doi: 10.1046/j.1529-8817.2003.00125.x

A Tale of Aborigines, Conquerors and Slaves: Alu Insertion Polymorphisms and the Peopling of Canary Islands

N. Maca-Meyer, J. Villar, L. Pérez-Méndez, A. Cabrera de León and C. Flores*

Unidad de Investigación, Hospital Universitario N. S. de Candelaria, Santa Cruz de Tenerife, Spain

Summary

Classical, mitochondrial DNA (mtDNA) and Y chromosome markers have been used to examine the genetic admixture in present day inhabitants of the Canary Islands. In this study, we report the analysis of ten autosomal Alu insertion polymorphisms in 364 samples from the seven main islands of the Archipelago, and their comparison to continental samples. The detection of population-specific alleles from the Iberian Peninsula and Northwest Africa, as well as their affinities on the basis of genetic distances and principal component analysis, support a clear link between these populations. Coincident with previous results, the Canarian gene pool can be distinguished as being halfway between those of its putative parents, although with a major Iberian contribution (62–78%). Both the substantial Northwest African contribution (23–38%), and the minor sub-Saharan African input (3%), suggest that the genetic legacy from the aborigines and slaves still persists in the Canary Islanders.

Keywords : Alu insertion polymorphisms, Canary Islands, Northwest Africa, Iberia

Introduction

The Canarian archipelago, off the Northwest coast of Africa (Figure 1), has undergone different population inputs. The European conquest of the Islands in the XVth century added an important Iberian contingent to the initial Northwest African aborigine substrate (Suárez *et al.* 1988). With time, the Iberian settlement brought the need for a farm labour force, giving rise to the importation of slaves from the West African coast (Lobo-Cabrera, 1982). Since the 1950s, the admixture in the Canarian population has been extensively analysed using genetic approaches. For classical markers (blood groups and red-blood-cell enzyme loci), the prevailing discovery was that Canarians were a typical European population with some African influence, most of it being due to Negroid alleles introduced by slaves (see Flores *et al.* 2001a for a review). More recently, uniparental markers (mtDNA and Y chromosome) have been used as important tools for interpreting demographic processes, taking advantage of their known phylogeny and higher population compartmentalisation (Semino *et al.* 2000, Richards *et al.* 2002). Particularly, in the case of the Canary Islands, both have provided a more reliable vision of the genetic composition of the present day population. From mtDNA data, besides the sub-Saharan African influence, a more important Northwest African one was detected (Pinto *et al.* 1996; Rando *et al.* 1999). Furthermore, ancient mtDNA has demonstrated that this Northwest African influence is not only due to slavery but mostly due to aborigine heritage (Maca-Meyer *et al.* 2004). Although no data is available to date on the ancient Y chromosome, present day Canarians show specific Northwest African paternal lineages, together with sub-Saharan African ones (Flores *et al.* 2003). However, although studies with autosomal markers are limited, they point to an overwhelming European influence (Esteban *et al.* 1998; Flores *et al.* 2001b).

Although all variation studies agree that current Canarians are an amalgam of Iberian, Northwest African

*Address for Correspondence: Carlos Flores Unidad de Investigación Hospital Universitario N. S. de Candelaria Carretera del Rosario s/n, 38010 Santa Cruz de Tenerife, Spain Phone: 34-922-600545 FAX: 34-922-600562 E-mail: cflores@ull.es

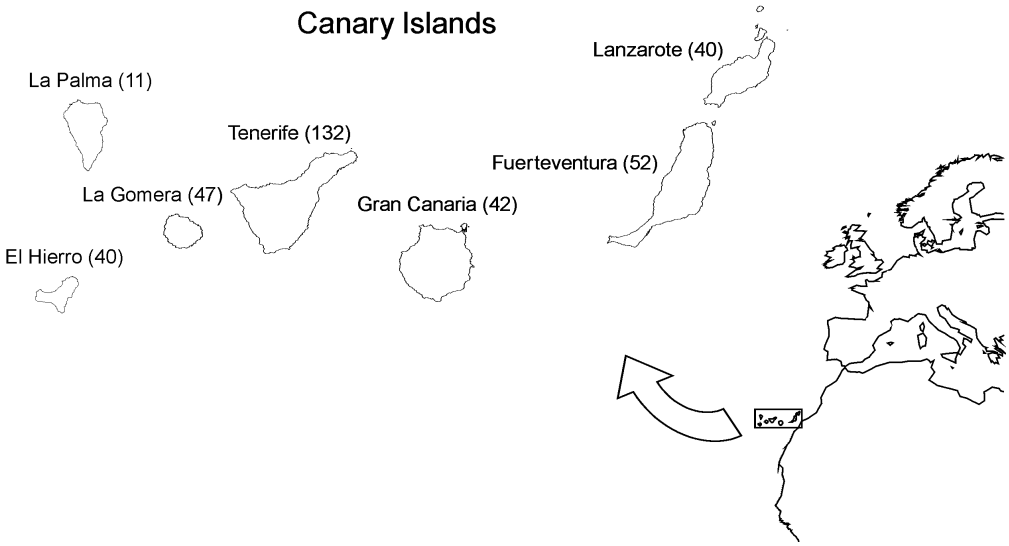

Figure 1 Map showing the approximate location of the Canary Islands. Sample sizes for each Island are indicated in parenthesis.

and sub-Saharan African populations, the extent to which each population contributed varies depending on the markers studied (Flores *et al.* 2001a). However, all studies agree with a principal Iberian input, followed by Northwest African and reduced sub-Saharan African input. Taking the Northwest African contribution as a proxy of the aboriginal substrate, the highest proportion is observed for maternal lineages, and the lowest for the paternal ones, with classical markers in between. These results have been interpreted as a consequence of a directional mating that can be explained by the way the Islands were conquered (Pinto *et al.* 1996).

Although uniparental markers have been very informative for demographic studies, their sexual inheritance and higher susceptibility to drift and selection could lead to biased information (Jobling & Tyler-Smith, 2003). Therefore, biparental markers are necessary to obtain a complete picture of the history and dispersals of human populations. Polymorphic Alu insertions (PAI), scattered throughout the human genome, occur as unique events in our evolution and are apparently selectively neutral (Batzer & Deininger, 2002). These markers have been used extensively for population structure and evolution, both at global (Bazter *et al.* 1996; Antúnez-de-Mayolo *et al.* 2002; Romualdi *et al.* 2002; Bamshad *et al.* 2003) and regional levels (Majumder *et al.* 1999; Comas *et al.* 2000, 2004; Martínez de Pancorbo *et al.* 2001; Nasidze *et al.* 2001). In this study, we genotyped 10 PAI in samples from the seven main islands of the Canarian archipelago, in order to fill in the gaps in autosomal studies. We compared their genetic composition to that of their putative parentals, and established the genetic admixture in this population.

Materials and Methods

Blood samples from a total of 364 unrelated healthy donors with at least two generations of Canarian ancestors were analysed. They are a subset of more than 6,000 samples collected in the cohort study called "CDC of the Canary Islands." This cohort was randomly selected as a representative sample of the Canarian population aged between 18 and 75 years. Sampling was stratified among the seven islands with no gender differences. Informed consent was obtained from all individuals through personal interviews. All samples were typed for ten PAI (A25, B65, ACE, D1, APO, FXIIIB, PV92, TPA25, HS3.23 and HS4.65) as previously described (Comas *et al.* 2000). Additional genotypic data from three Iberian (Basques, Catalans and Andalusians) and six Northwest African (Tunisians, Algerians, North Moroccans, Western Moroccans, Southeast Moroccans and Saharawi) populations were kindly provided by D. Comas and used for comparisons.

For the characterization of the large PV92 allele, the agarose bands with amplified products were excised

and purified (QIAGEN Gel extraction kit). The region was sequenced (BigDye® v3.1 Terminator Cycle Sequencing kit, Applied Biosystems) using the amplification primers under the manufacturer's recommendations. Products were ethanol precipitated and run on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems).

ARLEQUIN 2000 (Schneider *et al.* 2000) was used to calculate allele frequencies and to test for Hardy-Weinberg equilibrium by means of an exact test (Guo & Thomson, 1992). The genetic differentiation was examined by two complementary methods: a) hierarchical analysis of molecular variance (AMOVA), determined by ARLEQUIN 2000; b) Principal component analysis (PCA) of Alu insertion frequencies, as performed by SPSS ver. 11 package (SPSS, Inc.).

Three frequency-based admixture estimators were used to calculate the genetic contribution from Iberians and Northwest Africans: m_L (Long, 1991), m_R (Roberts & Hiorns, 1965) and m_C (Chakraborty *et al.* 1992). The ADMIX.PAS program (kindly provided by Dr. Jeffrey Long) was used to implement m_L, a weighted least-squares estimator that takes into account sampling error and drift. The m_R and m_C estimators were performed by means of the ADMIX 1.0 program (Bertorelle & Excoffier, 1998). The m_R, the first proposed frequency-based estimator, is a least-squares estimator that neglects stochastic effects apart from the sampling of the hybrid population. The m_C is a closed-form expression of m_L. The m_L estimator was also used in a model with three parentals. In this case, pooled allele frequency data from Bantu and Niger-Congo samples (http://alfred.med.yale.edu/alfred/index.asp) were used to represent the sub-Saharan African input. We also obtained individual admixtures using the STRUCTURE program (Pritchard *et al.* 2000), assuming that there are two populations (no multilocus genotype data were available for sub-Saharan Africans) contributing to the gene pool of Canarians. For comparison, individual admixture estimates were also assessed using mtDNA (Flores *et al.* 2001a) and Y chromosome data (Flores *et al.* 2003). All runs were replicated three times, employing a "no admixture" model for individual ancestry (as suggested in the software documentation for detecting subtle structure), and a burn-in period of length 20,000 followed by 50,000 iterations.

Results and Discussion

All Alu insertions typed were found to be polymorphic in all Canarian populations, except HS4.65 in La Palma (Table 1). No loci showed significant departure from Hardy-Weinberg equilibrium after Bonferroni correction (data not shown). Although all ten markers have been described as biallelic polymorphisms, we found three alleles for the locus PV92. In four heterozygous individuals (one from Tenerife, one from Gran Canaria and two from Fuerteventura) a larger-than-expected Alu insertion was found. The sequences of these alleles correspond to an Alu insertion within a pre-existing PAI, which has previously been found in two Basque and one northern Moroccan chromosomes from a broad set of populations worldwide (Comas *et al.* 2001, Romualdi *et al.* 2002, Bamshad *et al.* 2003). Its presence in the current Canarians evidences the clear and strong relationship of this population with those from Iberia and Northwestern Africa.

Table 1 Alu insertion frequencies in the Canary Islands

		Polymorphisms									
Population	2N[a]	A25	B65	ACE	D1	APO	FXIIIB	PV92	TPA25	HS3.23	HS4.65
La Palma	22	0.091	0.636	0.409	0.455	0.955	0.364	0.091	0.591	0.909	0
El Hierro	80	0.213	0.638	0.350	0.288	0.925	0.300	0.225	0.575	0.875	0.013
La Gomera	94	0.181	0.574	0.362	0.330	0.957	0.447	0.245	0.723	0.883	0.032
Tenerife	264	0.136	0.545	0.348	0.352	0.951	0.356	0.273*	0.629	0.894	0.030
Gran Canaria	84	0.214	0.548	0.393	0.333	0.976	0.464	0.262*	0.476	0.845	0.012
Lanzarote	80	0.125	0.475	0.350	0.413	0.913	0.338	0.275	0.588	0.838	0.013
Fuerteventura	104	0.125	0.577	0.433	0.337	0.962	0.433	0.221*	0.606	0.846	0.010
Total	728	0.155	0.559	0.370	0.348	0.949	0.383	0.250*	0.609	0.872	0.021

[a]Number of chromosomes analysed. *The four double Alu insertion alleles are included.

Table 2 F_{ST} genetic distances between Island and continental populations

Population	sub-Saharan Africa	Northwest Africa	Iberia
La Palma	0.0910***	0.0146	0
El Hierro	0.0550***	0.0011	0.0069
La Gomera	0.1027***	0.0144***	0.0017
Tenerife	0.0795***	0.0080***	0.0058**
Gran Canaria	0.0734***	0.0122**	0
Lanzarote	0.0722***	0.0084*	0.0077*
Fuerteventura	0.0895***	0.0134***	0
Total Canaries	0.0821***	0.0099***	0.0029**
Iberia	0.0953***	0.0215***	-
NW Africa	0.0624***	-	-

Significance of pairwise distances are *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

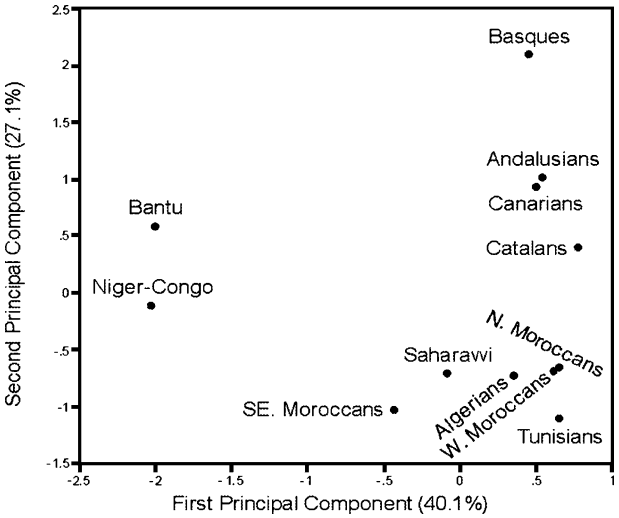

Figure 2 Plot of the first two principal components of the ten Alu insertion allele frequencies in Canarian and continental populations.

Analysis of molecular variance (AMOVA) between the Canarian populations showed that they constitute a genetically homogeneous group (0.02% of variance between populations, $P = 0.405$) for the analysed loci, except for a limited heterogeneity found for TPA25 (1.12% of variance between populations, $P = 0.041$). Although archaeological and anthropological data point to differentiation between the western and eastern islands, this heterogeneity was not detected (0.13% of variance between groups, $P = 0.219$). Although there are differences among islands, when Canarians are compared through F_{ST} genetic distances to pools of continental populations (Table 2), they appeared to be about three times more similar to Iberians than to Northwest Africans. In all cases, the most distant population was sub-Saharan Africa. An exceptional case was El Hierro, for which the genetic distances, although non-significant, point to a closer affinity with Northwest Africans than Iberians. This distinctiveness could be due to a higher sub-Saharan African influence in this Island, as has been observed for mtDNA (Rando *et al.* 1999, Flores *et al.* 2001a) and the Y chromosome (Flores *et al.* 2003). However, it should be noted that the distance between Canarians and Northwest Africans were two times smaller than those between Iberians and Northwest Africans (Table 2). Using Canarians as a single population, we then compared their genetic pools using PCA (Figure 2). The PCA plot, with the first two components accounting for 67.2% of genetic variance, cluster populations into three groups: sub-Saharan Africans, Iberians and Northwest Africans, although with some heterogeneity within the latter (Comas *et al.* 2000). Canarians are clearly clustered with Iberian populations.

Admixture estimates have been obtained using Iberians and Northwest Africans as the main contributors to the present-day Canarians (Table 3). The estimations with m_L and m_C gave similar results, with an Iberian contribution of over 70%. However, m_R estimated this contribution to be around 62%. Using an admixture model with three parentals, a minimal 3% of sub-Saharan African input was calculated. This value is intermediate between those obtained for Y chromosome and mtDNA ($\approx$1% and 5-21%, respectively). Since a possible sub-Saharan African influence has been attributed to Saharawi and Southeast Moroccans (Comas *et al.* 2000), calculations were repeated without these populations, resulting in virtually the same estimates (not shown). El Hierro, with $16.7 \pm 7.9\%$ sub-Saharan African contribution, stands out again from the rest of the islands (data not shown). However, it must be noted that this value should be interpreted with caution, since the pools of sub-Saharan African populations used for this calculation may depart significantly from that of the slaves introduced in historic times, who were mostly imported from Western parts of the African continent (Lobo-Cabrera, 1982). The distinctive pattern of El Hierro can be explained by the way the island was conquered and re-colonized: it was one of the

Table 3 Admixture estimates (± SD) for the Canarian population

	Iberian	Northwest African	sub-Saharan African
m_R	0.617 ± 0.078	0.383 ± 0.078	n.t.
m_L	0.723 ± 0.085	0.278 ± 0.085	n.t.
	0.742 ± 0.087	0.227 ± 0.106	0.031 ± 0.045
m_C	0.779 ± 0.085	0.221 ± 0.085	n.t.

n.t., not tested

most castigated islands in relation to aboriginal conservation, and the increase in African raids during the XV century enhanced the slave trade to the island (Suárez *et al*. 1988). As a whole, a significant Northwest African contribution was detected in the Canarian population, which is in the range of the 20-30% calculated from classical markers (Flores *et al*. 2001a). We then explored whether these estimates are homogeneous for all individuals of the Canarian population. Using STRUCTURE, we calculated the proportion of an individual's genome that can be assigned to the Iberian gene pool. As shown in Figure 3, most individuals presented an intermediate ancestry, although some showed a high resemblance to either one of the parentals (i.e. high or low Iberian ancestry). However, this result can be explained by the reduced geographic differentiation of these autosomal loci, as evidenced by comparison of these populations using mtDNA and Y chromosome markers, which are more sensitive to levels of population subdivision. Specifically, with PAI, the percentage of individuals that had high ($\geq$ 0.7) or low ($<$ 0.3) levels of Iberian ancestry was 36% (23% with high and 13% with low). When mtDNA was used for such inference, 84% of individuals had high or low Iberian ancestry (63% with high and 21% with low), while the Y chromosome yielded 73% (59% with high and 15% with low). Thus, although not conclusive, the present autosomal data point to some heterogeneity in the individual ancestries of Canary Islanders.

In summary, the typing of these ten autosomal markers has evidenced an important similarity between present day Canarians and the Iberian population. However, a notable African influence, mostly from the Northwestern region, was also detected confirming previous findings with classical and uniparental markers. Altogether, these results are concordant with the way the Islands were conquered and colonised, and with the maintenance of a substantial proportion of aboriginal heritage in the current population of the Archipelago.

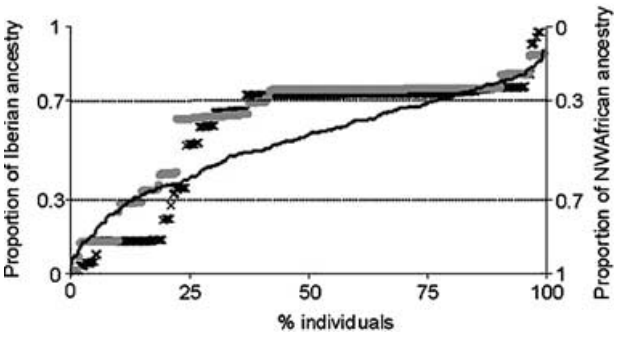

Figure 3 Plot of individual admixtures obtained from Alu insertion data compared to mtDNA (×) and the Y chromosome (●). Note that, since the set of points derived from the Alu insertion data is dense, it gives the impression of a continuum line in the plot. Left hand Y-axis represents Iberian ancestry whilst right hand represents the complementary Northwest African ancestry.

Acknowledgements

We would like to thank V.M. Cabrera and three anonymous reviewers for their helpful comments. This work was supported by grants 37/02 from the Fundación Canaria de Investigación y Salud (FUNCIS) to J.V., 99/0361 and 02/1189 from the Fondo de Investigación Sanitaria (FIS), to L. P.-M. and A.C., respectively. N.M.-M. and C.F. are FUNCIS postdoctoral fellows.

References

Antúnez-de-Mayolo, G., Antúnez-de-Mayolo, A., Antúnez-de-Mayolo, P., Papiha S. S., Hammer, M., Yunis, J. J., Yunis, E. J., Damodaran, C., Martínez de Pancorbo, M., Caeiro, J. L., Puzyrev, V. P. & Herrera, R. J. (2002) Phylogenetics of worldwide human populations as determined by polymorphic Alu insertions. *Electrophoresis* **23**, 3346–3356.

Bamshad, M. J., Wooding, S., Watkins, W. S., Ostler, C. T., Batzer, M. A. & Jorde, L. B. (2003) Human Population Genetic Structure and Inference of Group Membership. *Am J Hum Genet* **72**, 578–589.

Batzer, M. A. & Deininger, P. L. (2002) Alu repeats and human genomic diversity. *Nat Rev Genet* **3**, 370–379.

Batzer, M. A., Arcot, S. S., Phinney, J. W., Alegria-Hartman, M., Kass, D. H., Milligan, S. M., Kimpton, C., Gill, P., Hochmeister, M., Ioannou, P. A., Herrera, R. J., Boudreau, D. A., Scheer, W. D., Keats, B. J., Deininger, P. L. & Stoneking, M. (1996) Genetic variation of recent Alu insertions in human populations. *J Mol Evol* **42**, 22–29.

Bertorelle, G. & Excoffier, L. (1998) Inferring admixture proportions from molecular data. *Mol Biol Evol* **15**, 1298–1311.

Chakraborty, R., Kamboh, M. I., Nwankwo, M. & Ferrell, R. E. (1992) Caucasian genes in American Blacks: new data. *Am J Hum Genet* **50**, 145–155.

Comas, D., Calafell, F., Benchemsi, N., Helal, A., Lefranc, G., Stoneking, M., Batzer, M. A., Bertranpetit, J. & Sajantila, A. (2000) Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary though the Gibraltar straits. *Hum Genet* **107**, 312–319.

Comas, D., Plaza, S., Calafell, F., Sajantila, A. & Bertranpetit, J. (2001) Recent insertion of an *Alu* element within a polymorphic human-specific *Alu* insertion. *Mol Biol Evol* **18**, 85–88.

Comas, D., Schmid, H., Braeuer, S., Flaiz, C., Busquets, A., Calafell, F., Bertranpetit, J., Scheil, H. G., Huckenbeck, W., Efremovska, L. & Schmidt, H. (2004) Alu insertion polymorphisms in the Balkans and the origins of the Aromuns. *Ann Hum Genet* **68**, 120–127.

Esteban, E., Dugoujon, J. M., Valveny, N., González-Reimers, E. & Moral, P. (1998) Spanish and African contributionto the genetic pool of the Canary islanders:data from GM and KM haplotypes and RFLPs in the immunoglobulin IGHG loci. *Ann Hum Genet* **62**, 33–45.

Flores, C., Larruga, J. M., González, A. M., Hernández, M., Pinto, F. M.& Cabrera, V. M. (2001a) The origin of the Canary Island aborigines and their contribution to the modern population: a molecular genetics perspective. *Curr Anthropol* **42**, 749–755.

Flores, C., Maca-Meyer, N., Pérez, J. A. & Cabrera, V. M. (2001b) The peopling of the Canary Islands: a CD4/Alu microsatellite haplotype perspective. *Hum Immunol* **62**, 949–953.

Flores, C., Maca-Meyer, N., Pérez, J. A., González, A. M., Larruga, J. M. & Cabrera, V. M. (2003) A predominant European ancestry of paternal lineages from Canary Islands. *Ann Hum Genet* **67**, 138–152.

Guo, S. & Thompson, E. (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**, 361–372.

Jobling, M. A. & Tyler-Smith, C. (2003) The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* **4**, 598–612.

Lobo-Cabrera, M. (1982) *La esclavitud en las Canarias orientales en el siglo XVI*. Las Palmas de Gran Canaria, Excmo. Cabildo Insular de Gran Canaria.

Long, J. C. (1991) The genetic structure of admixed populations. *Genetics* **127**, 417–428.

Maca-Meyer, N., Arnay, M., Rando, J. C., Flores, C., González, A. M., Cabrera, V. M. & Larruga, J. M. (2004) Ancient mtDNA analysis and the origin of the Guanches. *Eur J Hum Genet* **12**, 155–162.

Majumder, P. P., Roy, B., Banerjee, S., Chakraborty, M., Dey, B., Mukherjee, N., Roy, M., Thakurta, P. G. & Sil, S. K. (1999) Human-specific insertion/deletion polymorphisms in Indian populations and their possible evolutionary implications. *Eur J Hum Genet* **7**, 435–446.

Martínez de Pancorbo, M., López-Martínez, M., Martínez-Bouzas, C., Castro, A., Fernández-Fernández, I., Antúnez de Mayolo, G., Antúnez de Mayolo, A., Antúnez de Mayolo, P., Rowold, D., & Herrera, R. (2001) The Basques according to polymorphic Alu insertions. *Hum Genet* **109**, 224–233.

Nasidze, I., Risch, G. M., Robichaux, M., Sherry, S. T., Batzer, M. A. & Stoneking, M. (2001) Alu insertion polymorphisms and the genetic structure of human populations from the Caucasus. *Eur J Hum Genet* **9**, 267–272.

Pinto, F., González, A. M., Hernández, M., Larruga, J. M. & Cabrera, V. M. (1996) Genetic relationship between the Canary Islanders and their African and Spanish ancestors inferred from mitochondrial DNA sequences. *Ann Hum Genet* **60**, 321–330.

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

Rando, J. C., Cabrera, V. M., Larruga, J. M., Hernández, M., González, A. M., Pinto, F. & Bandelt, H. J. (1999) Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. *Ann Hum Genet* **63**, 413–428.

Richards, M., Macaulay, V., Torroni, A. & Bandelt, H. J. (2002) In search of geographical patterns in European mitochondrial DNA. *Am J Hum Genet* **71**, 1168–1174.

Roberts, D. & Hiorns, R. (1965) Methods of analysis of the genetic composition of a hybrid population. *Hum Biol* **37**, 38–43.

Romualdi, C., Balding, D., Nasidze, I. S., Risch, G., Robichaux, M., Sherry, S. T., Stoneking, M., Batzer, M. A. & Barbujani, G. (2002) Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms. *Genome Res* **12**, 602–612.

Schneider, S., Roessli, D. & Excoffier, L. (2000) *Arlequin ver. 2000: A software for population genetics data analysis*. Switzerland, Genetics and Biometry Laboratory, University of Geneva.

Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., Marcikiae, M., Mika, A., Mika, B., Primorac, D., Santachiara-Benerecetti, A. S., Cavalli-Sforza, L. L. & Underhill, P. A. (2000) The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* **290**, 1155–1159.

Suárez, J. J., Rodríguez, F. & Quintero, C. L. (1988) *Historia popular de Canarias. Vol. 2. Conquista y colonización*. Tenerife, Centro de la Cultura Popular Canaria.

Received: 21 April 2004
Accepted: 23 June 2004

