Paternal Genetic History of the Basque Population of Spain

KRISTIN L. YOUNG,[1]* GUANGYUN SUN,[2] RANJAN DEKA,[2] AND MICHAEL H. CRAWFORD[3]

Abstract This study examines the genetic variation in Basque Y chromosome lineages using data on 12 Y-short tandem repeat (STR) loci in a sample of 158 males from four Basque provinces of Spain (Alava, Vizcaya, Guipuzcoa, and Navarre). As reported in previous studies, the Basques are characterized by high frequencies of haplogroup R1b (83%). AMOVA analysis demonstrates genetic homogeneity, with a small but significant amount of genetic structure between provinces (Y-short tandem repeat loci STRs: 1.71%, $p = 0.0369$). Gene and haplotype diversity levels in the Basque population are on the low end of the European distribution (gene diversity: 0.4268; haplotype diversity: 0.9421). Post-Neolithic contribution to the paternal Basque gene pool was estimated by measuring the proportion of those haplogroups with a Time to Most Recent Common Ancestor (TMRCA) previously dated either prior (R1b, I2a2) or subsequent to (E1b1b, G2a, J2a) the Neolithic. Based on these estimates, the Basque provinces show varying degrees of post-Neolithic contribution in the paternal lineages (10.9% in the combined sample).

The debate concerning the genetic composition of the modern European gene pool is currently a matter of degree. Did the Paleolithic inhabitants of Europe contribute to the modern gene pool? If so, how much? Or are modern Europeans the result of a replacement of the Paleolithic groups by technologically advanced Neolithic farmers? Two models have been proposed to explain the patterns of gene distribution found in Europe. The demic diffusion model (DDM) states that the majority of genetic variation present in modern Europeans is the result of Neolithic farmers spreading their technology (and genes) into Europe with the advent of agriculture (Ammerman and Cavalli-Sforza 1984). The cultural diffusion model (CDM) posits that while technology spread into Europe 10,000 years ago, people did not, leaving the Paleolithic gene pool largely intact

[1]Department of Family Medicine, Research Division, University of Kansas Medical Center, Kansas City, KS 66160.

[2]Department of Environmental Health, University of Cincinnati Medical Center, Cincinnati, OH 45267.

[3]Department of Anthropology, University of Kansas, Lawrence, KS 66045.

*Correspondence to: Kristin L. Young, Department of Family Medicine, Research Division, University of Kansas Medical Center, 3901 Rainbow Boulevard, Mailstop 3064, Kansas City, KS 66160. E-mail: kyoung@kumc.edu.

Human Biology, August 2011, v. 83, no. 4, pp. 455–475.

KEY WORDS: SPAIN, BASQUES, Y-CHROMOSOME, Y-SHORT TANDEM REPEAT LOCI, HAPLOTYPES, POPULATION GENETICS.

(Novelletto 2007). Genetic evidence is used in support of both models. Synthetic gene map analyses based on classical markers exhibit a cline for several loci in the first principal component, accounting for approximately 27% of the total variation and spreading from southeast to northwest through Europe (Cavalli-Sforza et al. 1994). This cline has been interpreted as a genetic signature of the DDM model, with a correlation of 0.89 between the first principal component of gene frequencies and temporal spread of agriculture into Europe (Ammerman and Cavalli-Sforza 1984). Similar clines are reported for other autosomal markers, HLA-DQA plus six short tandem repeat loci (STRs), indicating "directional population expansion" (Chikhi et al. 1998: 9055). Analysis of Y-chromosome STRs produce comparable results, with clinal variation in the frequencies of 12 of 27 alleles revealing maximum effective divergence times (τ_{max}) ranging from 281–10,296 years, well within the Neolithic (Casalotti et al. 1999). This suggests that most of the variation present in the European Y-chromosome dates to this period, but research in Southeastern Europe suggests that this Neolithic expansion includes endogenous European haplogroups (Battaglia et al. 2009).

In support of the CDM hypothesis, evidence comes from paleoanthropology as well as genetic analyses. Advocates note that the Paleolithic expansion into Europe occurred from the same area as the Neolithic expansion, implying that the gradient seen in some classical markers might instead be a Paleolithic signal (Barbujani et al. 1998). In the Basque region, there is strong archaeological evidence of Paleolithic human occupation (Bertranpetit et al. 1995). Analysis of prehistoric skeletal remains in Iberia show little transition in cranial morphology or evidence of dental caries between the Mesolithic and Neolithic, a transition which would be expected if the Mesolithic hunter/gatherers had been replaced by farmers with different dietary patterns (Jackes et al. 1997). Y-chromosome analyses reveal one haplotype (R1*M173), which appears to show evidence of expansion after the Last Glacial Maximum (Wells et al. 2001) and a "high degree of non-Neolithic ancestry" in populations of Iberia (Flores et al. 2004). Studies of Y-chromosome haplogroups among the Basques have demonstrated a high frequency of R1b, the most common haplogroup in Western Europe (Alonso et al. 2005; Calderon et al. 2003; Flores et al. 2004; Lucotte and Hazout 1996; Lucotte and Loirat 1999; Quintana-Murci et al. 1999). Examination of Y-haplogroup diversity in Southeastern Europe demonstrated that the timing of the spread of agriculture in the region overlaps with the expansion of European Haplogroup I-M423, suggesting a cultural diffusion of agricultural technologies by autochthonous groups (Battaglia et al. 2009). Analysis of mitochondrial HVS-I sequences in Europe present little clinal variation, and divergence dates suggest that many of the mitochondrial haplotypes in Europe have pre-Neolithic origins (Richards et al. 1996). Haplogroup V has been proposed as a signal of population expansion after the Last Glacial Maximum (LGM), but the absence of this haplogroup in an ancient sample from the Basque region called this hypothesis into question (Izagirre and de la Rua 1999; Torroni et al. 1998). In addition, the presence of 15 individuals in the same sample belonging to

mitochondrial haplogroup J, which has been linked to the expansion of Neolithic groups into Europe, suggests some level of admixture.

In the biological anthropology literature, the Basque population has been studied using craniometrics (Broca 1864; Calafell and Bertranpetit 1994), linguistics (Trask 1997), blood group antigens (Boyd and Boyd 1937; Chalmers et al. 1948; Chalmers et al. 1949; Etcheverry 1945; Levine 1977; Levine et al. 1974; MacClancy 1993; Mourant 1948), erythrocytic enzymes (Aguirre et al. 1989; Aguirre et al. 1991; Manzano et al. 1996), plasma proteins (Manzano et al. 1993), HLA antigens and haplotypes (Calderon et al. 1998; Dugoujon et al. 1989; Esteban et al. 1998; Hazout et al. 1991), autosomal STRs (Alonso et al. 1995; Arrieta et al. 1997; Garcia et al. 1998; Garcia et al. 2001; Iriondo et al. 1997; Iriondo et al. 1999; Perez-Lezaun et al. 2000; Perez-Miranda et al. 2005a; Perez-Miranda et al. 2005b; Zlojutro et al. 2006), mitochondrial haplogroups and sequences (Alfonso-Sanchez et al. 2008; Alzualde et al. 2006; Alzualde et al. 2005; Bertranpetit et al. 1995; Corte-Real et al. 1996; Garcia et al. 2011; Gonzalez et al. 2003; Gonzalez et al. 2006; Izagirre and de la Rua 1999), and Y-chromosome markers (Adams et al. 2008; Alonso et al. 2005; Bosch et al. 1999; Calderon et al. 2003; Flores et al. 2004; Garcia et al. 2004; Gonzalez-Neira et al. 2000; Lucotte and Hazout 1996; Quintana-Murci et al. 1999; Rosser et al. 2000; Shi et al. 2010). The Basque language, *Euskara*, is most widely accepted as an isolate, unrelated to any other extant language in Europe. Genetically, the Basques are outliers in the European distribution for several classical markers, including blood groups ABO, Rhesus, and MNS, erythrocytic enzyme adenylate kinase (AK), and immunoglobulin GM (Izagirre et al. 2001). Given the linguistic and genetic distinctiveness of the Basques in Europe, and their relative isolation and distance from the epicenter of agricultural development during the Neolithic (beginning about 8500 BC), they seem an ideal population in which to explore the DDM and CDM models.

The Y-chromosome provides a singular perspective on questions of population history. Y markers are paternally inherited, do not undergo recombination, and reflect male lineages and migration. Despite a high mutation rate and consequent homoplasy, Y-STR haplotype variation appears mostly compartmentalized or is restricted to particular SNP backgrounds (haplogroups) (Bosch et al. 1999). Y-STR data have also been found useful for studies of interpopulation heterogeneity, with STR haplotype grouping demonstrating geographic clines in Europe (Gusmao et al. 2003). This investigation examines the paternal genetic structure of the Basque population of Spain using Y-STR haplotypes. Frequencies of haplogroups which have been identified as having a TMRCA dating subsequent to the Neolithic (Balaresque et al. 2010; Capelli et al. 2007; Chikhi et al. 2002; Martinez et al. 2007; Semino et al. 2000a; Semino et al. 2000b) are also estimated from the Y-STR haplotypes to measure their relative contribution to the modern Basque gene pool. Heterogeneity within the Basques is explored to determine whether the provinces have undergone different rates of potential admixture since the Neolithic.

Figure 1. Map of Basque provinces in Spain and France, with locations of provincial capitals highlighted.

Materials and Methods

Sampling. During summer field sessions between 2000 and 2002, DNA samples were collected in mountain villages throughout the four Basque provinces of northern Spain (Figure 1). Buccal samples were collected from 158 adult male autochthonous participants (those who claimed four Basque grandparents). Six villages were sampled in Alava ($n = 41$), 17 villages were sampled in the province of Vizcaya ($n = 46$), 10 villages were sampled in Guipuzcoa ($n = 58$), and two villages were sampled in Navarre ($n = 13$). Of the 158 male samples collected in the survey, 126 amplified without allelic dropout (Pawlowski and Maciejewska 2000) and were included in further analyses. This study was approved by the University of Kansas Human Subjects Committee (HSCL #11,955), and participants provided written informed consent.

DNA Analysis. DNA extraction was performed using a standard phenol:chloroform protocol. Y-STR profiles were characterized using the Y-Plex™ 12 kit (Reliagene Technologies, Inc., New Orleans, Louisiana) to type 11 Y-STR loci (detailed in Table 1) plus the sex-determining amelogenin locus, according to

Table 1. Y-Chromosome STRs Used in the Present Analysis

STR Locus	*Repeat*	*Allele Range*
DYS392	TAT	6–18
DYS390	TCTA/TCTG	17–28
DYS385a/b	GAAA	7–25
DYS393	AGAT	8–17
DYS389I	TCTG/TCTA	10–17
DYS389II	TCTG/TCTA	24–34
DYS391	TCTA	6–14
DYS19	TAGA	10–19
DYS439	GATA	8–15
DYS438	TTTTC	6–14
Amelogenin	X, Y	—

manufacturer's instructions. Amplified products were detected using an ABI 377 DNA sequencer. Fragments were sized and genotyped using GeneScan 3.1 and Genotyper 2.5 software (Applied Biosystems).

Comparative Population Data. Y-chromosome STR allele frequency and haplotype data on 36 additional European populations, as listed in Table 5, were compiled from the literature for comparative purposes: other Basque populations (4) Ireland; Spain (8) Portugal (2); France; Estonia; Italy (3) Belgium; Sweden Germany; Romania; Greece; Bulgaria; Hungary; Latvia; Lithuania; Serbia; Russia; (2) Poland; Croatia; Bosnia. Because not all groups had been typed for the 11 Y-STRs examined in the study population, the seven loci common to all populations (DYS*19, DYS*389I, DYS*389II, DYS*390, DYS*391, DYS*392, DYS*393) were used in the interpopulation analyses.

Statistical Analysis. Allele frequencies and haplotype diversity were calculated according to the methods of Nei (1987). Y haplogroups were determined from Y-STR haplotypes, using a Bayesian approach with NW Europe priors to estimate the probability of assignment to a particular haplogroup (Athey 2005, 2006). To test the efficacy of the haplogroup predictor algorithm, 116 Basque males typed by Adams et al. (2008) with both SNPs and STRs were used to estimate an error rate assuming the same NW Europe priors. Each Basque Y-STR haplotype was entered into the predictor using the same 11 STR loci typed in the present study, and the resulting predicted haplogroup was compared to the haplogroup determined by SNP analysis (Adams et al. 2008). Of the 116 haplotypes measured, only one could not be unambiguously predicted. IP889, listed as J2 in Adams et al. (2008), was given a 43.2% probability of J2 and a 50.6% probability of L. Based on these data, the error rate for the haplogroup predictor algorithm was estimated at 0.86%. For this population, the predictor algorithm appears to be a reliable method of estimating haplogroups from Y-STR haplotypes.

Population structure between Basque provinces was explored using analysis of molecular variance (AMOVA) with Arlequin 3.5 (Excoffier and Lischer 2010). Y-STR haplotype data were used to measure genetic distances between populations with Slatkin's R_{ST} (1995). Relationships between populations were visualized using an Multidimensional Scaling (MDS) plot in NTSYSpc 2.1 (Rohlf 2005). Genetic barriers were inferred using the SAMOVA algorithm, which establishes groupings of populations that maximize F_{CT} (Dupanloup et al. 2002). Post-Neolithic contribution to the paternal Basque gene pool was estimated by measuring the proportion of those haplogroups with a TMRCA dated subsequent to the Neolithic (E1b1b, G2a, J2a) (Capelli et al. 2007; Cinnioglu et al. 2004; Rosser et al. 2000; Semino et al. 2004).

Results

Haplotype frequencies, haplogroup assignments, as well as allele frequency goodness-of-fit and probability for the 89 Y-STR haplotypes found in the four Basque Provinces are presented in Table 2. While several haplotypes are shared between provinces (H1 is found in Alava, Guipuzcoa, and Navarre), 92% of the haplotypes are unique to a single province. All but two haplotypes were assigned to a single haplogroup with greater than 95% probability. Seventy-four of the 89 haplotypes (84.1%) are R1b. Six other haplogroups were identified in the Basque sample: E1b1b (5.6%), J2a (4.0%), I2 (3.2%), G2a (1.5%), L (0.8%), and T (0.8%). Haplogroup frequencies by province are shown in Table 3. Only 8% of haplotypes were found in more than one province, and the AMOVA analysis of provinces (Table 4) shows that while a small but significant amount of variation (1.71%, $p = 0.0369$) is accounted for between provinces, 98.29% of the total variation is found between males within provinces.

Gene and haplotype diversity values are presented in Table 5. Gene diversity for the Basque populations range from 0.3966, the lowest value in the overall group, to 0.5137 (Brion et al. 2003; Garcia et al. 2004). The study population has a gene diversity value of 0.4268. The highest gene diversity is found in the Romanian sample (0.6735). Among the Basques, haplotype diversity ranges from 0.9148, again the lowest value in the overall group, to 0.9631 (excluding the mixed Basque residents sample). The haplotype diversity of the study population is 0.9421. The highest haplotype diversity is found in Greece (0.9983).

The first axis of the MDS plot (Figure 2) shows a separation of Eastern and Western European populations. The first two axes account for 94.90% of the total variation in the sample (I: 75.98%; II: 18.92%), and the STRESS=0.11528 ($p < 0.01$) and matrix correlation ($r = 0.95688$) indicate a good fit with the original distance matrix. The Western European populations radiate from the upper left corner of the plot, where the population of Ireland and the Basque groups cluster. SAMOVA results confirm the East–West separation (represented by the black line on the MDS plot), with the highest F_{CT} value (0.06313) seen when the populations are divided into two groups.

Table 2. Basque Haplogroup Identification from Haplotype Definition, with Goodness of Fit and Probability

Haplotype	*Definition*[a]	*n*	*Haplogroup*	*Goodness of Fit*	*Probability*
H1	14-14-16-24-11-13-13-11-14-12-11	7	R1b	77	100.0%
H2	14-14-16-23-9-11-13-12-17-10-12	3	J2a	97	95.0%
H3	13-12-17-24-10-11-12-17-18-10-11	3	E1b1b	44	100.0%
H4	14-13-17-24-11-13-13-12-14-12-11	2	R1b	62	100.0%
H5	14-13-16-24-11-13-13-11-14-12-11	4	R1b	91	100.0%
H6	14-12-16-24-11-13-13-11-14-12-12	2	R1b	78	100.0%
H7	15-13-17-23-11-13-13-11-16-12-12	1	R1b	49	99.9%
H8	15-13-16-24-12-13-13-11-15-12-11	1	R1b	51	100.0%
H9	14-14-17-23-10-13-13-11-11-12-13	1	R1b	38	100.0%
H10	14-14-16-24-11-13-13-12-14-12-11	1	R1b	61	100.0%
H11	14-14-16-24-11-13-13-11-15-12-12	1	R1b	75	100.0%
H12	14-14-16-24-11-13-13-11-14-12-12	2	R1b	85	100.0%
H13	14-14-16-24-11-13-13-11-13-12-12	1	R1b	73	100.0%
H14	14-14-16-24-10-13-13-11-14-12-12	1	R1b	78	100.0%
H15	14-13-16-25-12-13-13-11-14-12-12	1	R1b	69	100.0%
H16	14-13-16-25-11-13-13-11-14-12-11	1	R1b	81	100.0%
H17	14-13-16-24-11-13-13-12-14-12-11	2	R1b	72	100.0%
H18	14-13-16-24-11-13-13-11-14-12-13	4	R1b	88	100.0%
H19	14-13-16-24-11-13-13-11-14-12-12	10	R1b	100	100.0%
H20	14-13-16-24-10-13-13-11-15-12-12	1	R1b	81	100.0%
H21	14-13-15-24-11-13-13-11-14-12-14	1	R1b	59	100.0%
H22	14-12-17-24-11-13-13-11-14-12-11	1	R1b	61	100.0%
H23	14-12-16-24-10-13-13-11-14-12-12	1	R1b	71	100.0%
H24	14-14-16-23-10-13-13-11-11-12-13	3	R1b	44	100.0%
H25	16-13-17-23-9-12-12-13-13-9-12	1	J2a	46	99.5%
H26	15-13-16-24-9-11-12-13-15-10-10	1	J2a	100	96.1%
H27	15-13-16-24-11-13-13-12-14-12-11	1	R1b	59	99.9%
H28	15-13-16-24-11-13-13-11-15-12-12	1	R1b	72	100.0%
H29	14-14-16-24-11-13-13-11-15-12-11	1	R1b	68	100.0%
H30	14-14-16-24-10-13-13-11-14-12-11	2	R1b	71	100.0%
H31	14-14-16-23-10-13-13-11-11-12-14	1	R1b	37	100.0%
H32	14-14-15-25-11-13-13-11-14-12-12	1	R1b	61	100.0%
H33	14-13-17-24-11-13-13-11-15-12-12	1	R1b	76	100.0%
H34	14-13-17-23-11-13-13-11-15-12-13	1	R1b	61	100.0%
H35	14-13-16-25-10-13-12-11-13-12-12	1	R1b	58	100.0%
H36	14-13-16-24-11-14-13-11-14-12-12	1	R1b	84	100.0%
H37	14-13-16-24-11-13-13-12-14-12-12	1	R1b	80	100.0%
H38	14-13-16-24-11-13-13-11-11-12-12	1	R1b	71	100.0%
H39	14-13-16-24-11-13-12-11-14-13-12	1	R1b	69	100.0%
H40	14-13-16-24-10-13-13-11-13-12-12	1	R1b	79	100.0%
H41	14-13-16-23-11-13-13-11-16-12-13	1	R1b	61	100.0%
H42	14-13-15-24-11-13-13-11-15-14-12	1	R1b	49	100.0%
H43	13-12-17-25-10-11-12-17-18-10-11	1	E1b1b	39	99.9%
H44	14-13-16-24-11-13-14-11-14-12-12	3	R1b	76	100.0%
H45	17-13-15-24-9-11-13-12-12-10-13	2	I2a2	36	95.6%
H46	15-14-17-24-11-15-13-11-14-12-12	2	R1b	41	99.4%
H47	14-13-16-24-10-13-13-11-14-12-12	2	R1b	92	100.0%
H48	17-13-16-23-11-12-13-14-16-10-11	1	I2	63	93.6%

Table 2. (*continued*)

Haplotype	*Definition*[a]	*n*	*Haplogroup*	*Goodness of Fit*	*Probability*
H49	15-14-17-23-10-14-11-13-17-10-13	1	L	78	100.0%
H50	15-14-16-24-11-13-13-12-14-12-12	1	R1b	56	100.0%
H51	15-13-17-24-11-13-13-12-15-12-11	1	R1b	45	91.1%
H52	15-13-16-24-11-13-13-11-15-12-13	1	R1b	63	100.0%
H53	15-13-16-24-10-13-13-11-14-12-12	1	R1b	75	100.0%
H54	15-13-16-23-11-13-13-11-15-12-13	1	R1b	58	100.0%
H55	15-12-19-22-10-11-14-14-14-10-11	1	G2a	55	100.0%
H56	15-12-18-22-10-10-14-15-16-10-12	1	G2a	33	99.9%
H57	15-12-15-24-10-14-14-13-15-9-12	1	T	42	100%
H58	15-12-15-23-10-11-13-12-12-10-12	1	I2a2	42	97.0%
H59	14-15-16-24-11-13-13-11-15-12-11	1	R1b	49	100.0%
H60	14-14-18-24-10-13-13-11-14-12-12	1	R1b	55	100.0%
H61	14-14-16-24-11-13-13-11-15-13-11	1	R1b	58	100.0%
H62	14-14-16-24-10-13-13-11-11-12-12	1	R1b	55	100.0%
H63	14-13-17-25-10-13-13-11-14-12-12	1	R1b	71	100.0%
H64	14-13-17-24-11-13-13-10-14-12-12	1	R1b	65	100.0%
H65	14-13-16-25-11-13-13-11-14-12-12	1	R1b	90	100.0%
H66	14-13-16-25-10-13-12-11-13-12-13	1	R1b	51	100.0%
H67	14-13-16-24-11-13-14-12-14-12-12	1	R1b	61	100.0%
H68	14-13-16-24-11-13-13-11-15-12-12	1	R1b	88	100.0%
H69	14-13-16-24-11-13-13-11-14-11-11	1	R1b	77	100.0%
H70	14-13-16-24-11-12-13-11-14-12-12	1	R1b	71	100.0%
H71	14-13-16-24-10-13-13-12-14-12-11	1	R1b	66	100.0%
H72	14-13-16-23-11-13-13-11-14-12-11	1	R1b	83	100.0%
H73	14-13-16-23-10-13-13-11-11-12-13	1	R1b	52	100.0%
H74	14-13-15-24-11-13-13-12-14-12-12	1	R1b	64	100.0%
H75	14-13-15-23-11-13-13-11-14-12-12	1	R1b	73	100.0%
H76	14-12-16-24-11-13-13-11-14-12-11	1	R1b	70	100.0%
H77	14-12-16-23-11-13-14-11-14-12-13	1	R1b	47	100.0%
H78	14-11-16-25-11-13-13-11-13-12-11	1	R1b	43	100.0%
H79	13-12-18-25-11-11-13-17-18-10-11	1	E1b1b	38	100.0%
H80	13-12-18-25-10-11-13-17-18-10-11	1	E1b1b	47	100.0%
H81	15-14-17-24-10-14-13-11-14-12-12	1	R1b	46	99.5%
H82	14-14-17-24-10-13-13-11-14-12-11	1	R1b	61	100.0%
H83	14-14-17-24-10-11-13-17-19-10-14	1	E1b1b	37	100.0%
H84	14-14-16-24-11-13-14-11-14-12-12	1	R1b	64	100.0%
H85	14-14-16-24-11-13-13-12-14-12-13	1	R1b	59	100.0%
H86	14-14-16-24-10-13-13-11-13-12-11	1	R1b	61	100.0%
H87	14-14-15-24-11-13-13-12-14-12-12	1	R1b	54	100.0%
H88	14-13-16-24-11-13-12-12-14-12-11	1	R1b	59	100.0%
H89	13-14-16-24-9-11-13-13-14-10-10	1	E1b1b	35	99.9%

a. Haplotype definition locus order: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385a, DYS385b, DYS438, DYS439.

Discussion

This study demonstrates the homogeneity of the Basque population in terms of Y chromosome lineages, with the majority of variation found between individuals within provinces. Comparison with other European populations also

Table 3. Y-Chromosome Haplogroups by Province

Province	*R1b*	*J2a*	*E1b1b*	*G2a*	*I2*	*L*	*T*
Alava ($n = 32$)	26	3	3	0	0	0	0
Vizcaya ($n = 25$)	22	2	1	0	0	0	0
Guipuzcoa ($n = 58$)	48	0	2	2	4	1	1
Navarre ($n = 11$)	10	0	1	0	0	0	0
Total ($n = 126$)	106	5	7	2	4	1	1

confirms their uniqueness from other Iberian groups, contrary to recent genome-wide SNP analysis, likely underpowered or faulty because of biased marker selection (Laayouni and Bertranpetit 2009), but in agreement with other, more densely scanned, genomic studies (Nothnagel et al. 2010; Rodriguez-Ezpeleta et al. 2010).

Previous analysis of Y-haplogroups in the Basque country found 86% R1b (Alonso et al. 2005), which is consistent with the present study, where 84% of the Y-chromosomes were identified as this most common Western European haplogroup. Prior research has also reported the presence of haplogroups E, G2, Q, and I in the Basque population and noted the absence of J in Guipuzcoa (Adams et al. 2008; Brion et al. 2003; Cruciani et al. 2004; Flores et al. 2004; Rootsi et al. 2004). The present analysis found haplogroup E1b1b in all four provinces and haplogroups G2a (3.44%) and I2 (6.90%) in Guipuzcoa. Haplogroup J was absent in Guipuzcoa, though J2a was detected in both Alava (9.38%) and Vizcaya (8%). In addition, Guipuzcoa also had low frequencies of haplogroups L (1.72%), and T (1.72%), which have not been reported previously among the Basques. It is possible that these individuals represent mixed Basque descent; that haplogroups L and T are found at low frequencies throughout Southern Europe, including among the Basques; or that these haplotypes were misassigned by the haplogroup predictor algorithm and are actually a rare haplotype in a more common European haplogroup.

Two haplogroups that have been considered markers of non-Neolithic genetic contribution (R1b and I2a2) are found at appreciable frequencies (>5%) in the present Basque population. Haplogroup R1b shows the highest frequencies in Western European populations, including France (52.2%), the Netherlands (70.4%), Germany (50%), Italy (62%), and Britain (68.8%) (Semino et al. 2000a). In the British Isles, this haplogroup is found in 79–82% of males, while in Iberia it ranges from 56% in Portugal to 68% in Spain (Rosser et al. 2000). Also found in haplogroup R1b is the Atlantic Modal Haplotype (AMH)—

Table 4. AMOVA Based on Y-STRs in Four Basque Provinces

Source	*Sum of Squares*	*Variance*	*% Variation*	*p*
Among provinces	11.229	0.044	1.71	0.0369
Within provinces	305.612	2.505	98.29	—
Total	316.841	2.549		

Table 5. Gene and Haplotype Diversity (Standard Error) Values for 37 Populations Based on Y-STR Data

Population	*n*	*Gene Diversity*	*Haplotype Diversity*	*Reference*
Basque	128	0.4268	0.9421 ± 0.0127	Present study
Andalusia	224	0.5387	0.9806 ± 0.0042	Adams et al. (2008); $n = 168$ Gonzalez-Neira et al. (2000); $n = 56$
Aragon	34	0.5990	0.9943 ± 0.0090	Adams et al. (2008)
Asturias	20	0.5692	0.9789 ± 0.0245	Adams et al. (2008)
Basque	29	0.5137	0.9631 ± 0.0204	Brion et al. (2003)
Basque	167	0.3966	0.9148 ± 0.0149	Garcia et al. (2004)
Basque Country	116	0.4135	0.9465 ± 0.0110	Adams et al. (2008)
Basque Residents	60	0.4822	0.9791 ± 0.0079	Pena et al. (2006)
Belgium	113	0.5830	0.9771 ± 0.0073	Mertens et al. (2007)
Bosnia	181	0.4621	0.9734 ± 0.0054	Klaric et al. (2005)
Bulgaria	126	0.5604	0.9874 ± 0.0047	Zaharova et al. (2001)
Castilla la Mancha	63	0.5226	0.9794 ± 0.0089	Adams et al. (2008)
Castile	131	0.5343	0.9799 ± 0.0059	Adams et al. (2008)
Catalonia	80	0.4688	0.9680 ± 0.0114	Adams et al. (2008)
Croatia	166	0.5317	0.9801 ± 0.0048	Ljubkovic et al. (2008)
Estonia	133	0.5968	0.9869 ± 0.0038	Lessig et al. (2001)
Extremadura	52	0.5677	0.9864 ± 0.0070	Adams et al. (2008)
France	100	0.5412	0.9834 ± 0.0052	Keyser-Tracqui et al. (2003)
Galicia	141	0.5630	0.9834 ± 0.0049	Brion et al. (2003); $n = 53$ Adams et al. (2008); $n = 88$
Germany	439	0.5907	0.9918 ± 0.0012	Ploski et al. (2002)
Greece	69	0.6662	0.9983 ± 0.0028	Parreira et al. (2002)
Hungary	116	0.6441	0.9963 ± 0.0017	Furedi et al. (1999)
Ireland	151	0.4636	0.9679 ± 0.0075	Ballard et al. (2006)
Latvia	145	0.5841	0.9907 ± 0.0027	Lessig et al. (2001)
Lithuania	152	0.5814	0.9884 ± 0.0031	Lessig et al. (2001)
Moscow	85	0.5687	0.9776 ± 0.0079	Ploski et al. (2002)
North Italy	104	0.5644	0.9733 ± 0.0095	Cerri et al. (2005)
North Portugal	115	0.5443	0.9822 ± 0.0064	Gonzalez-Neira et al. (2000); $n = 55$ Adams et al. (2008); $n = 60$
Piedmont Italy	233	0.5914	0.9771 ± 0.0037	Cerutti et al. (2006)
Poland	919	0.5428	0.9879 ± 0.0012	Ploski et al. (2002)
Romania	91	0.6735	0.9941 ± 0.0027	Egyed et al. (2006)
Rome	125	0.6595	0.9961 ± 0.0017	Caglia et al. (1998)
Serbia	185	0.6121	0.9797 ± 0.0041	Veselinovic et al. (2008)
South Portugal	78	0.6096	0.9942 ± 0.0035	Adams et al. (2008)
Sweden	344	0.6187	0.9837 ± 0.0028	Karlsson et al. (2006)
Valencia	213	0.5548	0.9825 ± 0.0033	Aler et al. (2001); $n = 140$ Adams et al. (2008); $n = 73$
Western Russia	543	0.5974	0.9866 ± 0.0017	Roewer et al. (2008)

defined as DYS19*14-DYS390*24-DYS391*11-DYS392*13-DYS393*13 (Wilson et al. 2001)—which is observed at a frequency of 50% in the present Basque sample and found at frequencies ranging from 44% in Ireland to 70% in Wales. R1b is considered a Western European-specific haplogroup, which

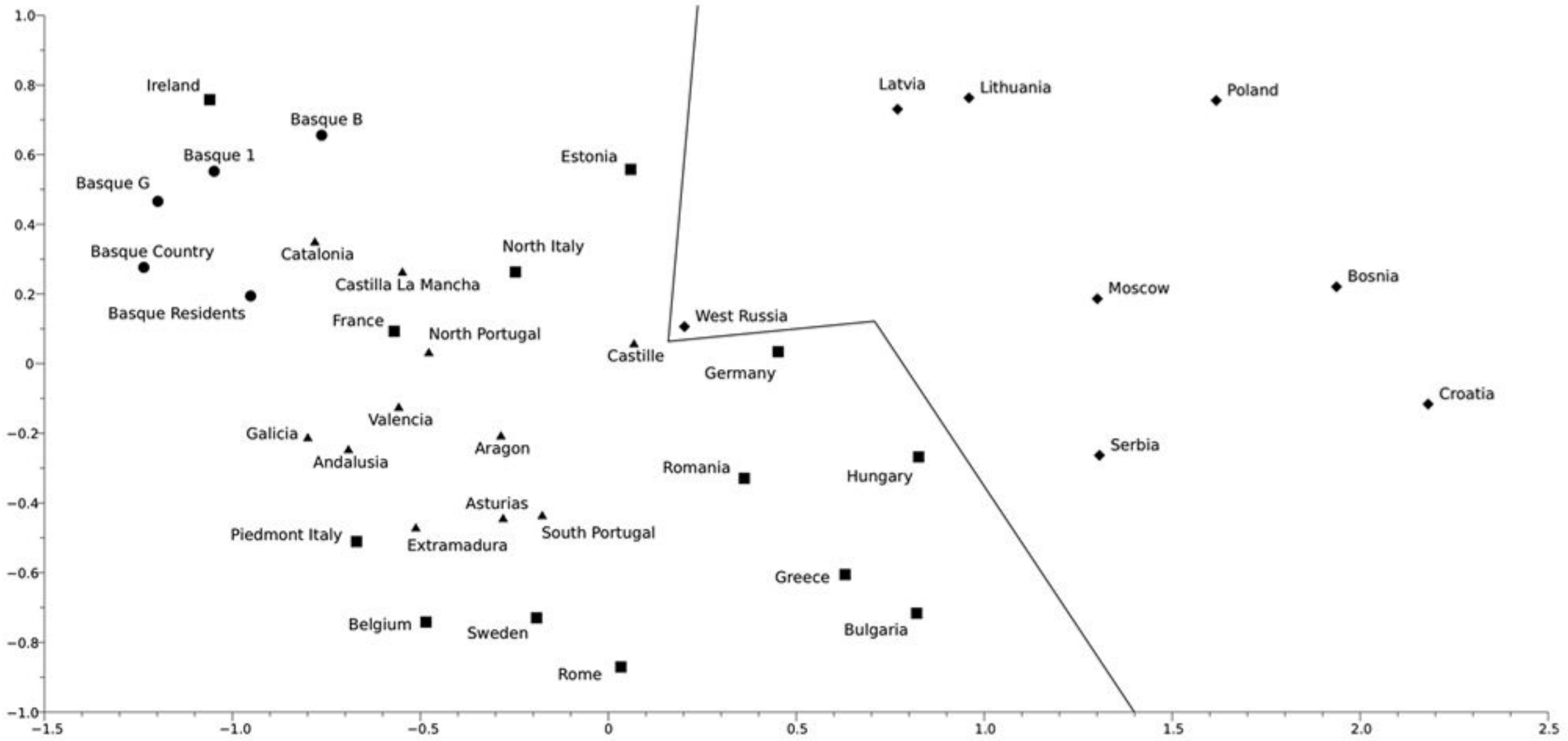

Figure 2. MDS plot of 37 European populations based on Slatkin's distance. Total variation on the first two axes = 94.90% (STRESS = 0.11528, $r = 0.95688$. Basque populations (circles), Iberian groups (triangles), other Western European populations (squares), Eastern European populations (diamonds). Genetic barriers detected by SAMOVA (black line), $F_{CT} = 0.06388$).

Table 6. Divergence Time Estimates for Various European Y-Chromosome Haplogroups

Haplogroup (Population)	*Divergence Time*	*Reference*
R1b (Basques)	17,900–21,300 BP	Alonso et al. (2005)
E1b1b	7,000–14,000 BP	Semino et al. (2004)
G2a	9,500–17,000 BP	Cinnioglu et al. (2004), Semino et al. (2000a)
I2a2 (Pyrenees)	11,900 BP	Lopez-Parra et al. (2009)
J2a	4,700–10,000 BP	Di Giacomo (2004)
L2 (Turkey)	14,600 BP	Cinnioglu et al. (2004)
T	20,700 BP	King et al. (2007)

diverged during the Upper Paleolithic (see Table 6), with the high frequencies found along the Atlantic Fringe being attributed to genetic drift during the Last Glacial Maximum (LGM) (Alonso et al. 2005; Quintana-Murci et al. 1999; Rootsi et al. 2004; Semino et al. 2000a; Wells et al. 2001). Recently, based on an analysis of the geographic spread of variation in R1b, Balaresque et al. (2010) suggested that the majority of paternal lineages present in modern European populations originate during the Neolithic. In particular, they suggest that R1b has its origin in the Anatolian Neolithic (TMRCA 7000–7989 YA; 95% CI: 4,423–11,014). TMRCA estimates for European populations in this study were generally younger than those in Anatolia, but TMRCA for R1b in Baie de Somme in France (7384 YA (95% CI: 5,259–10,131) and Bavaria in Germany (7282 YA; 95% CI: 5,059–10,139) are within the range of Anatolian estimates. While 116 Basques were typed for this study, TMRCA estimates are not provided, and they were excluded from many of the published analyses. Further analysis of R1b Y-STR haplotype variation in Europe and Anatolia using the additional DYS461 locus revealed two core R1b haplotypes (DYS393*13/DYS461*12 in Western Europe vs. DYS393*12/DYS461*11 in Eastern Europe). The majority of Anatolian R1b samples are of the Eastern European type, while those in Sardinia are predominantly Western European and belong to the Atlantic Modal Haplotype (Wilson et al. 2001). In addition, this analysis estimated TMRCA for R1b ranging from 32,600 YA in Iberia (95% CI: 25,000–80,700) to 19,600 YA in Anatolia (95% CI: 19,400–44,400), well before the advent of agriculture (Morelli et al. 2010). The discrepancy in TMRCA estimates between the two studies was attributed to the mutation rates used to estimate coalescent times. Morelli et al. (2010) used an evolutionary effective mutation rate (Zhivotovsky 2001; Zhivotovsky et al. 2004), whereas Balaresque et al. (2010) used an STR-specific germline mutation rate, which can bias estimates downward under conditions of constant population size.

Haplogroup I2a2 has previously been reported among Basques (6%), the populations of Castile (19%), Bernais (7.7%), and Normandy (2.4%), as well as the Irish (2.6%), with a high of 40.9% in Sardinia (Flores et al. 2004; Rootsi et

al. 2004). This European-specific haplogroup is believed to have originated in the Pyrenees before the LGM (Karafet et al. 2008; Lopez-Parra et al. 2009; Novelletto 2007; Rootsi et al. 2004).

Several Y haplogroups with TMRCA estimates dated after the advent of agriculture were also detected in the present study: E1b1b, G2a, and J2a (Table 6). Haplogroup E1b1b has been reported in many other European populations, including Portugal (4.0%), Spain (3.2%), France (4.7%), Italy (10%), and Sardinia (3.5%). Among the French Basques, the frequency of E1b1b was 6.3% (Cruciani et al. 2004). Haplogroup E1b1b has a complex history, with evidence of several demographic expansion events (Cruciani et al. 2004; Cruciani et al. 2006; Cruciani et al. 2007). The most common variant of this haplogroup in Europe is defined by mutation E-M78, which is believed to have originated in the Horn of Africa, from which it spread to the Middle East and then into Southern Europe during the Neolithic (Semino et al. 2004). Haplogroup G2a has been reported in Italy (10%), Sardinia (14.1%), Catalonia (8.3%), and Andalusia (2%) and reaches its highest frequency in Palestine (75%) (Francalacci et al. 2003; Semino et al. 2000a; Shen et al. 2004). Like E1b1b, G2a is considered to have diverged during the Neolithic (Cinnioglu et al. 2004), spreading from the Middle East into Europe with the advent of agriculture. Haplogroup J2a has been previously reported among the French Basques (13.6%), French (13%), Italians (12–16%), and Greeks (21%), with higher frequencies in Turkey (40%) and Lebanon (29%) (Semino et al. 2004; Semino et al. 2000a). It is also considered a marker of Neolithic expansion (Capelli et al. 2007; Rosser et al. 2000), but a more recent maritime route has been proposed for the distribution of J2a from the Middle East through the Mediterranean (Di Giacomo et al. 2004).

Haplogroup L has been divided into three subclades, L1, L2, and L3. L1 is found at moderate frequency in India (6.3%), whereas L3 is more frequent in Pakistan (6.8%). The majority of European and Anatolian L samples are believed to be L2 (Sengupta et al. 2006). Haplogroup L2 has been detected in several other European groups, including Andalusia (3.4%), Italy (5.4%), Greece (1.3%), and Hungary (2.2%). It is also found at low frequencies in Turkey (1.6–4.2%) and Syria (3.2%) (Cinnioglu et al. 2004; Semino et al. 2000b). The parental haplogroup L originated in East Africa around 30,000 years ago, whereas L2 dates to 14,600 years ago in Turkish populations assuming a model of continuous growth (Cinnioglu et al. 2004).

Haplogroup T has been found in several other Iberian populations, including Huelva (4.5%), Seville (4.5%), Cadiz (10.7%), Valencia (3.2%), North Portugal (0.9%), and Cantabria (4.3%) (Flores et al. 2004). It is present at low frequencies in other Western European groups, including Britain (0.5%), Western France (0.5%), Sardinia (1.3%), and Greece (1.3%). Haplogroup T is also found in African populations in Egypt (7.2%), Ethiopia (4.8%), and Somalia (10.5%) and in Middle Eastern groups in Iraq (7.2%), Turkey (2.5%), and Oman (8.3%) (King et al. 2007). This haplogroup has been dated to 20,700 yrs ago, and

network analysis suggests that haplogroup T is "an ancient and diverse indigenous European lineage, rather than recent immigrants from the Middle East or Africa" (King et al. 2007:585). Overall, the evidence from the paternal Basque lineages demonstrate a 11% Post-Neolithic contribution based on previously published TMRCA estimates for haplogroups E1b1b, G2a, and J2a (See Table 6), with Alava showing the greatest Post-Neolithic genetic contribution (6 of 32 lineages), whereas Guipuzcoa has four of 58 and Vizcaya has three of 25.

A study of admixture rates in contemporary European populations, using 22 Y-SNPs and putative Paleolithic (Basque) and Neolithic (Turkey, Iraq, Syria, and Lebanon) parental populations, found that the percentage of Neolithic admixture varies with distance from the Levant (Chikhi et al. 2002). Greece, Albania, and Macedonia were described as 100% Neolithic, while Sardinia, the Netherlands, and Andalusia have 0–16% Neolithic admixture. The authors interpret these results as evidence for the demic diffusion model, but archaeological evidence suggests that modern populations in Greece would have few non-Neolithic markers because of a lack of Paleolithic occupation of that region, though Pinhasi (2000) has suggested this may be the result of sampling bias of the archaeological record in Greece. Frequency and variance of Y-chromosome haplogroups in Central Europe suggest the diffusion of agriculture into this area was accomplished by populations bearing the indigenous I-M423 haplogroup, rather than haplogroups linked to the Neolithic transition (G2a and J2a) (Battaglia et al. 2009). The average non-Neolithic contribution in 17 populations was found to be 51% for the Y-chromosome, which argues strongly against complete replacement. A meta-analysis of eight data sets, including mtDNA, Y SNPs, and autosomal markers, estimated Neolithic admixture rates in 34 European populations (Dupanloup et al. 2004). As in the previous study, Basques were chosen to represent the Paleolithic parental population, whereas Near East and Anatolian groups were used as a proxy for the Neolithic parental population. The Neolithic contribution in Europe varied with distance from the Levant, ranging from 80% in the Balkans to only 21% in the British Isles.

During the Paleolithic and into the Mesolithic, the Iberian Peninsula was sparsely populated, particularly in the interior. Most of the sites were coastal, reflecting the comparative wealth of resources (Straus 1991b; Zilhao 2000). The northern coast of Iberia, known as the Franco-Cantabrian region, was one of two refuge areas during the Last Glacial Maximum, the second being in the Ukraine. Paleolithic populations would have retreated to these areas during the height of the glaciation, and the archaeological record demonstrates an increase in occupation sites between 21,000–16,000 BP in this region of Iberia (47 sites compared with only 26 for the entire Early Upper Paleolithic) (Housley et al. 1997; Straus 1991a). Divergence times for Y-chromosome haplogroups R1b and I2a2 have also been dated to this period, suggesting that these haplogroups have a pre-Neolithic origin. Rather than demonstrating evidence of a "wave of

advance," the Neolithic archaeological record in Iberia instead shows a leapfrog settlement pattern, again along the coasts (Zilhao 2000). Mesolithic populations were well established in these areas, and it is likely that small groups of Neolithic "maritime pioneers" would have assimilated, or been assimilated by, the local Mesolithic groups (Lahr et al. 2000; Zilhao 2000; Zilhao 2003). Based on the archaeological evidence, the entire Iberian Peninsula appears little affected by the Neolithic Transition, lending credence to the idea that the Basques are a reservoir of pre-Neolithic Y lineages, although one which has experienced some admixture since the end of the Last Glacial Maximum.

Comment. Data set available from the first author upon request.

Acknowledgments We thank the Basque participants, Dr. Arantza Apraiz for sample collection, and Dr. Rohina Rubicz for her assistance with DNA extraction. This work is supported in part by National Geographic Society Grant (Project 6935–00) to the University of Kansas Laboratory of Biological Anthropology.

Received 14 October 2010; revision accepted for publication 4 April 2011.

Literature Cited

Adams, S. M., E. Bosch, P. L. Balaresque et al. 2008. The genetic legacy of religious diversity and intolerance: Paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am. J. Hum. Genet.* 83(6):725–736.

Aguirre, A. I., A. Vicario, L. I. Mazon et al. 1989. AK1, PGD, GC, and HP frequencies in the Basque population: A review. *Gene Geogr.* 3(1):41–51.

Aguirre, A. I., A. Vicario, L. I. Mazon et al. 1991. Acid phosphatase, adenosine deaminase and esterase D polymorphisms in the Spanish Basque population. *Hum. Hered.* 41(2):93–102.

Aler, M., A. Salas, P. Sanchez-Diz et al. 2001. Y-chromosome STR haplotypes from a Western Mediterranean population sample. *Forensic Sci. Int.* 119(2):254–257.

Alfonso-Sanchez, M. A., S. Cardoso, C. Martinez-Bouzas et al. 2008. Mitochondrial DNA haplogroup diversity in Basques: A reassessment based on HVI and HVII polymorphisms. *Am. J. Hum. Biol.* 20(2):154–164.

Alonso, S., A. Castro, I. Fernandez et al. 1995. Population study of 3 STR loci in the Basque Country (northern Spain). *Int. J. Legal Med.* 107(5):239–245.

Alonso, S., C. Flores, V. Cabrera et al. 2005. The place of the Basques in the European Y-chromosome diversity landscape. *Eur. J. Hum. Genet.* 13(12):1293–1302.

Alzualde, A., N. Izagirre, S. Alonso et al. 2006. Insights into the "isolation" of the Basques: mtDNA lineages from the historical site of Aldaieta (6th–7th centuries AD). *Am. J. Phys. Anthropol.* 130(3):394–404.

Alzualde, A., N. Izagirre, S. Alonso et al. 2005. Temporal mitochondrial DNA variation in the Basque Country: Influence of post-Neolithic events. *Ann. Hum. Genet.* 69(6):665–679.

Ammerman, A. J., and L. L. Cavalli-Sforza. 1984. *The Neolithic Transition and the Genetics of Populations in Europe*. Princeton, NJ: Princeton University Press.

Arrieta, M. I., B. Martinez, J. M. Millan et al. 1997. Study of a trimeric tandem repeat locus (SBMA) in the Basque population: Comparison with other populations. *Gene Geogr.* 11(1):61–72.

Athey, T. W. 2005. Haplogroup prediction from Y-STR values using an allele-frequency approach. *Journal of Genetic Genealogy.* 1:1–7.

Athey, T. W. 2006. Haplogroup prediction from Y-STR values using a Bayesian-allele-frequency approach. *Journal of Genetic Genealogy.* 2:34–39.

Balaresque, P., G. R. Bowden, S. M. Adams et al. 2010. A predominantly Neolithic origin for European paternal lineages. *PLoS Biol.* 8(1):e1000285.

Ballard, D. J., C. Phillips, C. R. Thacker et al. 2006. Y chromosome STR haplotype data for an Irish population. *Forensic Sci. Int.* 161(1):64–68.

Barbujani, G., G. Bertorelle, and L. Chikhi. 1998. Evidence for Paleolithic and Neolithic gene flow in Europe. *Am. J. Hum. Genet.* 62(2):488–492.

Battaglia, V., S. Fornarino, N. Al-Zahery et al. 2009. Y-chromosomal evidence of the cultural diffusion of agriculture in Southeast Europe. *Eur. J. Hum. Genet.* 17(6):820–830.

Bertranpetit, J., J. Sala, F. Calafell et al. 1995. Human mitochondrial DNA variation and the origin of Basques. *Ann. Hum. Genet.* 59(1):63–81.

Bosch, E., F. Calafell, F. R. Santos et al. 1999. Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. *Am. J. Hum. Genet.* 65(6):1623–1638.

Boyd, W. C., and L. G. Boyd. 1937. New data on blood groups and other inherited factors in Europe and Egypt. *Am. J. Phys. Anthropol.* 23:49–70.

Brion, M., A. Salas, A. Gonzalez-Neira et al. 2003. Insights into Iberian population origins through the construction of highly informative Y-chromosome haplotypes using biallelic markers, STRs, and the MSY1 minisatellite. *Am. J. Phys. Anthropol.* 122(2):147–161.

Broca, P. 1864. Letter to the Society regarding Basque crania. *Journal of the Anthropological Society of London.* 2:cclxvii–cclxxiii.

Caglia, A., M. Dobosz, I. Boschi et al. 1998. Increased forensic efficiency of a STR-based Y-specific haplotype by addition of the highly polymorphic DYS385 locus. *Int. J. Legal Med.* 111(3):142–146.

Calafell, F., and J. Bertranpetit. 1994. Principal component analysis of gene frequencies and the origin of Basques. *Am. J. Phys. Anthropol.* 93(2):201–215.

Calderon, R., M. Carrion, A. Perez-Miranda et al. 2003. Allele variation of DYS19 and Y-Alu insertion (YAP) polymorphisms in Basques: An insight into the peopling of Europe and the Mediterranean region. *Hum. Biol.* 75(1):117–127.

Calderon, R., C. Vidales, J. A. Pena et al. 1998. Immunoglobulin allotypes (GM and KM) in Basques from Spain: Approach to the origin of the Basque population. *Hum. Biol.* 70(4):667–698.

Capelli, C., F. Brisighelli, F. Scarnicci et al. 2007. Y chromosome genetic variation in the Italian peninsula is clinal and supports an admixture model for the Mesolithic-Neolithic encounter. *Mol. Phylogenet. Evol.* 44(1):228–239.

Casalotti, R., L. Simoni, M. Belledi et al. 1999. Y-chromosome polymorphisms and the origins of the European gene pool. *Proceedings: Biological Sciences* 266(1432):1959–1965.

Cavalli-Sforza, L. L., P. Menozzi, and A. Piazza 1994. *The History and Geography of Human Genes.* Princeton, NJ: Princeton University Press.

Cerri, N., A. Verzeletti, B. Bandera et al. 2005. Population data for 12 Y-chromosome STRs in a sample from Brescia (northern Italy). *Forensic Sci. Int.* 152(1):83–87.

Cerutti, N., A. Marin, C. Di Gaetano et al. 2006. Population data for Y-chromosome STR haplotypes from Piedmont (Italy). *Forensic Sci. Int.* 158(2–3):238–243.

Chalmers, J. N., E. W. Ikin, and A. E. Mourant. 1948. Basque blood groups. *Nature.* (162):27.

Chalmers, J. N., E. W. Ikin, and A. E. Mourant. 1949. The ABO, MN and Rh blood groups of the Basque people. *Am. J. Phys. Anthropol.* 7(4):529–544.

Chikhi, L., G. Destro-Bisol, G. Bertorelle et al. 1998. Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *Proc. Natl. Acad. Sci. USA* 95(15):9053–9058.

Chikhi, L., R. A. Nichols, G. Barbujani et al. 2002. Y genetic data support the Neolithic demic diffusion model. *Proc. Natl. Acad. Sci. USA* 99(17):11008–11013.

Cinnioglu, C., R. King, T. Kivisild et al. 2004. Excavating Y-chromosome haplotype strata in Anatolia. *Hum. Genet.* 114(2):127–148.

Corte-Real, H. B., V. A. Macaulay, M. B. Richards et al. 1996. Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. *Ann. Hum. Genet.* 60(4): 331–350.

Cruciani, F., R. La Fratta, P. Santolamazza et al. 2004. Phylogeographic analysis of haplogroup E3b (E-M215) Y chromosomes reveals multiple migratory events within and out of Africa. *Am. J. Hum. Genet.* 74(5):1014–1022.

Cruciani, F., R. La Fratta, A. Torroni et al. 2006. Molecular dissection of the Y chromosome haplogroup E-M78 (E3b1a): A posteriori evaluation of a microsatellite-network-based approach through six new biallelic markers. *Hum. Mutat.* 27(8):831–832.

Cruciani, F., R. La Fratta, B. Trombetta et al. 2007. Tracing past human male movements in northern/eastern Africa and western Eurasia: New clues from Y-chromosomal haplogroups E-M78 and J-M12. *Mol. Biol. Evol.* 24(6):1300–1311.

Di Giacomo, F., F. Luca, L. O. Popa et al. 2004. Y chromosomal haplogroup J as a signature of the post-Neolithic colonization of Europe. *Hum. Genet.* 115(5):357–371.

Dugoujon, J. M., J. Clayton, A. Sevin et al. 1989. Immunoglobulin (Gm and Km) allotypes in some Pyrenean populations of France. *Collegium Antropologicum* 13:43–50.

Dupanloup, I., G. Bertorelle, L. Chikhi et al. 2004. Estimating the impact of prehistoric admixture on the genome of Europeans. *Mol. Biol. Evol.* 21(7):1361–1372.

Dupanloup, I., S. Schneider, and L. Excoffier. 2002. A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11(12):2571–2581.

Egyed, B., S. Furedi, and Z. Padar. 2006. Population genetic study in two Transylvanian populations using forensically informative autosomal and Y-chromosomal STR markers. *Forensic Sci. Int.* 164(2–3):257–265.

Esteban, E., J. M. Dugoujon, E. Guitard et al. 1998. Genetic diversity in northern Spain (Basque Country and Cantabria): GM and KM variation related to demographic histories. *Eur. J. Hum. Genet.* 6(4):315–324.

Etcheverry, M. A. 1945. El factor rhesus, su genetica e importancia clinica. *El Dia Medico* 17:1237–1259.

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. *Molecular Ecology Resources* 10(3):564–567.

Flores, C., N. Maca-Meyer, A. M. Gonzalez et al. 2004. Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: Implications for population demography. *Eur. J. Hum. Genet.* 12(10):855–863.

Francalacci, P., L. Morelli, P. A. Underhill et al. 2003. Peopling of three Mediterranean islands (Corsica, Sardinia, and Sicily) inferred by Y-chromosome biallelic variability. *Am. J. Phys. Anthropol.* 121(3):270–279.

Furedi, S., J. Woller, Z. Padar et al. 1999. Y-STR haplotyping in two Hungarian populations. *Int. J. Legal Med.* 113(1):38–42.

Garcia, O., R. Fregel, J. M. Larruga et al. 2011. Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge. *Heredity* 106(1):37–45.

Garcia, O., P. Martin, B. Budowle et al. 1998. Basque Country autochthonous population data on 7 short tandem repeat loci. *Int. J. Legal Med.* 111(3):162–164.

Garcia, O., P. Martin, L. Gusmao et al. 2004. A Basque Country autochthonous population study of 11 Y-chromosome STR loci. *Forensic Sci. Int.* 145(1):65–68.

Garcia, O., I. Uriarte, P. Martin et al. 2001. STR data from Basque country autochthonous population. *Forensic Sci. Int.* 115(1–2):111–112.

Gonzalez, A. M., A. Brehm, J. A. Perez et al. 2003. Mitochondrial DNA affinities at the Atlantic fringe of Europe. *Am. J. Phys. Anthropol.* 120(4):391–404.

Gonzalez, A. M., O. Garcia, J. M. Larruga et al. 2006. The mitochondrial lineage U8a reveals a Paleolithic settlement in the Basque country. *BMC Genomics* 7(1):124.

Gonzalez-Neira, A., L. Gusmao, M. Brion et al. 2000. Distribution of Y-chromosome STR defined haplotypes in Iberia. *Forensic Sci. Int.* 110(2):117–126.

Gusmao, L., P. Sanchez-Diz, C. Alves et al. 2003. Grouping of Y-STR haplotypes discloses European geographic clines. *Forensic Sci. Int.* 134(2–3):172–179.

Hazout, S., J. M. Dugoujon, F. Loirat et al. 1991. Genetic similarity maps and immunoglobulin allotypes of eleven populations from the Pyrenees (France). *Ann. Hum. Genet.* 55(Pt 2):161–174.

Housley, R. A., C. S. Gamble, M. Street et al. 1997. Radiocarbon evidence for the lateglacial human recolonisation of Northern Europe. *Proceedings of the Prehistoric Society* 63:25–54.

Iriondo, M., M. C. Barbero, N. Izagirre et al. 1997. Data on six short-tandem repeat polymorphisms in an autochthonous Basque population. *Hum. Hered.* 47(3):131–137.

Iriondo, M., C. de la Rua, M. C. Barbero et al. 1999. Analysis of 6 short tandem repeat loci in Navarre (northern Spain). *Hum. Biol.* 71(1):43–54.

Izagirre, N., S. Alonso, and C. de la Rua. 2001. DNA analysis and the evolutionary history of the Basque population: A review. *Journal of Anthropological Research* 57:325–344.

Izagirre, N., and C. de la Rua. 1999. An mtDNA analysis in ancient Basque populations: Implications for haplogroup V as a marker for a major Paleolithic expansion from southwestern Europe. *Am. J. Hum. Genet.* 65(1):199–207.

Jackes, M., D. Lubell, and C. Meiklejohn. 1997. On physical anthropological aspects of the Mesolithic-Neolithic transition in the Iberian Peninsula. *Current Anthropology* 38(5): 839–846.

Karafet, T. M., F. L. Mendez, M. B. Meilerman et al. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* 18(5):830–838.

Karlsson, A. O., T. Wallerstrom, A. Gotherstrom et al. 2006. Y-chromosome diversity in Sweden: A long-time perspective. *Eur. J. Hum. Genet.* 14(8):963–970.

Keyser-Tracqui, C., F. X. Ricaut, P. Blandin et al. 2003. French allele frequencies and haplotypes of nine Y-specific STRs. *J. Forensic Sci.* 48(1):242–244.

King, T. E., G. R. Bowden, P. L. Balaresque et al. 2007. Thomas Jefferson's Y chromosome belongs to a rare European lineage. *Am. J. Phys. Anthropol.* 132(4):584–589.

Klaric, I. M., L. B. Lauc, M. Pericic et al. 2005. Evaluation of Y-STR variation in Bosnian and Herzegovinian population. *Forensic Sci. Int.* 154(2–3):252–256.

Laayouni, H., and J. Bertranpetit. 2009. From the detection of population structure to the reconstruction of population history: The historical reading of the human genome. *Heredity* 103(5):362–363.

Lahr, M. M., R. A. Foley, and R. Pinhasi 2000. Expected regional patterns of Mesolithic-Neolithic human population admixture in Europe based on archaeological evidence. In *Archaeogenetics: DNA and the Population Prehistory of Europe*, C. Renfrew and K. Boyle, eds. Cambridge, U.K.: McDonald Institute for Archaeological Research, 81–88.

Lessig, R., and J. Edelmann. 2001. Population data of Y-chromosomal STRs in Lithuanian, Latvian and Estonian males. *Forensic Sci. Int.* 120(3):223–225.

Levine, M. H. 1977. A hematological approach to Basque isolation in two French Basque Villages. *Ann. N. Y. Acad. Sci.* 293:185–193.

Levine, M. H., V. von Hagen, J. C. Quilici et al. 1974. Anthropology of a Basque village: A new hematypological study. *Cahiers d'anthropologie et d'ecologie humaine* II:159–171.

Ljubkovic, J., A. Stipisic, D. Sutlovic et al. 2008. Y-chromosomal short tandem repeat haplotypes in southern Croatian male population defined by 17 loci. *Croat. Med. J.* 49(2):201–206.

Lopez-Parra, A. M., L. Gusmao, L. Tavares et al. 2009. In search of the pre- and post-Neolithic genetic substrates in Iberia: Evidence from Y-chromosome in Pyrenean populations. *Ann. Hum. Genet.* 73(1):42–53.

Lucotte, G., and S. Hazout 1996. Y-chromosome DNA haplotypes in Basques. *J. Mol. Evol.* 42(4):472–475.

Lucotte, G., and F. Loirat 1999. Y-chromosome DNA haplotype 15 in Europe. *Hum. Biol.* 71(3):431–437.

MacClancy, J. 1993. Biological Basques, sociologically speaking. In *Social and Biological Aspects of Ethnicity*, M. Chapman, ed. Oxford, U.K.: Oxford University Press.

Manzano, C., A. I. Aguirre, M. Iriondo et al. 1996. Genetic polymorphisms of the Basques from Gipuzkoa: Genetic heterogeneity of the Basque population. *Ann. Hum. Biol.* 23(4):285–296.

Manzano, C., P. Moral, C. de la Rua. 1993. Serum protein polymorphisms (GC, TF, and PI subtypes) in the Basque population of Alava. *Hum. Hered.* 43(2):121–125.

Martinez, L., P. A. Underhill, L. A. Zhivotovsky et al. 2007. Paleolithic Y-haplogroup heritage predominates in a Cretan highland plateau. *Eur. J. Hum. Genet.* 15(4):485–493.

Mertens, G., E. Jehaes, G. Leijnen et al. 2007. Twelve-locus Y-STR haplotypes in the Flemish population. *J. Forensic. Sci.* 52(3):755–757.

Morelli, L., D. Contu, F. Santoni et al. 2010. A comparison of Y-chromosome variation in Sardinia and Anatolia is more consistent with cultural rather than demic diffusion of agriculture. *PLoS ONE* 5(4):e10419.

Mourant, A. E. 1948. The blood groups of the Basques. *Nature* 160:505–506.

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

Nothnagel, M., T. T. Lu, M. Kayser et al. 2010. Genomic and geographic distribution of SNP-defined runs of homozygosity in Europeans. *Hum. Mol. Genet.* 19(15):2927–2935.

Novelletto, A. 2007. Y chromosome variation in Europe: Continental and local processes in the formation of the extant gene pool. *Ann. Hum. Biol.* 34(2):139–172.

Parreira, K. S., M. V. Lareu, P. Sanchez-Diz. 2002. DNA typing of short tandem repeat loci on Y-chromosome of Greek population. *Forensic Sci. Int.* 126(3):261–264.

Pawlowski, R., and A. Maciejewska. 2000. Forensic validation of a multiplex containing nine STRs–population genetics in northern Poland. *Int. J. Legal Med.* 114(1–2):45–49.

Pena, J. A., S. Garcia-Obregon, A. M. Perez-Miranda et al. 2006. Gene flow in the Iberian Peninsula determined from Y-chromosome STR loci. *Am. J. Hum. Biol.* 18(4):532–539.

Perez-Lezaun, A., F. Calafell, J. Clarimon et al. 2000. Allele frequencies of 13 short tandem repeats in population samples from the Iberian Peninsula and northern Africa. *Int. J. Legal Med.* 113(4):208–214.

Perez-Miranda, A. M., M. A. Alfonso-Sanchez, A. Kalantar et al. 2005a. Microsatellite data support subpopulation structuring among Basques. *J. Hum. Genet.* 50(8):403–414.

Perez-Miranda, A. M., M. A. Alfonso-Sanchez, J. A. Pena et al. 2005b. Genetic polymorphisms at 13 STR loci in autochthonous Basques from the province of Alava (Spain). *Legal Medicine (Tokyo, Japan)* 7(1):58–61.

Pinhasi, R., R. A. Foley, and M. M. Lahr. 2000. Spatial and temporal patterns in the Mesolithic-Neolithic archaeological record of Europe. In *Archaeogenetics: DNA and the Population Prehistory of Europe*, C. Renfrew and K. Boyle, eds. Cambridge, UK: McDonald Institute for Archaeological Research, 45–56.

Ploski, R., M. Wozniak, R. Pawlowski et al. 2002. Homogeneity and distinctiveness of Polish paternal lineages revealed by Y chromosome microsatellite haplotype analysis. *Hum. Genet.* 110(6):592–600.

Quintana-Murci, L., O. Semino, E. Minch et al. 1999. Further characteristics of proto-European Y chromosomes. *Eur. J. Hum. Genet.* 7(5):603–608.

Richards, M., H. Corte-Real, P. Forster et al. 1996. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. *Am. J. Hum. Genet.* 59(1):185–203.

Rodriguez-Ezpeleta, N., J. Alvarez-Busto, L. Imaz et al. 2010. High-density SNP genotyping detects homogeneity of Spanish and French Basques, and confirms their genomic distinctiveness from other European populations. *Hum. Genet.* 128(1):113–117.

Roewer, L., S. Willuweit, C. Kruger et al. 2008. Analysis of Y chromosome STR haplotypes in the European part of Russia reveals high diversities but non-significant genetic distances between populations. *Int. J. Legal Med.* 122(3):219–223.

Rohlf, F. J. 2005. *NTSYS-PC: Numerical Taxonomy and Multivariate Analysis System 2.1 ed.* New York, NY: Exeter Software.

Rootsi, S., C. Magri, T. Kivisild et al. 2004. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am. J. Hum. Genet.* 75(1):128–137.

Rosser, Z. H., T. Zerjal, M. E. Hurles et al. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am. J. Hum. Genet.* 67(6):1526–1543.

Semino, O., C. Magri, G. Benuzzi et al. 2004. Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: Inferences on the Neolithization of Europe and later migratory events in the Mediterranean area. *Am. J. Hum. Genet.* 74(5):1023–1034.

Semino, O., G. Passarino, P. J. Oefner et al. 2000a. The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: A Y chromosome perspective. *Science* 290(5494):1155–1159.

Semino, O., G. Passarino, L. Quintana-Murci et al. 2000b. mtDNA and Y chromosome polymorphisms in Hungary: Inferences from the Palaeolithic, Neolithic and Uralic influences on the modern Hungarian gene pool. *Eur. J. Hum. Genet.* 8(5):339–346.

Sengupta, S., L. A. Zhivotovsky, R. King et al. 2006. Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *Am. J. Hum. Genet.* 78(2):202–221.

Shen, P., T. Lavi, T. Kivisild et al. 2004. Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation. *Hum. Mutat.* 24(3):248–260.

Shi, W., Q. Ayub, M. Vermeulen et al. 2010. A worldwide survey of human male demographic history based on Y-SNP and Y-STR data from the HGDP-CEPH populations. *Mol. Biol. Evol.* 27(2):385–393.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139(1):457–462.

Straus, L. G. 1991a. Human geography of the Late Upper Paleolithic in Western Europe: Present state of the question. *Journal of Anthropological Research* 47(2):259–278.

Straus, L. G. 1991b. Southwestern Europe at the Last Glacial Maximum. *Current Anthropology* 32(2):189–199.

Torroni, A., H. J. Bandelt, L. D'Urbano et al. 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am. J. Hum. Genet.* 62(5):1137–1152.

Trask, R. L. 1997. *The History of Basque*. London,U.K.: Routledge.

Veselinovic, I. S., D. M. Zgonjanin, M. P. Maletin et al. 2008. Allele frequencies and population data for 17 Y-chromosome STR loci in a Serbian population sample from Vojvodina province. *Forensic Sci. Int.* 176(2–3):e23–e28.

Wells, R. S., N. Yuldasheva, R. Ruzibakiev et al. 2001. The Eurasian heartland: A continental perspective on Y-chromosome diversity. *Proc. Natl. Acad. Sci. USA* 98(18):10244–10249.

Wilson, J. F., D. A. Weiss, M. Richards et al. 2001. Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc. Natl. Acad. Sci. USA* 98(9):5078–5083.

Zaharova, B., S. Andonova, A. Gilissen et al. 2001. Y-chromosomal STR haplotypes in three major population groups in Bulgaria. *Forensic Sci. Int.* 124(2–3):182–186.

Zhivotovsky, L. A. 2001. Estimating divergence time with the use of microsatellite genetic distances: Impacts of population growth and gene flow. *Mol. Biol. Evol.* 18(5):700–709.

Zhivotovsky, L. A., P. A. Underhill, C. Cinnioglu et al. 2004. The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. *Am. J. Hum. Genet.* 74(1):50–61.

Zilhao, J. 2000. From the Mesolithic to the Neolithic in the Iberian Peninsula. In *Europe's First Farmers*, T. D. Price, ed. Cambridge, U.K.: Cambridge University Press, 144–182.

Zilhao, J. 2003. The Neolithic transition in Portugal and the role of demic doffusion in the spread of agriculture across West Mediterranean Europe. In *The Widening Harvest: The Neolithic Transition in Europe*, A. J. Ammerman and P. Biagi, eds. Boston, MA: Archaeological Institute of America, 207–223.

Zlojutro, M., R. Roy, J. Palikij et al. 2006. Autosomal STR variation in a Basque population: Vizcaya Province. *Hum. Biol.* 78(5):599–618.