

## ORIGINAL ARTICLE

## Genetic uniqueness of the Waorani tribe from the Ecuadorian Amazon

S Cardoso<sup>1</sup>, MA Alfonso-Sánchez<sup>1</sup>, L Valverde<sup>1</sup>, D Sánchez<sup>2</sup>, MT Zarrabeitia<sup>3</sup>, A Odriozola<sup>1</sup>, B Martínez-Jarreta<sup>4</sup> and MM de Pancorbo<sup>1</sup>

South America and especially the Amazon basin is known to be home to some of the most isolated human groups in the world. Here, we report on a study of mitochondrial DNA (mtDNA) in the Waorani from Ecuador, probably the most warlike human population known to date. Seeking to look in more depth at the characterization of the genetic diversity of this Native American tribe, molecular markers from the X and Y chromosomes were also analyzed. Only three different mtDNA haplotypes were detected among the Waorani sample. One of them, assigned to Native American haplogroup A2, accounted for more than 94% of the total diversity of the maternal gene pool. Our results for sex chromosome molecular markers failed to find close genetic kinship between individuals, further emphasizing the low genetic diversity of the mtDNA. Bearing in mind the results obtained for both the analysis of the mtDNA control region and complete mitochondrial genomes, we suggest the existence of a 'Waorani-specific' mtDNA lineage. According to current knowledge on the phylogeny of haplogroup A2, we propose that this lineage could be designated as subhaplogroup A2s. Its wide predominance among the Waorani people might have been conditioned by severe genetic drift episodes resulting from founding events, long-term isolation and a traditionally small population size most likely associated with the striking ethnography of this Amazonian community. In all, the Waorani constitute a fine example of how genetic imprint may mirror ethnopsychology and sociocultural features in human populations.

*Heredity* (2012) 108, 609–615; doi:10.1038/hdy.2011.131; published online 11 January 2012

**Keywords:** human isolate; complete sequencing; mitochondrial DNA; genetic drift; population bottleneck; native American

## INTRODUCTION

Isolated populations living in remote and/or inaccessible parts of the world are regarded as biological treasures from the genetic viewpoint. Many of these isolated human groups have remained relatively unknown until very recent times, so that the information provided by population genetic studies can help the scientists in the partial reconstruction of their demographic and evolutionary histories. The Waorani of the Ecuadorian Amazon (the singular and the adjectival form is Wao) are one such population. The Waorani community is an indigenous tribe currently numbering ~1500–2000 individuals. They inhabit the Yasuní National Forest, a Pleistocene forest refugium located between the Napo and Curaray rivers in the eastern Ecuadorian provinces of Napo, Orellana and Pastaza (Figure 1). The Waorani are descended from hunter-gatherer groups that once inhabited much of the lowland tropical Amazon forests (Wallis, 1973) and they remain essentially a society based on hunting and horticulture. The Waorani are a highly inbred, homogenous population (Larrick *et al.*, 1985), where bilateral cross-cousin marriage is prescribed and is traditionally arranged by the parents of the young couple, often without their knowledge (Robarchek and Robarchek, 1992; Beckerman *et al.*, 2009).

The Waorani have been considered as the most warlike society yet described owing to exceptionally high homicide rates (Yost, 1981; Beckerman *et al.*, 2009). Disputes over marriage arrangements, blood

feuds and retaliations following past killings are commonplace, even among closely related groups. On the basis of extensive genealogies (up to five generations back) collected in the 1970s, Yost (1981) estimated that more than 60% of adult deaths were the result of warfare: 17% due to external raiding and 44% to spearing in fighting within the tribe. In such a context, violence is seen as a major determinant of population mortality (Robarchek and Robarchek, 1992; Boster *et al.*, 2004; Beckerman *et al.*, 2009).

The Waorani were the last indigenous tribe living in the Ecuadorian Amazon to be contacted by western cultures in the 20th century. The first sustained peaceful contact was made by US missionaries in 1958 (Wallis, 1973). Since that time the Waorani have had limited but increasing contact with outsiders, although even now-a-days some Waorani groups survive completely isolated in the forest (CODENPE, 2002; Boster *et al.*, 2004).

The Waorani speak Wao tiriro (*Wao tededo*), a language that is apparently unrelated to any other in the region (Peeke, 1973). Thus, both the ethnopsychology and the linguistic singularity of the Waorani have led to a marked population isolation, which makes them an exceptional anthropological group for evolutionary and genetic studies. Seeking to look in more depth at the genetic characterization of this Native American population, we tackle herein the analysis of maternal (mitochondrial DNA, mtDNA) and

<sup>1</sup>BIOMiCs Research Group, Centro de Investigación y Estudios Avanzados 'Lucio Lascaray', Universidad del País Vasco UPV/EHU, Vitoria-Gasteiz, Spain; <sup>2</sup>Laboratorio de Genética Molecular DIAGEN, Metropolitan Hospital, Quito, Ecuador; <sup>3</sup>Unidad de Medicina Legal, Universidad de Cantabria, Santander, Spain and <sup>4</sup>Departamento de Medicina Forense, Facultad de Medicina, Universidad de Zaragoza, Zaragoza, Spain

Correspondence: Professor MM de Pancorbo, BIOMiCs Research Group, Centro de Investigación y Estudios Avanzados "Lucio Lascaray", Universidad del País Vasco UPV/EHU, Avda. Miguel de Unamuno, 3, Vitoria-Gasteiz, Alava 01006, Spain.

E-mail: marian.mdepancorbo@ehu.es

Received 25 March 2011; revised 30 October 2011; accepted 7 November 2011; published online 11 January 2012



**Figure 1** Map of Ecuador showing the geographic location of the Yasuní National Forest (dashed line), the homeland of the Waorani people, located in the eastern provinces of Napo, Orellana and Pastaza.

paternal (Y chromosome) lineages, as well as X chromosome short tandem repeat (STR) markers in a Waorani population sample.

The analysis of mtDNA at the highest level of resolution (complete genomes) revealed the existence of a ‘Waorani-specific’ mtDNA lineage that we have called A2s. In the main, the genetic characterization carried out in this study has brought to light the extremely low genetic diversity of the Waorani. This limited variability might have been generated by genetic drift associated with founder effects during the early settlement of the Americas, and also with recurrent population bottleneck episodes resulting from the intense demographic stress generated by internal fights, retaliation homicides and warfare, which have been commonplace among the Waorani until very recent times.

## MATERIALS AND METHODS

A total of 36 Waorani individuals living in various communal longhouses or *nanicabos* (traditional basic units or domestic groups) were included in the sample. Genealogical reconstruction is a complex task in the Waorani society, because every individual has several names (typically 4 to 6), and may often use different names at different stages of life and in different social and geographical contexts (Beckerman *et al.*, 2009). We estimate that the Wao individuals included in our study belong to at least 11 different familial clans, according to the surname composition of the sample. All voluntary donors gave their informed consent before inclusion in the sample, following the ethical principles and guidelines of the Declaration of Helsinki for the protection of human subjects of research. The study protocol was approved by the Institutional Review Board from Universidad del País Vasco.

### Maternal and paternal lineages

DNA was extracted from blood samples using the Wizard Genomic DNA Purification Kit System (Promega Corp., Madison, WI, USA). To analyze mtDNA we amplified the hypervariable segments HVS-I and HVS-II of the control region, following the methodology described in a previous publication (Alfonso-Sánchez *et al.*, 2008). In all, 11 individuals with different surnames

were selected for complete sequencing of the mitochondrial genome. The procedure of complete sequencing entailed PCR amplification of 15 overlapping mtDNA fragments that were sequenced by using 38 primers (Supplementary Table S1). DNA sequences were edited using Sequence Scanner v1.0 software (Applied Biosystems, Foster City, CA, USA). They were subsequently aligned and compared with the revised Cambridge Reference Sequence (Andrews *et al.*, 1999), using the Clustal X v2.0 program (Larkin *et al.*, 2007). All the polymorphic positions detected were confirmed in chromatograms. Haplogroups were assigned based on the phylogenetic tree of global human mtDNA variation available at [www.phylotree.org](http://www.phylotree.org) (van Oven and Kayser, 2009). All the sequences reported in this article are available online at GenBank under accession nos. GQ398432-GQ398479 and HQ709171-HQ709172 (25 HVSI–HVSII sequences), and GQ398480-GQ398490 (11 complete genomes).

For paternal lineages, we analyzed Y-chromosome haplotypes in the 11 men included in the sample through the amplification of 17 Y-STR loci in a single PCR amplification using the AmpFSTR Yfiler PCR Amplification Kit, as described by the manufacturer (Applied Biosystems, Foster City, CA, USA). The Y Chromosome Haplotype Reference Database, YHRD (Willuweit and Roewer, 2007) was used to check whether the Y chromosome haplotypes found in the Waorani existed in other Native American populations. The findings on Y-STR haplotype diversity in the Wao population sample were further examined in a broader geographic context. To that end, Y-STR data for other Native American groups were compiled from previous studies. With this integrative approach we sought to assess the genetic affinities of the Amerindian populations included in the analysis based on paternal lineages (Y-STR haplotypes). Bearing in mind the paucity of population data including the 17 STRs analyzed herein, we focused the comparative analysis on those Native American samples with data for 11 STRs. Information regarding ethnic group, geographic origin, sample size and data source for the Native American groups included in the analysis is given in Supplementary Table S2. Fst genetic distances (Reynolds *et al.*, 1983) were computed between all pairs of populations from haplotype frequency data using the Phylip v3.69 program (Felsenstein, 1989). Non-metric multidimensional scaling (MDS) was then applied to represent the Fst genetic distance matrix obtained in a two-dimensional space. MDS was performed using SPSS v16.0 statistical package (SPSS Inc., Chicago, IL, USA).

### X-chromosome molecular markers

We typed 10 X-chromosome STRs, including DXS8378, DXS9902, DXS7132, DXS9898, DXS6809, DXS6789, DXS7133, GATA172D05, GATA31E08 and DXS7423 using previously described methods (Gusmão *et al.*, 2009). Analyses were performed as described elsewhere (Zarrabeitia *et al.*, 2009). Allele frequencies were estimated by direct counting. Hardy–Weinberg equilibrium was tested in the female subsample by utilizing Genepop 3.4 (Raymond and Rousset, 1995). Gene diversity in females was calculated using Arlequin v3.0 (Excoffier *et al.*, 2005).

## RESULTS

### Maternal lineages

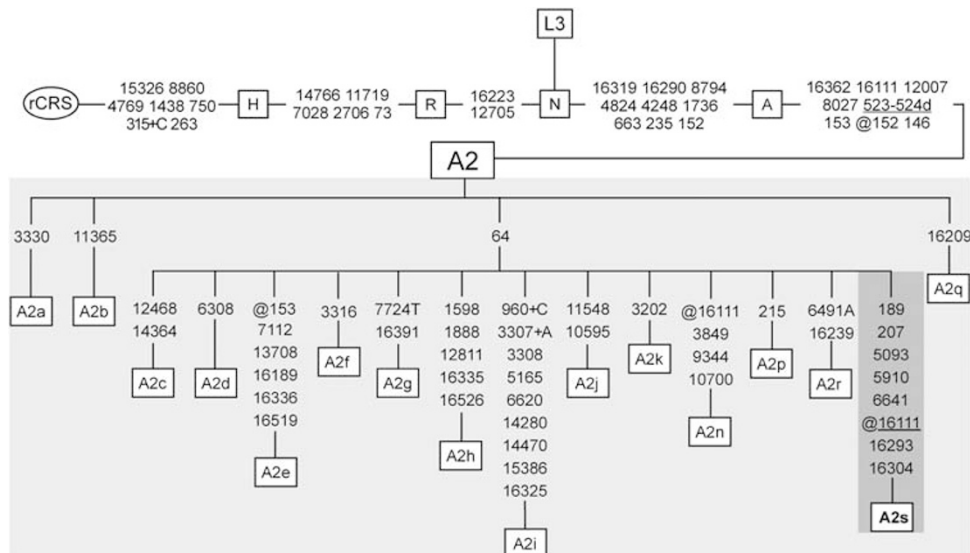
The mitochondrial genome of the Wao population sample showed an extremely low diversity: only three different haplotypes were observed in the analysis of HVS-I/HVS-II sequences. The diagnostic HVS-I/HVS-II haplotype of 35 individuals (97.2% over total) was defined by motifs C16223T, C16290T, G16319A and T16362C at the HVS-I region, and C64T, A73G, T146C, A153G, A235G, A263G, 309.1C and 315.1C at HVS-II. Of them, 34mtDNA sequences (94.4% over total) were characterized by a transition at np 16111. The lack of this transition defined the second haplotype, found in only one (2.8%) sequence (sample ECWA24). Based on the root haplotype, these individuals were all assigned to haplogroup A2. Interestingly, the root HVS-I/HVS-II haplotype harbored four additional polymorphisms, namely A16293G, T16304C, A189G and G207A. The third haplotype, defined by mutations T16172C, C16223T, T16325C, T16362C, C67T, D105-110, C150T, A263G and 315.1C was assigned to haplogroup D1 (sample ECWA34).

To extend the characterization of the leading haplogroup among the Waorani, complete sequencing of the mitochondrial genome of 11 individuals was performed. The analysis of the mtDNA coding region confirmed haplogroup assignment based on control region. Thus, excepting ECWA24, who showed no transition at np 16111, a single haplotype was shared by all the individuals selected for complete sequencing. Three specific mutations defined this ‘Waorani haplotype’: two synonymous base substitutions (T5093C and T6641C), and

a nonsynonymous transition (G5910A). Such nucleotide substitutions had been previously reported by Kivisild *et al.*, 2006 in the single Wao individual included in their study (referred to as Auca Indian, Am17). A survey of the 2704 complete sequences compiled in the mtDB (Ingman and Gyllensten, 2006), and of the 7843 entire human mtDNA sequences accessible through the phylotree webpage (van Oven and Kayser, 2009) confirmed that these three mutations only appear simultaneously within the A2 lineage described herein (Figure 2). These results strongly suggest that the cited haplotype may be a Waorani-specific maternal lineage, whose prevalence in the population might have been reinforced by population isolation and by the distinctive ethnographic characteristics of the Waorani people. Bearing in mind that this lineage delineates a specific branch within the phylogeny of haplogroup A2, and considering the increase in numbers of the Wao community as they were contacted by missionaries in 1958 (Peeke, 1973), we propose that this Waorani-specific maternal lineage could be designated as A2s.

### Paternal lineages of the Y-chromosome

To rule out the idea that the extremely low diversity of maternal lineages could result from biases in the sampling of the Waorani community, we analyzed the paternal lineages of the males included in the sample. We identified a total of five different haplotypes among 11 males, suggesting the existence of at least five paternal lineages (Table 1). Of these lineages, 2 (18.2% over total males) were individual specific and three were shared by three individuals apiece. Two of the most common haplotypes identified in this study (Hp1 and Hp5) had been previously observed in two Waorani samples available at the Y-Chromosome Haplotype Reference Database, YHRD (Willuweit and Roewer, 2007) under accession numbers YA003552 and YA003553 (Geppert *et al.*, 2011); yet, the cited haplotypes were not found in the remaining 245 world populations (with 17 Y-STRs) searchable in YHRD release 37. Diversity estimates for the Y chromosome ( $0.836 \pm 0.070$ ) indicated low variability of paternal lineages, which might be evidence for the existence of certain kinship relationships



**Figure 2** Phylogenetic tree of mtDNA haplogroup A2 summarized from van Oven and Kayser, 2009 and rooted in African haplogroup L3. The phylogeny includes the specific branch delineated by the Waorani (dark-shaded), denoted as subhaplogroup A2s. Mutations are transitions unless marked otherwise: suffixes indicate indels (+, d) or transversion to T at np 7724 in A2g. Mutations back to the revised Cambridge Reference Sequence (rCRS) nucleotide are prefixed with @. The back mutation at np 16111 within subhaplogroup A2s was observed only in one of 35 individuals analyzed. Recurrent mutational events are underlined. The position of the rCRS (Andrews *et al.*, 1999) is indicated for reading off sequence motifs.

among individuals. Nonetheless, the identification of several haplotypes suggested that such a low diversity was not a direct consequence of close kinship between the Wao males examined.

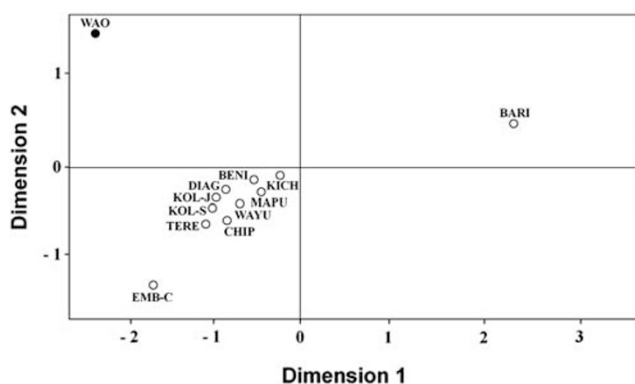
Results of the MDS applied on the  $F_{ST}$  genetic distance matrix computed from Y-STR haplotype frequencies are shown in Figure 3. The MDS plot proved to be statistically robust, as the total variance accounted for was 99.9% and the coefficient of stress was 2.6%. The main cluster comprised 10 (out of 12) populations, all of them positioned in the quadrant delimited by the negative segments of both Dimension 1 and 2. The great majority of these populations plotted relatively close to the centroid of the distribution, with the exception of the Emberá–Chamí from Colombia. No geographical trend was detected in the MDS topology: the only North American

group (Chippewa from Wisconsin) clustered with South American populations from different geographic origins such as Diaguita, Kolla and Mapuche from Argentina, Wayuu from Venezuela, Kichwa from Ecuador, Terena from Brazil and a multiethnic sample (BENI) from Beni, Bolivia. On the other hand, the most differentiated human groups from the Y-chromosome viewpoint proved to be the Barí (Venezuela), and the study population, the Waorani from Ecuador. Both collections plotted in the positive semi axis of Dimension 2, each of them occupying a distinctive quadrant and clearly distant from the centroid, which is indicative of a high genetic heterogeneity, and therefore, of a low genetic affinity with the rest of the Native American populations included in the analysis.

**Table 1** Lineages of Y-chromosome STRs observed among the 11 Wao males included in the sample

Haplotype	1	2	3	4	5
Frequency	3	1	1	3	3
DYS19	14	13	14	14	13
DYS389I	13	12	13	13	13
DYS389II	30	30	30	30	28
DYS390	23	25	23	23	23
DYS391	10	10	10	10	10
DYS392	15	14	15	15	14
DYS393	13	14	13	13	13
DYS385	14,16	14,19	14,16	14,16	15,15
DYS438	12	11	12	12	11
DYS439	12	12	13	11	10
DYS437	14	15	14	14	14
DYS448	21	20	21	21	20
DYS456	15	18	15	15	15
DYS458	17	17	17	17	16
DYS635	22	22	22	22	24
YGATAH4	12	11	12	12	12

Abbreviation: STR, short tandem repeat.



**Figure 3** Non-metric MDS applied on a Reynolds ( $F_{ST}$ ) genetic distance matrix for 12 Native American populations. Genetic distances were computed from Y-STR haplotype frequencies (considering 11 Y-STRs). Population are Barí (BARI) and Wayuu (WAYU) from Venezuela, Terena (TERE) from Brazil, Kichwa (KICH) and Waorani (WAO) from Ecuador, Emberá–Chamí (EMB-C) from Antioquia, Colombia, Chippewa (CHIP) from Wisconsin, USA, a multiethnic sample (BENI) from Beni, Bolivia, Kolla from Jujuy (KOL-J) and from Salta (KOL-S), Argentina, and Diaguita (DIAG) and Mapuche (MAPU) also from Argentina. The study population is highlighted (solid circle). Total variance accounted for: 99.9%. Coefficient of stress: 2.6%.

### X-chromosome molecular markers

The variability of paternal lineages seems to indicate that the Waorani included in the sample were not closely related through the paternal line. Nevertheless, we also analyzed X-chromosome molecular markers to ensure that the markedly low diversity of the Wao maternal lineages was not the consequence of close matrilineal kinship. Female genotypes and male haplotypes are shown in Supplementary Table S3, while allele frequencies of X-STR markers are listed in Supplementary Table S4. Fisher's exact tests performed on data from females found no departure from Hardy–Weinberg equilibrium expectations. The marker DXS9902 showed a marginally significant uncorrected  $P$ -value (0.045), but it was well below the significance threshold after applying the Bonferroni procedure for multiple comparisons (Bland and Altman, 1995). Average diversity was 0.569, the lowest value observed to date in populations from South America (Gusmão *et al.*, 2009). The paucity of data on Native American populations for the X-decaplex utilized herein prevented us from performing a comparative analysis of allele distribution between the Waorani and other Amazonian populations. However, we could compare the allele frequency distributions for markers DXS8378, DXS9898 and DXS7423 between the Waorani and five Native American populations from Gran Chaco (Catanesi *et al.*, 2007) through a likelihood ratio ( $G$ ) test (Table 2). Allele distribution of DXS9898 showed no significant differences between the Waorani and three out of the five Gran Chaco native populations. On the other hand, allele distribution of DXS7423 showed significant differences in three of the populations, and the differences at DXS8378 were statistically significant for all the populations compared with the Waorani. As expected, when the five populations from Gran Chaco were pooled gene diversity increased and the allele distributions of all three markers were significantly different from those of the Waorani. In all, the lack of significant differences for some of the nuclear markers compared between the Waorani and panmictic Amerindian populations from Gran Chaco,

**Table 2** Frequency distribution of DXS8378, DXS9898 and DXS7423 in various Amerindian populations

	Mocoví	Chorote	Wichí	Lengua	Ayoreo	Total
DXS8378	27.97	26.02	48.78	21.58	40.49	41.87
	$P < 0.0001$	$P < 0.0001$	$P < 0.0001$	$P = 0.0024$	$P < 0.0001$	$P < 0.0001$
DXS9898	16.15	<b>5.95</b>	<b>4.72</b>	<b>7.59</b>	23.58	18.91
	$P = 0.0129$	<b><math>P = 0.1140</math></b>	<b><math>P = 0.1931</math></b>	<b><math>P = 0.0055</math></b>	$P < 0.0001$	$P = 0.0021$
DXS7423	<b>4.2</b>	<b>12.55</b>	15.12	20.97	48.78	31.03
	<b><math>P = 0.523</math></b>	$P = 0.051$	$P = 0.0017$	$P = 0.0003$	$P < 0.0001$	$P < 0.0001$

Results of the comparisons of the allelic distribution of markers DXS8378, DXS9898 and DXS7423 between the Waorani and five Native American populations from Gran Chaco. Values for the likelihood ratio test ( $G$ -test) and their corresponding significance tests ( $P$ -values) are displayed. Comparisons rendering no statistically significant differences are highlighted in bold.

along with the variability of the temporary haplotypes of the X chromosome of the Wao males might suggest the absence of close kinship between the Waorani individuals analyzed herein, thereby reinforcing the hypothesis on the singularity of their maternal lineages.

## DISCUSSION

One major issue to be considered when a small population is analyzed for population genetics purposes is the potential bias that may result from close genetic kinship between the individuals sampled. Avoiding the effects of sampling bias is a very complex task in populations such as the Waorani because, in addition to the small demographic size, researchers must also face other problems that hinder checks on the representativeness of the sample. In our specific case, important limitations might be, for instance, the isolation experienced by the Waorani until recent times, their tradition of using up to six different names at different stages of life (Beckerman *et al.*, 2009) and the difficulty of communicating with them. All these limitations along with the distinctive socio-cultural practices of this Native American community make the Waorani an invaluable model for studying the effect of genetic drift in the gene pool of human populations.

The Waorani are descended from hunter-gatherer groups (Wallis, 1973; Beckerman *et al.*, 2009) that have remained isolated in the depth Amazonian forest. Such populations may exhibit low genetic diversity coupled with high frequencies of divergent mtDNA types not found in neighboring human groups, as a result of long-term isolation and small population sizes (Oota *et al.*, 2005). This is particularly true for the Waorani. The genetic variability observed through the analysis of 10 X-chromosome molecular markers (0.569) showed the lowest diversity observed to date in any South American population (Gusmão *et al.*, 2009), and is consistent with previous findings in isolated populations (Zarrabeitia *et al.*, 2009). In contrast to some Amazonian groups such as the Waorani, populations from Gran Chaco have been reported to show a greater history of interpopulation contacts, and therefore, of genetic admixture (Cabana *et al.*, 2006). However, a comparative analysis of the allele distribution of X-chromosome markers between the Waorani and Gran Chaco panmictic populations showed some similarities, especially for marker DXS9898, characterized by the presence of very few alleles. The wide allele distribution of this marker in different South American populations (Gusmão *et al.*, 2009) supports the consistency of the similarities observed between isolated Waorani and native people from Gran Chaco. But even though some nuclear markers may show genetic similarities to other Amerindian groups, the Waorani are characterized by a conspicuously reduced diversity in the lineage markers examined. Thus, analysis of Y-chromosome STR haplotypes seems to confirm the high endogamy levels and the low genetic diversity of this Native American group. Similar results regarding the low genetic diversity of Y-chromosome markers in Wao people were previously reported by González-Andrade *et al.*, 2009, though this study only included 12 Y-STRs. Interestingly, the haplotype diversity estimated for the 47 Wao haplotypes of 17 Y-STRs available at the YHRD (ANs YA003552 and YA003553; Geppert *et al.*, 2011) proved to be  $0.676 \pm 0.052$ . This figure is substantially lower than the haplotype diversity obtained in this study ( $0.836 \pm 0.070$ ), which suggests that much of the genetic variability of the Waorani people might be represented in the sample examined herein, at least at the level of Y chromosome. Beckerman *et al.* (2009) pointed out that less aggressive warriors could have higher indices of reproductive success, in contrast to another well-known warlike South American people, the Yanomamo. Thus, the most common paternal lineages found in the Waorani might corre-

spond to the less belligerent individuals, whose Y-chromosome would therefore be better preserved over the generations. Along these lines, Y-SNPs analysis indicated that all the Waorani males belong to haplogroup Q (under publication). This is the major lineage among the Native Americans (Karafet *et al.*, 2008), with Q-M3 (Q1a3a) being almost completely restricted to the Americas (Zegura *et al.*, 2004). Indeed, a recent publication on the paternal lineages of the Waorani stated that more than 90% of the males belong to subhaplogroup Q1a3a (Geppert *et al.*, 2011).

Mitochondrial DNA showed extremely low diversity among the Waorani. The pan-American haplogroup A2 (Achilli *et al.*, 2008) represented more than 97% of Wao lineages. In fact, only two A2 haplotypes were observed. Bearing in mind that nucleotide position 16111 represents a well known hotspot (Meyer *et al.*, 1999), we assume that the lack of transition C16111T (a diagnostic mutation for haplogroup A2) in ECWA24 could be a back mutation in the native haplotype. Complete sequencing of ECWA24 corroborates haplogroup assignment. Haplogroup A2 has been reported as highly predominant in other indigenous populations such as Inuits from Greenland and Canada, with frequencies of 94.7% and 87.5%, respectively (Helgason *et al.*, 2006), Ngöbes from Panama (63.6%) and Arsarios from Colombia (63.8%) (Tamm *et al.*, 2007). However, in the Cayapa from Ecuador, the geographically closest population to the Waorani for which mtDNA data is available, haplogroup A2 represents only 33.3% of total maternal lineages (Rickards *et al.*, 1999).

Interestingly, the issue of the extremely low genetic diversity of the Waorani is apparent both in maternal and paternal lineages, a phenomenon that could be indicative of relatively similar demographic histories for both sexes. Among Native American groups, the Yanomamo possess similar behavioral characteristics to the Waorani, including fierceness and quarrels within the tribe. Nevertheless, one key difference regarding the sociocultural rules that guide these two belligerent South American peoples might help to understand the causes of the low diversity of mtDNA and Y-chromosome STRs in Waorani. Thus, for instance, among the Yanomamo male deaths from warfare and/or homicide have been estimated to be 5 to 10 times more numerous than female deaths. In contrast, among the Waorani, male deaths (at all ages) resulting from violence, internal fights and revenge raids have been estimated to be only 1.4 times more numerous than female deaths from the same cause (see Beckerman *et al.*, 2009 for a review). This cultural peculiarity of killing women and girls seems to be crucial in explaining the historical failure of the Waorani to grow as a population, and therefore, in explaining the potential fixation of one or few maternal lineages in the Waorani. The results obtained in this study from the analysis of the haploid genome are consistent with previous findings on autosomal (diploid) loci such as polymorphic Alu insertions (Gómez-Pérez *et al.*, 2011) and autosomal STRs (González-Andrade *et al.*, 2007), which also mirrored the genetic singularity and reduced genetic diversity of the Wao population.

As can be inferred, the strikingly low diversity of the mitochondrial genome in the Waorani population is, in all probability, the consequence of intense genetic drift episodes associated with both founder effects during the early settlement and with population bottlenecks. As mentioned above, in such a conflictive sociocultural and psychological context, deaths resulting from internal quarrels and revenge killings could have led to dramatic population bottlenecks, thereby promoting recurrent genetic drift events. In this scenario, some factors would have reinforced the impact of genetic drift: (i) the permanently low demographic size of the population, due to notably high mortality rates until recent times (Yost, 1981; Beckerman *et al.*, 2009); and (ii) the remarkable isolation of the Waorani as a consequence of their

singularity regarding ethnopsychology and language (Peeke, 1973; Wallis, 1973; Robarchek and Robarchek, 1992). It is worth mentioning that at the time of the first peaceful contact with the Waorani, in 1958, the demographic size of the population showed a marked declining trend due to incessant internal warfare, with ~500 people living in an area of ~20 000 km<sup>2</sup>, which represents a population density of around 0.025 person per km<sup>2</sup> (Beckerman *et al.*, 2009).

The main consequence of the genetic drift would have been a minimization of the diversity of maternal lineages, with wide predominance of one or few haplogroups by fixation of all the specific substitutions described herein for subhaplogroup A2. Unfortunately, we were unable to estimate the time to the most recent common ancestor (TMRCA), as the coding region sequences of all the individuals examined were identical. Furthermore, it is not known from what populations the Waorani descend, and therefore, it is not possible to set a reference source population. Despite of the limitations in estimating TMRCA, it seems obvious that the Waorani root haplotype originated thousands of years ago, bearing in mind the number of polymorphic positions accumulated. What remains unclear is when the Waorani appeared as a tribe, as episodes of violence and population bottlenecks, among other factors, hinder this calculation. Further studies on mtDNA complete genomes, especially from the scarcely analyzed regions of Central and South America, might help to find the intermediate lineages that most likely gave rise to the Wao lineage from the root haplotype of A2. What seems to be quite clear from our findings on mitochondrial genome diversity is the existence of a 'Waorani-specific' maternal lineage, which we have designated A2s.

## CONCLUSIONS

The extremely low genetic diversity observed in the Waorani population might be explained by genetic drift events promoted by founder effects and a historically low population size probably caused by their warlike customs, in which homicides were usually followed by revenge raids within the tribe (Beckerman *et al.*, 2009). The findings of our study constitute a good example of how ethnopsychology and sociocultural features may have influenced the genetic make-up of human populations. To reinforce the singularity of the Waorani community, results obtained for both the analysis of the mtDNA control region (HVS-I/HVS-II) and the sequencing of the complete mitochondrial genome suggest the existence of a Waorani-specific maternal lineage. According to current knowledge on the phylogeny of haplogroup A2 (van Oven and Kayser, 2009), this mitochondrial lineage could be designated as subhaplogroup A2s, the maternal legacy of a history of warfare, revenge raids and consanguineous marriages led by the Waorani, the last human isolate in the Ecuadorian Amazon.

## DATA ARCHIVING

Sequence data have been submitted to GenBank: accession numbers GQ398432-GQ398479 and HQ709171-HQ709172 (25 HVSI-HVSI sequences), and GQ398480-GQ398490 (11 complete genomes).

Data have also been deposited at Dryad: doi:10.5061/dryad.f6hs62hq.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

Funds were provided by the Basque Government (IT424-07). Technical and human support provided by SG Iker (UPV/EHU, MICINN, GV/EJ, ERDF and

ESF) is gratefully acknowledged. LV receives a grant from the University of the Basque Country. The authors are deeply indebted to the Waorani people who voluntarily participated in this study.

- Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, Woodward SR *et al.* (2008). The phylogeny of the four pan-American mtDNA haplogroups: implications for evolutionary and disease studies. *PLoS One* **3**: e1764.
- Alfonso-Sánchez MA, Cardoso S, Martínez-Bouzas C, Peña JA, Herrera RJ, Castro A *et al.* (2008). Mitochondrial DNA haplogroup diversity in Basques: a reassessment based on HVI and HVII polymorphisms. *Am J Hum Biol* **20**: 154–164.
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* **23**: 147.
- Beckerman S, Erickson PI, Yost J, Regalado J, Jaramillo L, Sparks C *et al.* (2009). Life histories, blood revenge, and reproductive success among the Waorani of Ecuador. *Proc Natl Acad Sci USA* **106**: 8134–8139.
- Bland JM, Altman DG (1995). Multiple significance tests: the Bonferroni method. *BMJ Statistics Notes* **310**: 170.
- Boster JS, Yost J, Peeke C (2004). Rage, revenge, and religion: Honest signaling of aggression and nonaggression in Waorani coalitional violence. *Ethos* **31**: 471–494.
- Cabana GS, Merriwether DA, Hunley K, Demarchi DA (2006). Is the genetic structure of Gran Chaco populations unique? Interregional perspectives on native South American mitochondrial DNA variation. *Am J Phys Anthropol* **131**: 108–119.
- Catanesi CI, Martina PF, Giovambattista G, Zukas P, Vidal-Rioja L (2007). Geographic structure in Gran Chaco Amerindians based on five X-chromosome STRs. *Hum Biol* **79**: 463–474.
- CODENPE (2002). *Sistema de Indicadores de Nacionalidades y Pueblos indígenas del Ecuador (SIDENPE)*. [http://www.codenpe.gov.ec/].
- Excoffier L, Laval G, Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform* **1**: 47–50.
- Felsenstein J (1989). PHYLIP: Phylogeny inference package (version 3.2). *Cladistics* **5**: 164–166.
- Geppert M, Baeta M, Núñez C, Martínez-Jarreta B, Zweynert S, Cruz OW *et al.* (2011). Hierarchical Y-SNP assay to study the hidden diversity and phylogenetic relationship of native populations in South America. *Forensic Sci Int Genet* **5**: 100–104.
- Gómez-Pérez L, Alfonso-Sánchez MA, Sánchez D, García-Obregón S, Espinosa I, Martínez-Jarreta B *et al.* (2011). 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.
- González-Andrade F, Roewer L, Willuweit S, Sánchez D, Martínez-Jarreta B (2009). Y-STR variation among ethnic groups from Ecuador: Mestizos, Kichwas, Afro-Ecuadorians and Waoranis. *Forensic Sci Int Genet* **3**: e83–e91.
- González-Andrade F, Sánchez D, González-Solórzano J, Gascón S, Martínez-Jarreta B (2007). Sex-specific genetic admixture of Mestizos, Amerindian Kichwas, and Afro-Ecuadorians from Ecuador. *Hum Biol* **78**: 51–78.
- Gusmão L, Sánchez-Diz P, Alves C, Gomes I, Zarrabeitia MT, Abovich M *et al.* (2009). A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations. *Int J Legal Med* **123**: 227–234.
- Helgason A, Pálsson G, Pedersen HS, Angulalik E, Gunnarsdóttir ED, Yngvadóttir B *et al.* (2006). mtDNA variation in Inuit populations of Greenland and Canada: migration history and population structure. *Am J Phys Anthropol* **130**: 123–134.
- Ingman M, Gyllenstein U (2006). mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. *Nucleic Acids Res* **34**: D749–D751.
- Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008). New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* **18** (5): 830–838.
- Kivisild T, Shen P, Wall DP, Do B, Sung R, Davis K *et al.* (2006). The role of selection in the evolution of human mitochondrial genomes. *Genetics* **172**: 373–387.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**: 2947–2948.
- Larrick JW, Yost J, Gourley C, Buckley 3rd CE, Plato CC, Pandey JP *et al.* (1985). Markers of genetic variation among the Waorani Indians of the Ecuadorian Amazon headwaters. *Am J Phys Anthropol* **66**: 445–453.
- Meyer S, Weiss G, von Haeseler A (1999). Pattern of nucleotide substitution and rate heterogeneity in the hypervariable regions I and II of human mtDNA. *Genetics* **152**: 1103–1110.
- Oota H, Pakendorf B, Weiss G, von Haeseler A, Pookajorn S, Settheetham-Ishida W *et al.* (2005). Recent origin and cultural reversion of a hunter-gatherer group. *PLoS Biol* **3**: e71.
- Peeke C (1973). *Preliminary grammar of Auca*. The Summer Institute of Linguistics: Norman, Oklahoma.
- Raymond F, Rousset M (1995). Testing heterozygote excess and deficiency. *Genetics* **140**: 1413–1419.
- Reynolds J, Weir BS, Cockerham CC (1983). Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**: 767–779.

- Rickards O, Martínez-Labarga C, Lum JK, De Stefano GF, Cann RL (1999). mtDNA history of the Cayapa Amerinds of Ecuador: detection of additional founding lineages for the Native American populations. *Am J Hum Genet* **65**: 519–530.
- Robarchek CA, Robarchek CJ (1992). Cultures of war and peace: a comparative study of Waorani and Semai. In: Silverberg J, Gray P (eds). *Aggression and peacefulness in humans and other primates*. Oxford University Press: Oxford. pp 189–213.
- Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ *et al*. (2007). Beringian standstill and spread of Native American founders. *PLoS One* **2**: e829.
- van Oven M, Kayser M (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* **30**: E386–E394.
- Wallis EE (1973). *Aucas downriver*. Harper and Row: New York.
- Willuweit S, Roewer L (2007). Y chromosome haplotype reference database (YHRD): Update. *Forensic Sci Int Genet* **1**: 83–87.
- Yost JA (1981). Twenty years of contact: The mechanisms of change in Huao (Auca) culture. In: Whitten NA (eds). *Cultural transformations and ethnicity in modern Ecuador*. University of Illinois Press: Urbana. pp 677–704.
- Zarrabeitia MT, Pinheiro F, de Pancorbo MM, Cainé L, Cardoso S, Gusmão L *et al*. (2009). Analysis of 10 X-linked tetranucleotide markers in mixed and isolated populations. *Forensic Sci Int Genet* **3**: 63–66.
- Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF (2004). High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol Biol Evol* **21**: 164–175.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)