipierri et *al.*, 1998, 2000; Gómez-Pérez et *al.*, 2011a), thereby promoting genetic heterogeneity in the area by genetic drift effects.

In a previous work based on polymorphic *Alu* elements, Gómez-Pérez et *al.* (2011a) provided evidence of association between altitude and spatial distribution of *Alu* insertion frequencies across Jujuy province. Thus, the five study regions of Jujuy showed disparate admixture levels of Native Americans with individuals of either European or African ancestry depending basically on height above sea level. High-altitude populations (> 2500 m) showed almost 100% of Amerindian ancestry, with a negligible genetic contribution of other ethnic groups. Genetic diversity therefore showed a decreasing trend from the high-altitude regions (Quebrada Baja, Quebrada Alta and Puna) to those at lower altitudes (Valle and Selva). Virtually no genetic heterogeneity was observed within the high-altitude populations, as expected according to their similarity in admixture levels and ethnic ancestry.

The present study was designed to refine the analysis of the genetic diversity in Jujuy by focusing attention in the three high-altitude populations of this Argentine province, which presumably would have preserved almost intact their Native American gene pool (Cardoso *et al*. 2013). To that end, we take advantage of the high level of polymorphism of STR markers, whose hypervariable nature provides a higher resolution than biallelic *Alu* markers to detect genetic signatures of recent demographic phenomena in these human communities. First, this work was intended to confirm the hypothetical lack of genetic differentiation among subpopulations of the highlands by contrasting the results generated by the highly-polymorphic microsatellite loci utilized herein with those previously reported from evolutionarily conservative markers (*Alu* repeats). Furthermore, given that high-altitude populations of Jujuy stand out by extreme living conditions and remarkably low population densities, our study was then addressed to gain insights into the microevolutionary changes that might have modelled the genetic diversity of the region, exploring the interaction between gene flow and genetic drift in shaping the gene pool of these human communities.

MATERIAL AND METHODS

Autosomal microsatellite loci (D6S2792, D6S265, D6S105 and D6S2666) located at chromosome 6 were screened in a population sample Jujuy province, NW Argentina. The fact that all microsatellite markers are located at the same genomic region of chromosome 6 (6p21.3) enables carrying out haplotype-based analyses. The usefulness of such analyses lies in that haplotypes might have potential value as lineage markers in clinical, forensic and/or evolutionary investigations (Dunn et *al.*, 2007; García-Obregón et *al.*, 2010, 2011). STRs are noncoding DNA markers and, therefore, selectively neutral. Nevertheless, their position within the Major Histocompatibility Complex (MHC) genomic region could be very helpful for detecting signals of natural selection, owing to genetic hitchhiking with adjacent genes involved in immune response (Shiina et *al.*, 2009). As a consequence of the phenomenon of genetic hitchhiking, the frequency of an evolutionarily neutral allele can be altered by selective processes operating upon contiguous, closely linked genes (Hedrick, 1982; Kaplan et *al.*, 1989).

All individuals included in the sample lived in different population centers from the province's highlands, which are characterized by their prominent height above sea level (> 2500 m, see Fig. 1). For a refined approach of the genetic diversity of the human communities settled in the highlands of Jujuy, samples were divided into three groups according to altitude: *i*) Puna (N = 41; mean altitude > 3,500 m), *ii*) Quebrada Alta (N = 31; mean altitude ~ 3,300), and *iii*) Quebrada Baja (N = 30; mean altitude ~ 2,500 m). For the sake of simplicity, the population cluster formed by these three subregions will be called hereafter 'Jujuy highlands'. Voluntary donors were asked to provide personal and family information useful to our study, such as surnames and geographical origin of their parents and grandparents, among others. The final sample consisted of 102 healthy, unrelated Native American individuals, whose ancestors (until the third generation back) were also born in Jujuy highlands.

Ethical guidelines for research with human beings were adhered to as stipulated by the institutions implicated in the research project: Universidad del País Vasco (UPV/EHU), Spain, and Universidad Nacional de Jujuy, Argentina. Blood donors gave their informed consent before inclusion in the sample. The study protocol was approved by the Institutional Review Board from the UPV/EHU.

PCR amplification and STR typing

Genomic DNA was extracted from bloodstains using the standard phenol:chloroform procedure (Sambrook et *al.*, 1989). Typing of the STR markers was performed by PCR amplification, according to methods previously described (see García-Obregón et *al.*, 2010). Positive and negative PCR controls were run to assess the quality of the amplification process. Amplified STR fragments were further detected and separated by capillary electrophoresis, using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). To estimate the size of the amplified products, an internal size standard (GeneScan TAMRA-500, Applied Biosystems) was included. Fragment sizes were determined automatically using GeneMapper Analysis software v3.7, and by comparison with the supplied allelic ladders.

Statistical analysis

Allele frequencies were calculated by direct counting. To test for Hardy-Weinberg equilibrium (HWE), a Fisher's exact probability test using the Markov-chain Monte Carlo method was conducted (Guo and Thompson, 1992). Statistical significance of HWE results was checked by estimating *P*-values after 100,000 steps in the Markov chain. Genetic variability of the analyzed STR markers was measured by calculating observed heterozygosity (H_o), number of alleles per locus, and Nei's unbiased average gene diversity (GD) across all loci for each sample collection (Nei, 1987).

Bearing in mind that all microsatellite loci are located in chromosome 6, linkage disequilibrium values (Slatkin and Excoffier, 1996) were computed for each loci pair. Haplotypes were inferred with Arlequin v3.5 software (Excoffier and Lischer 2010), using genotype data for each individual (D6S2792-D6S265-D6S105-D6S2666). Once the haplotype combinations were constructed, we obtained further data on the genetic variability of the study area by estimating haplotype frequencies and haplotype diversity (Nei and Tajima, 1981).

Departure from the patterns of polymorphism predicted by the neutral model was examined using the Ewens-Watterson test (Ewens, 1972; Watterson, 1978). Selective neutrality was tested for all microsatellites involved in the study.

Potential structuring of the genetic variation between the different subregions of Jujuy highlands was explored by comparing within population variation with among population variation through the analysis of molecular variance (AMOVA) for each locus. Subsequently, we performed an overall test to check the statistical significance of the fixation indices by combining the separate probability values for each locus through the equation $\chi^2_{[2k]} = - 2 \sum \ln p_i$, where k is the number of loci and p_i the separate probability value associated with the fixation index for each i locus (Sokal and Rohlf, 1997). For a more thorough analysis, both allele and haplotype frequencies were further utilized to perform F_{ST} pairwise comparisons (Reynolds et *al.*, 1983), and exact tests of population differentiation (Raymond and Rousset, 1995) for genotypic data with unknown gametic phase (Goudet et al., 1996). All statistical analyses were carried out with the Arlequin v3.5 program.

RESULTS

Allele frequencies of the STRs, both for the three subpopulations (Quebrada Baja, Quebrada Alta and Puna) and for the whole study region (Jujuy highlands) are listed in Table 1. A total of 10 alleles for the D6S2792 locus, 14 alleles for D6S105, 7 alleles for D6S265 and 6 for D6S2666 were identified in the sample surveyed. Overall, the more frequent alleles per loci coincided in the three subpopulations. These were the cases of allele 115 of D6S2792, allele 120 of D6S105, and allele 124 of D6S265. Exception to the above was the D6S2666 locus, with allele 170 being the most frequent in Quebrada Baja and Quebrada Alta and allele 144 predominating in Puna.

The Fisher's exact probability test revealed no significant deviation from the expected allele frequencies when HWE was checked for each single subpopulation (Table 2). In this analysis, results indicated that only D6S105 in Quebrada Alta was very close to the statistical significance after applying the Bonferroni correction for multiple testing. However, when the same analysis was performed considering the whole study area (Jujuy highlands) we found significant departures from HWE for D6S2792 and D6S105 microsatellite loci. In general terms, the lowest values of genetic diversity in Jujuy highlands appeared in Puna, which featured a mean value of gene diversity (GD, considering all STR loci) of 0.68. The mean GD values estimated for Quebrada Alta and Quebrada Baja proved to be 0.70 and 0.72, respectively. Such results disclose a perceptible decreasing trend of the genetic diversity from Puna (> 3,500 m) to Quebrada Baja (~ 2,500 m), which are the populations located at the extremes of altitude variation.

To evaluate potential genetic hitchhiking with nearby coding genes involved in the immune response, departure from selective neutrality was measured through the Ewens–Watterson test (data not shown). No evidence was found to reject the null hypothesis of a random spatial distribution of the allele frequencies. Thus, the lack of statistically significant associations ($P > 0.05$ in all cases) was revealing of no signals of selective pressures acting on the genomic region examined, although such a possibility should not be ruled out given the nature (noncoding markers) and the limited number of the STR loci analyzed. Obviously, much more data would be necessary to help in elucidating this topic.

Taking into account that all the analyzed loci share the same genomic region (6p21.3), pairwise linkage disequilibrium analysis was performed to assess the magnitude of the association between microsatellites. No more than 50% of the associations proved to be statistically significant (Supporting Information Table 1). These were the cases of D6S2792 and D6S105 loci ($P = 0.013$), D6S2792 and D6S2666 ($P = 0.006$), and mainly between D6S265 and D6S105 ($P < 0.001$). Results confirmed that the selected marker set is not fully linked, so that haplotype-based analyses could provide complementary information to that obtained from individual genes.

Spatial structuring of the genetic diversity was explored by hierarchical AMOVA to ascertain maximum genetic variance between groups (F_{ST}). AMOVA analyses were performed based on both STR allele and haplotype frequencies (Supporting Information Table 2). As can be noticed, three STR loci (D6S2792, D6S105 and D6S2666) contribute significantly to the geographic patterning of the genetic heterogeneity. The only exception was D6S265 ($F_{ST} = 0.0054$, $P = 0.8328$). Likewise, the overall significance test for F_{ST} (combining the separate probability values for each locus) was statistically significant ($P = 0.0042$), which indicates a substantial genetic substructuring in Jujuy highlands. As logically expected, AMOVA results based on haplotype frequencies were also revealing of spatial patterning of the genetic diversity in Jujuy highlands ($F_{ST} = 0.0225$, $P < 0.001$).

To refine the analysis of the spatial structuring of the genetic variation in the study area, population differentiation tests (F_{ST} pairwise tests) were further performed using allele and haplotype frequencies as well (Table 3). Several statistical significant differences were observed among the three subpopulations. Frequency differences for both the individual STR loci and the haplotype combinations were mostly detected between Puna and Quebrada Baja (D6S2792, D6S105, D6S2666 and the four-loci haplotype). As mentioned above, these subpopulations occupy the extremes of the variation in height above sea level. In addition, two microsatellite markers (D6S105 and D6S2666) were found to show significant frequency differences between Puna and Quebrada Alta, whereas only D6S2666 differed between Quebrada Baja and Quebrada Alta. Likewise, results derived from the exact tests of population differentiation coincided, for the most part, with Fst pairwise comparisons. The only exception was the locus D6S105, whose allele frequencies also showed significant differences ($P = 0.0053$) between the subpopulations of Quebrada Baja and Quebrada Alta.

DISCUSSION

Geographic distance is considered a crucial factor in human population structure, because of its limiting effects on migratory movements (see Jorde, 1980). However, such a simple model certainly excludes a number of parameters that could have an impact in the shaping of the genetic structure of a population. The developing field of landscape genetics takes also into account other ecological determinants influencing gene flow and, therefore, genetic variation, mostly associated with geographic landforms like rivers, lagoons and mountains, among others (Storfer et *al.*, 2007; Coia et *al.*, 2012; Messina et *al.*, 2015). Recent studies indicate that the spatial heterogeneity of the environment is key in explaining gene flow patterns and population differentiation processes even at a microgeographic scale (Manel et *al.*, 2003; McRae and Beier, 2007; Manel and Holderegger, 2013; Duforet-Frebourg and Blum, 2014).

An earlier study on the genetic diversity of Jujuy province based on polymorphic *Alu* insertions found that Native American ancestry increased with altitude above sea level since immigrants tended to settle at lower altitudes, leading to more genetic admixture. Populations from the highest altitudes (the so-called Jujuy highlands) featured almost 100% of Native American ancestry, whereas the gene pool of the lowest population nuclei showed both European and African contribution. In agreement with the above findings, these authors also identified an increasing trend of genetic diversity from Jujuy highlands' to the human communities settled at lower-altitude regions (Valle and Selva). Yet, *Alu* markers failed to detect local genetic differentiation among Jujuy highlands' subpopulations (for more information, see Gómez-Pérez et *al.*, 2011a).

Results of this study confirm the effectiveness of the resolving power of the STR markers in identifying population differentiation processes even at a microgeographic level, as can be deduced from data presented in Table 3 (STR markers). Conflicting with the hypothetical lack of genetic heterogeneity in Jujuy highlands derived from the analysis of *Alu* insertions (Gómez-Pérez *et al.*, 2011a), geographic distribution of the STR polymorphism suggests existence of genetic structuring among the three high-altitude human groups. Deviations from HWE (see Table 2) as well as results from AMOVAs, F_{ST} pairwise comparisons, and exact tests of population differentiation could be reflecting the effect of geographical barriers to random mating between the subpopulations of Puna (N=41), Quebrada Alta (N=31), and Quebrada Baja (N=30). In view of the relatively short linear geographic distances between the population nuclei of Quebrada Baja and Quebrada Alta (minimum ~ 5.0 km, maximum: ~ 35.0 km), substantial limitation of gene flow in Jujuy highlands – and therefore, of genetic heterogeneity – identified herein might be basically explained in terms of landscape characteristics, notably the steep orography predominating in the region.

As expected, genetic microdifferentiation was particularly conspicuous between those populations located at the edges of the altitude range (Quebrada Baja: ~ 2500 m; Puna: > 3500 m). Puna is the more remote region of Jujuy highlands. This factor, jointly with the hostile living conditions associated with extreme altitudes have undoubtedly reinforced its degree of isolation. In fact, population density in Puna is 1.34 inh/km^2 (INDEC, 2010), the lowest in Jujuy highlands (Quebrada Alta: 2.95 inh/km^2; Quebrada Baja: 5.64 inh/km^2). Obviously, a constant small population size would have propitiated a reduction of genetic diversity by severe genetic drift events. In addition, Puna's steep geomorphology would have substantially strengthened the partial population isolation generated, in principle, by geographic distance. Similar patterns of microevolutionary changes by lack of panmixis due to geographical barriers has been reported for several model and non-model organisms, suggesting that the influence of landscape on local dispersal is a leading evolutionary force (Ramachandran et *al.*, 2005).

Another scenario to account for the low population density in Jujuy highlands could be connected with socioeconomic factors, and specifically with a marked limitation in employment opportunities. Overall, lands of Jujuy's highest regions are far from suitable for agricultural and cattle activities owing to its dryness, and therefore, low fertility. The mainstay of the economy has traditionally been the herding of llamas, and to a lesser extent, of alpacas and sheep. All of these animals produce fleece that is useful for weaving and plaiting. This socioeconomic picture was captured in the statistics of the 1859 census, which revealed that the vast majority of the Puna population was occupied in the abovementioned activities (Gil, 2000-2001). For these reasons, Jujuy highlands (and particularly the Puna region) would have been barely attractive for both the European emigrants and the African slaves and their descendants, a fact that would have had an impact in demographic size and population isolation of these human communities. Socioeconomic characteristics then became a limiting factor for gene flow, and accordingly, a factor that would have promoted genetic drift episodes in the region.

Summarizing, the main evolutionary consequence of the altitude and, especially, of the complex geomorphology of Jujuy highlands would be an imbalance in the interplay between gene flow and genetic drift favoring the impact of this latter evolutionary force. On the one hand, extreme living conditions and poor economic opportunities typical of great altitudes limit the inmigrants arrival, thus reducing the gene flow effects and the genetic diversity within subpopulations. On the other hand, population isolation on account of a very steep local orography, together with strikingly low population densities would have promoted intense genetic drift processes. The main consequence of the combined effect of a restricted gene flow with a significant impact of genetic drift would therefore be local genetic differentiation between Jujuy highlands' subpopulations. The loss of genetic diversity, and therefore, the most significant genetic differentiation is mainly detectable in Puna, the human community located at a highest altitude. Finally, the hypervariable nature of microsatellite markers has proved to be highly advantageous for the detection of recent microevolutionary signatures in the gene pool of human populations, as postulated in previous works (Shepard and Herrera, 2006; García-Obregón et *al.*, 2007, 2010, Ray et *al.*, 2010). In contrast, *Alu* repeats are unique evolutionary events (Cordaux and Batzer, 2009) and thus much more conservative, since they are not exposed to the fluctuations of recurrent mutational events. In this way, polymorphic *Alu* elements could better detect the divergent ancestry of a set of geographically related populations (Terreros et *al.*, 2009; Gómez-Pérez et *al.*, 2011a). The combined examination of *Alu* and STR loci may then provide complementary perspectives in the reconstruction of the demographic and evolutionary histories of partially isolated human groups.

Acknowledgments: L.G-.P has a doctoral fellowship from the University of the Basque Country (UPV/EHU). Grant sponsors: ‘Programa de Subvención a Grupos Consolidados, Departamento de Educación, Universidades e Investigación’ of the Basque Government (Contract grant number IT-424-07). The authors are particularly indebted to all voluntary donors from Jujuy province (Argentina), who generously cooperated to the development of this research.

Received 21 March 2016; accepted 1 August, 2016.

LITERATURE CITED

Agrawal, S., and F. Khan. 2005. Reconstructing recent human phylogenies with forensic STR loci: a statistical approach. *BMC Genet.* 6:47.

Capocasa, M., L. Taglioli, P. Anagnostou et al. 2014. Determinants of marital behaviour in five Apennine communities of Central Italy inferred by surname analysis, repeated pairs and kinship estimates. *Homo* 65:64-74.

Cardoso, S., L. Palencia-Madrid, L. Valverde et al. 2013. Mitochondrial DNA control region data reveal high prevalence of Native American lineages in Jujuy province, NW Argentina. *Forensic Sci. Int.- Genet.* 7:e52-55.

Coia, V., I. Boschi, F. Trombetta et al. 2012. Evidence of high genetic variation among linguistically diverse populations on a micro-geographic scale: a case study of the Italian Alps. *J. Hum. Genet.* 57:254-260.

Cordaux, R., and M. A. Batzer. 2009. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* 10:691-703.

Dipierri, J. E., E. Alfaro, V. L. Martínez-Marignac et al. 1998. Paternal directional mating in two Amerindian subpopulations located at different altitudes in northwestern Argentina. *Hum. Biol.* 70:1001-1010.

Dipierri, J. E., E. Alfaro, J. A. Peña et al. 2000. GM, KM immunoglobulin allotypes and other serum Genetic markers (HP, GC, PI and TF) among South American populations living at different altitudes (Jujuy Province, Argentina): Admixture estimates. *Hum. Biol.* 72:305–319.

Duforet-Frebourg, N., and M. G. Blum. 2014. Nonstationary patterns of isolation-by-distance: inferring measures of local genetic differentiation with Bayesian kriging. *Evolution* 68:1110-1123.

Dunn, D. S., M. K. Choy, M. E. Phipps et al. 2007. The distribution of major histocompatibility complex class I polymorphic Alu insertions and their associations with HLA alleles in a Chinese population from Malaysia. *Tissue Antigens* 70:136-143.

Ewens, W. 1972. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* 3:87-112.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population Genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564-567.

Forster, P., A. Röhl, P. Lünnemann et al. 2000. A short tandem repeat-based phylogeny for the human Y chromosome. *Am. J. Hum. Genet.* 67:182-196.

García-Obregón, S., M. A. Alfonso-Sánchez, L. Gómez-Pérez L et al. 2011. Microsatellites and Alu elements from the human MHC in Valencia (Spain): analysis of genetic relationships and linkage disequilibrium. *Int. J. Immunogenet.* 38:483-491.

García-Obregón, S., M. A. Alfonso-Sánchez, A. M. Pérez-Miranda et al. 2007. Polymorphic Alu insertions and the genetic structure of Iberian Basques. *J. Hum. Genet.* 52:317-327.

García-Obregón, S., M. A. Alfonso-Sánchez, A. M. Pérez-Miranda, et al. 2010. Genetic variability in autochthonous Basques from Guipuzcoa: a view from MHC microsatellites. *Int. J. Immunogenet.* 37:279-287.

Gil, R. 2000-2001. Género y etnicidad en la "Puna de Jujuy". Población & Sociedad 8/9: 41-70.

Gómez-Pérez, L., M. A. Alfonso-Sánchez, J. E. Dipierri et al. 2011a. Microevolutionary processes due to landscape features in the province of Jujuy (Argentina). *Am. J. Hum. Biol.* 23:177-184.

Gómez-Pérez, L., M. A. Alfonso-Sánchez, D. Sánchez et al. 2011b. Alu polymorphisms in the Waorani tribe from the Ecuadorian Amazon reflect the effects of isolation and genetic drift. *Am. J. Hum. Biol*. 23:790-795.

Goudet, J., M. Raymond, T. de Meeüs et al. 1996. Testing differentiation in diploid populations. *Genetics* 144:1933-1940.

Guo, S. W., and E. A. Thompson. 1992. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics* 48:361–372.

Hedrick, P. W. 1982. Genetic hitchhiking: a new factor in evolution? *BioScience* 32:845-853.

INDEC. 2010. Censo Nacional de población, hogares y viviendas 2010. Censo del Bicentenario. Available at: http://www.indec.gov.ar/.

Jorde, L. B. 1980. The Genetic Structure of Subdivided Human Populations, a review. En: Mielke J, Crawford MD, editors. Current developments in Anthropological Genetics, Vol 1. Theory and Methods. New York, USA: Plenum Press. p 135-208.

Kaplan, N. L., R. R.Hudson, and C. H. Langley. 1989. The "hitchhiking effect" revisited. *Genetics* 123:887-899.

Krapovickas, P. 1981-82. Hallazgos incaicos en Tilcara y Yacoraite (Una reinterpretación). Relaciones de la Sociedad Argentina de Antropología 14:67-80.

Lorandi, A. M. 1988. Los diaguitas y el Tawantinsuyu: Una hipótesis de conflicto. In (Dillehay TD, Netherly P, eds.) La frontera del Estado inca. Quito: Editorial Abya-Yala, p. 235-259.

Lorandi, A. M. 1992. El mestizaje interétnico en el noroeste argentino. *Senri Ethnological Studies* 33:133-166.

Manel, S., and R. Holderegger. 2013. Ten years of landscape genetics. *Trends Ecol. Evol.* 28:614-621.

Manel, S., M. K. Schwartz, G. Luikart et al. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol. Evol.* 18:189-197.

McRae, B., and P. Beier. 2007. Circuit theory predicts gene flow in plant and animal populations. *Proc. Natl. Acad. Sci. USA* 104:19885-19890.

Messina, F., A. Finocchio, M. F. Rolfo et al. 2015. Traces of forgotten historical events in mountain communities in Central Italy: A genetic insight. *Am. J. Hum. Biol.* 27:508-519.

Nei, M., and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. *Genetics* 97:145–163.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Prates, L., G. Politis, and J. Steele. 2013. Radiocarbon chronology of the early human occupation of Argentina. *Quatern. Int.* 301:104-122.

Ramachandran, S., O. Deshpande, C. C. Roseman et al. 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc. Natl. Acad. Sci. USA* 102:15942-15947.

Raymond, M., and F. Rousset. 1995. An exact test for population differentiation. *Evolution* 49:1280-1283.

Ray, N., D. Wegmann, N. J. Fagundes et al. 2010. A statistical evaluation of models for the initial settlement of the american continent emphasizes the importance of gene flow with Asia. *Mol. Biol. Evol.* 27:337-345.

Reynolds, J., B. S. Weir, and C. C. Cockerham. 1983. Estimation of the coancestry coefficient: basis for a short-term Genetic distance. *Genetics* 105:767-779.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual. 2nd ed. Cold Spring Habor: Cold Spring Habor Laboratory Press.

Shepard, E. M, R. A. Chow, E. Suafo'a et al. 2005. Autosomal STR variation in five Austronesian populations. *Hum. Biol.* 77:825-851.

Shepard, E. M., and R. J. Herrera. 2006. Iranian STR variation at the fringes of biogeographical demarcation. *Forensic Sci. Int.* 158:140-148.

Shiina, T., K. Hosomichi, H. Inoko et al. 2009. The HLA genomic loci map: expression, interaction, diversity and disease. *J. Hum. Genet.* 54:15-39.

Slatkin, M., and L. Excoffier. 1996. Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity* 76:377-383.

Sokal, R. R., and F. J. Rohlf. 1997. Biometry. The principles and practice of statistics in biological research,3rd ed. New York: WH Freeman and Company.

Storfer, A., M. A. Murphy, J. S. Evans et al. 2007. Putting the "landscape" in landscape genetics. *Heredity* 98:128-142.

Terreros, M. C., M. A. Alfonso-Sánchez, G. E. Novick et al. 2009. Insights on human evolution: an analysis of Alu insertion polymorphisms. *J. Hum. Genet.* 54:603-611.

Watterson, G. A. 1978. The homozygosity test of neutrality. *Genetics* 88:405-417.

Zhivotovsky, L. A., P. A. Underhill, and C. Cinnioğlu. 2004. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am. J. Hum. Genet.* 74:50-61.

Table 1. Allele frequencies and standard error (± SE) for autosomal microsatellites from chromosome 6 in three high-altitude populations from Jujuy province (NW Argentina)

Locus	Alelle (bp)	Frequency ± SE			
		Quebrada Baja (2N = 60)	Quebrada Alta (2N = 62)	Puna (2N = 82)	Total[a] (2N = 204)
D6S2792	97	0.117 ± 0.037	0.097 ± 0.040	0.122 ± 0.036	0.113 ± 0.022
	99	0.100 ± 0.035	0.065 ± 0.036	0.098 ± 0.034	0.088 ± 0.021
	103	0.017 ± 0.016	-	-	0.005 ± 0.005
	105	0.017 ± 0.016	0.048 ± 0.025	0.012 ± 0.012	0.025 ± 0.011
	107	0.250 ± 0.049	0.210 ± 0.047	0.134 ± 0.037	0.191 ± 0.026
	109	0.150 ± 0.040	0.065 ± 0.029	0.024 ± 0.016	0.074 ± 0.017
	111	-	-	0.024 ± 0.016	0.010 ± 0.007
	115	0.333 ± 0.060	0.468 ± 0.064	0.561 ± 0.057	0.466 ± 0.037
	117	0.017 ± 0.016	0.016 ± 0.015	0.012 ± 0.012	0.015 ± 0.008
	123	-	0.032 ± 0.031	0.012 ± 0.012	0.015 ± 0.011
D6S105	108	0.233 ± 0.062	0.210 ± 0.052	0.073 ± 0.032	0.162 ± 0.029
	110	-	0.032 ± 0.031	0.024 ± 0.016	0.020 ± 0.012
	112	0.017 ± 0.016	-	-	0.005 ± 0.005
	114	-	0.048 ± 0.034	-	0.015 ± 0.011
	116	0.133 ± 0.045	0.226 ± 0.060	0.256 ± 0.050	0.211 ± 0.031
	118	0.067 ± 0.029	0.032 ± 0.021	0.073 ± 0.026	0.059 ± 0.015
	120	0.350 ± 0.056	0.355 ± 0.058	0.427 ± 0.053	0.382 ± 0.033
	122	0.033 ± 0.022	0.032 ± 0.021	-	0.020 ± 0.009
	124	0.033 ± 0.022	-	0.024 ± 0.016	0.020 ± 0.009
	126	0.017 ± 0.016	-	0.024 ± 0.016	0.015 ± 0.008

	128	0.067 ± 0.030	0.016 ± 0.015	0.024 ± 0.023	0.034 ± 0.014
	130	0.050 ± 0.026	0.016 ± 0.015	0.061 ± 0.030	0.044 ± 0.015
	132	-	0.032 ± 0.021	-	0.010 ± 0.007
	138	-	-	0.012 ± 0.012	0.005 ± 0.005
D6S265	116	0.083 ± 0.040	0.065 ± 0.029	0.024 ± 0.016	0.054 ± 0.017
	120	0.150 ± 0.051	0.210 ± 0.060	0.159 ± 0.034	0.172 ± 0.027
	122	0.050 ± 0.026	0.048 ± 0.026	0.037 ± 0.026	0.044 ± 0.015
	124	0.500 ± 0.069	0.483 ± 0.066	0.512 ± 0.057	0.500 ± 0.038
	126	0.217 ± 0.057	0.194 ± 0.051	0.244 ± 0.047	0.221 ± 0.030
	128	-	-	0.012 ± 0.012	0.005 ± 0.005
	132	-	-	0.012 ± 0.012	0.005 ± 0.005
D6S2666	144	0.233 ± 0.057	0.387 ± 0.062	0.439 ± 0.050	0.363 ± 0.034
	148	0.017 ± 0.016	-	-	0.005 ± 0.005
	164	0.083 ± 0.032	0.065 ± 0.036	0.195 ± 0.043	0.123 ± 0.024
	168	0.033 ± 0.022	0.016 ± 0.015	0.024 ± 0.016	0.025 ± 0.015
	170	0.600 ± 0.059	0.516 ± 0.065	0.342 ± 0.047	0.471 ± 0.035
	172	0.033 ± 0.032	0.016 ± 0.015	-	0.015 ± 0.011

2N, sample size in number of chromosomes analyzed

[a] Jujuy highlands (Quebrada Baja, Quebrada Alta and Puna)

Table 2. Number of alleles per locus, observed heterozygosity (Ho), gene diversity (GD) and Hardy-Weimberg equilibrium (HWE) for autosomal STRs from chromosome 6 in three high-altitude populations from Jujuy province (NW Argentina)

		Locus				
Population	Parameter	D6S2792	D6S105	D6S265	D6S2666	Mean values
Quebrada Baja	Alleles	8	10	5	6	
	Ho	0.833	0.467	0.733	0.500	0.633
	GD	0.793	0.682	0.805	0.586	0.717
	HWE	0.290	0.148	0.052	0.234	
Quebrada Alta	Alleles	8	10	5	5	
	Ho	0.613	0.613	0.548	0.452	0.557
	GD	0.728	0.785	0.689	0.589	0.698
	HWE	0.073	**0.037**	0.328	0.180	
Puna	Alleles	9	10	7	5	
	Ho	0.561	0.634	0.585	0.683	0.616
	GD	0.649	0.744	0.659	0.660	0.678
	HWE	0.138	0.121	0.094	0.583	
Total[a]	Alleles	10	14	7	6	
	Ho	0.657	0.657	0.539	0.559	0.603
	GD	0.723	0.778	0.670	0.634	0.701
	HWE	**0.009**	< **0.001**	0.058	**0.032**	

[a] Jujuy highlands (Quebrada Baja, Quebrada Alta and Puna)

Results of HWE are *P*-values generated by the Fisher's exact probability test (Guo and Thompson, 1992) after 100,000 steps in the Markov chain

Statistically significant *P*-values ($P < 0.05$) are highlighted in bold; significant *P*-values after Bonferroni correction for multiple testing (threshold value, $P < 0.0125$) are underlined

Table 3. Pairwise F_{ST} comparisons based on autosomal STR markers to test for genetic differentiation among three high-altitude populations from Jujuy province (NW Argentina)

STR marker	Statistic	QB / QA	QB / Puna	QA / Puna
D6S2792	F_{ST}	0.0095	**0.0506**	0.0075
	P	0.063 ± 0.023	< 0.001 ± < 0.001	0.126 ± 0.043
D6S105	F_{ST}	0.0035	**0.0257**	**0.0142**
	P	0.180 ± 0.027	< 0.001 ± < 0.001	0.045 ± 0.020
D6S265	F_{ST}	0.0064	0.0065	0.0033
	P	0.847 ± 0.028	0.766 ± 0.058	0.559 ± 0.052
D6S2666	F_{ST}	**0.0176**	**0.0868**	**0.0291**
	P	0.036 ± 0.015	< 0.001 ± < 0.001	0.009 ± 0.009
Haplotype	F_{ST}	0.0009	**0.0315**	0.0064
	P	0.676 ± 0.043	< 0.001 ± < 0.001	0.207 ± 0.036

P, statistical significance of F_{ST} indices

QB, Quebrada Baja; QA, Quebrada Alta

Statistically significant F_{ST} values among populations are highlighted in bold

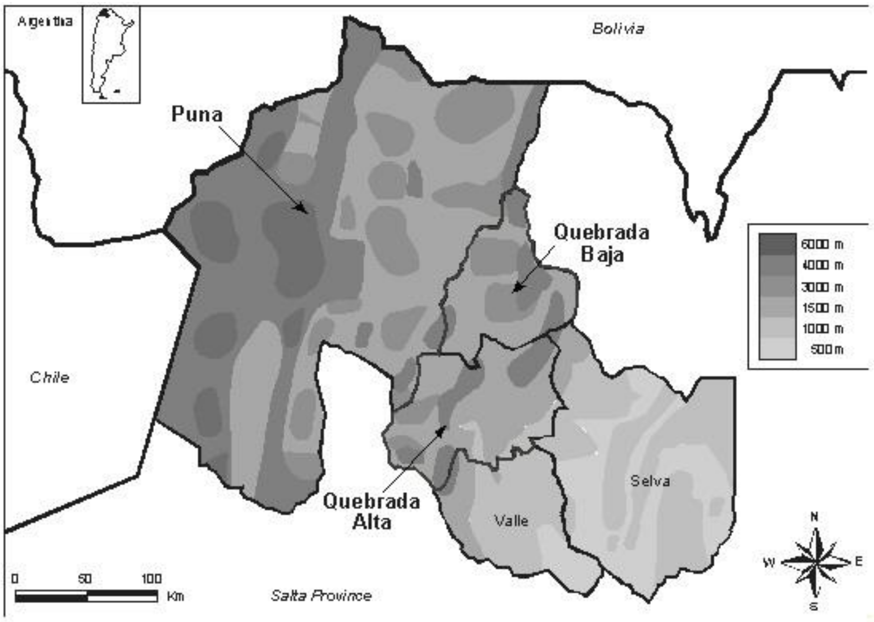

Fig. 1. Map of Jujuy province (NW Argentina) showing the geographic position of the three study populations (Puna, Quebrada Alta and Quebrada Baja). Altitude contour lines illustrate the height above sea level of the different regions.

Figure Legends

Fig. 1. Map of Jujuy province (NW Argentina) showing the geographic position of the three study populations (Puna, Quebrada Alta and Quebrada Baja). Altitude contour lines illustrate the height above sea level of the different regions.

Supporting Information Table 1: Linkage disequilibrium[a] analysis for autosomal STRs from chromosome 6 in population samples from Jujuy province (NW Argentina)

Locus	D6S2792	D6S265	D6S105
D6S265	0.204 ± 0.013	-	
D6S105	**0.013 ± 0.004**	**0.000 ± 0.000**	-
D6S2666	**0.006 ± 0.002**	0.324 ± 0.014	0.884 ± 0.008

[a] Slatkin and Excoffier (1996).

Figures are *P*-values ± standard errors (SE)

Statistically significant *P*-values are highlighted in bold

Supporting Information Table 2. Fixation indices (F_{ST}) generated by analysis of molecular variance (AMOVA) based on four STR markers and their haplotypic combination in three high-altitude populations from Jujuy province (NW Argentina)

STR marker	F_{ST}	*P*-value	± S.E.
D6S2792	0.02264	0.00098***	0.00009
D6S105	0.01437	0.01662*	0.00367
D6S265	0.00540	0.83284 n.s.	0.01012
D6S2666	0.04549	< 0.00001***	< 0.00001
Haplotype	0.02251	< 0.00001***	< 0.00001

F_{ST}, genetic variation among populations, S.E., standard error

n.s., not statistically significant

* statistically significant at $P < 0.05$

*** statistically significant at $P < 0.001$