

## Autosomal short tandem repeat genetic variation of the Basques in Spain

Kristin L. Young<sup>1</sup>,  
Guangyun Sun<sup>2</sup>, Ranjan  
Deka<sup>2</sup>, Michael H.  
Crawford<sup>3</sup>

<sup>1</sup>Department of Family Medicine,  
Research Division, University of  
Kansas Medical Center, Kansas City,  
Kan, USA

<sup>2</sup>Department of Environmental  
Health, University of Cincinnati  
Medical Center, Cincinnati, Ohio,  
USA

<sup>3</sup>Department of Anthropology,  
University of Kansas, Lawrence,  
Kan, USA

**Aim** To examine population genetic structure and hypotheses of the origin of the modern Basque population in Spain using autosomal short tandem repeat (STR) data from individuals living in 27 mountain villages in the provinces of Alava, Vizcaya, Guipuzcoa, and Navarre, by comparing Basque autosomal STR variation with that of neighboring populations in Europe, as well as proposed ancestral populations in North Africa and the Caucasus.

**Methods** Allele frequencies for 9 autosomal STR loci (D3S1358, D5S818, D7S820, D8S1179, D13S317, D18S51, D21S11, FGA, and vWA) and several population genetic parameters were determined for the 4 provinces in the Basque region of Spain (n=377). Heterozygosity within the Basque population was measured using a locus-by-locus analysis of molecular variance. Relationships between the Basques and other populations were examined using a multidimensional scaling (MDS) plot of Shriver's  $D_{sw}$  distance matrix.

**Results** Heterozygosity levels in the Basque provinces were on the low end of the European distribution (0.805-0.812). The MDS plot of genetic distances revealed that the Basques differed from both the Caucasian and North African populations with respect to autosomal STR variation.

**Conclusions** Autosomal STR analysis does not support the hypotheses of a recent common ancestor between the Basques and populations either from the Caucasus or North Africa.

Received: April 1, 2011

Accepted: May 17, 2011

**Correspondence to:**

Kristin L. Young  
Department of Family Medicine  
Research Division  
University of Kansas Medical Center  
4125 Rainbow Boulevard  
Mailstop 3064  
Kansas City, Kansas 66160  
[kyoung@kumc.edu](mailto:kyoung@kumc.edu)

The question of Basque origins has interested scholars since the 1800s, when Aranzadi suggested, based on cranial morphology, that the Basques were an ancient relict population (1). The Basque language, *Euskara*, is most widely accepted as an isolate, unrelated to any other extant language in Europe (2). Many hypotheses of a relationship between Basques and other populations have been put forward based on linguistic analyses. These proposed linguistic connections include the ancient languages of Iberian, Minoan, Etruscan, Pictish, Sumerian, and Aquitanian, as well as extant Uralic (such as Finnish), Caucasian (such as Georgian), African (especially Berber), and Native American languages, and even Japanese (3-10). It has been suggested that populations in the Basque region and the Caucasus, which speak non-Indo-European agglutinative languages, could be remnants of a Mesolithic European population and have been less affected than the rest of the continent by the Neolithic Revolution for the same reason – both inhabit mountainous regions that were less hospitable to agricultural pursuits (6). Alternatively, the Vasco-Iberian hypothesis holds that languages related to Basque were spoken throughout the Iberian Peninsula prior to Roman conquest. A genetic relationship between *Euskara* and Iberian was favored in the late 1700s, with Basque considered the last remnant of this larger language family, but discoveries of Iberian inscriptions which were not translatable using *Euskara* weakened this hypothesis on linguistic grounds. Because Iberians were believed to have migrated from North Africa, and a connection between Iberian and Basque had been proposed, genetic similarities between Basques and North Africans have also been sought (11). There are also linguists who conclude that Basque is an autochthonous language, which developed in situ in the Iberian Peninsula, and once had a wider range, but has also had contact with other languages in historical times (2).

Studies of human blood types in the mid-20th century bore out the distinctiveness of the Basques, distinguishing them from other European populations by a low frequency of ABO\*B (1.1%) and a high frequency of RH\*cde (between 30.5%-35.6%) (12-15). Since then, the Basque population has been characterized from a genetic perspective using blood group antigens (12-20), erythrocytic enzymes (21-23), plasma proteins (24), HLA antigens and haplotypes (25-30), Y-chromosome markers (31-36), mitochondrial haplogroups and sequences (37-44), whole genome single-nucleotide polymorphism (SNP) analyses (45-47), and autosomal microsatellites (48-58). Microsatellites are sequences of 2-6 bases tandemly repeated 10-30 times, which are found scattered throughout the genome. These

short tandem repeats (STR) are considered selectively neutral, and therefore appropriate for population genetic studies. Thirteen of these STR loci comprise the Combined DNA Index System (CODIS) used for forensic purposes, and have been widely implemented in anthropological genetics because forensic databases provide a wealth of comparative data.

The present study characterizes autosomal STR genotypes from the Basque region of Spain to examine population substructure and genetic relationships with other groups, including testing of the proposed genetic affinities between the Basques and populations in the Caucasus and North Africa. We predict that if the Basques share a common ancestor (or have experienced more recent migration and resultant gene flow) with either populations in the Caucasus or North Africa, allele frequencies of autosomal STR loci will be similar among the Basques and these proposed related populations and genetic distances between these populations will be low. Alternatively, if the Basques are an autochthonous European population (with no recent gene flow from groups in North Africa or the Caucasus), autosomal STR frequencies will be within the range of other populations on the European continent, and the Basques will be more genetically similar to other European groups. Previous studies presenting STR data from the Basque population either used small samples that were often collected in urban areas of a single province (when collection location was reported) to study relationships between the Basques and other populations (49,54,58-60). When larger samples were collected (48,50-52,57), the studies most often presented allele frequencies and population genetic parameters for only a few loci (Table 1). This study represents one of the most comprehensive samples of Basques yet analyzed for autosomal STR variation, with 377 individuals in 27 mountain villages from throughout the Basque region of Spain.

## MATERIALS AND METHODS

### Population

To test the hypotheses of genetic similarity between Basques and other populations, buccal DNA samples were collected from 652 autochthonous (those who claimed 4 Basque grandparents) participants of both sexes, in mountain villages throughout the Basque region of northern Spain. Six villages were sampled in Alava (n=143), 17 villages in Vizcaya (n=237), 10 villages in Guipuzcoa (n=220), and 2 villages in Navarre (n=52). A

sub-sample of individuals from 27 villages (N=478) was screened for autosomal STR genotypes (Table 1) (Figure 1). The samples were collected during summer field seasons

Figure 1.



Map of the Basque provinces in France and Spain. Sampling locations – black circles, provincial capitals – ball-and-stick.

between 2000 and 2002. Lab analysis was performed in 2003-2004, and statistical analysis was conducted between 2005-2007 (with some additional analyses done as part of KY's dissertation in 2009 and revised analyses for publication in 2011). This study was approved by the University of Kansas Human Subjects Committee (HSCL #11955), and participants provided written informed consent.

#### DNA analysis

DNA extraction was performed using a standard phenol:chloroform protocol. A portion of each sample was reserved for autosomal STR analysis using the Applied Biosystems Profiler Plus Kit (Foster City, CA; USA). The samples were characterized for 9 STR loci, including D3S1358, D5S818, D7S820, D8S1179, D13S317, D18S51, D21S11, FGA, and vWA, plus the sex-determining amelogenin locus, using a multiplex PCR procedure according to manufacturer's instructions (61). Amplified products were detected using an ABI 377 DNA sequencer. DNA fragments were sized and genotyped with GeneScan 3.1 and Genotyper 2.5 software (Applied Biosystems).

#### Comparative population data

To test hypotheses of population origins, autosomal STR data from our sample were compared to geographically and/or proposed related populations, including others

TABLE 1. Previous autosomal short tandem repeat studies in the Basque population.

Population	Location	N	Loci	Analyses	Reference
Rural autochthonous Basque females	Rural Basque region	57	1	allele frequencies, population genetic parameters, genetic distances, neighbor-joining (NJ) tree	(49)
Autochthonous Basques	Basque Country	326	6	allele frequencies	(52)
Basque autochthonous residents	Guipuzcoa	50	13	allele frequencies, population genetic parameters	(54)
Unrelated autochthonous Basques	Basque Country	206	5	allele frequencies, population genetic parameters	(51)
Basque Country autochthonous individuals	Basque Country	202-208	7	allele frequencies, population genetic parameters	(50)
Unrelated "native Basques"	Alava	101	13	allele frequencies, population genetic parameters, forensic parameters	(57)
Unrelated Basque country residents	Basque Country	100	3	allele frequencies, population genetic parameters, population comparison	(48)
Unrelated autochthonous Basque students	Navarre	107	3	allele frequencies, population genetic parameters, population comparison	(59)
Unrelated autochthonous Basques	Basque Country	200	3	allele frequencies, population genetic parameters, forensic parameters, population comparison	(60)
Unrelated autochthonous Basques	Vizcaya	73	13	allele frequencies, population genetic parameters, forensic parameters, NJ tree	(56)
Unrelated autochthonous Basques	Vizcaya	68	9	allele frequencies, population genetic parameters, NJ tree, multidimensional scaling plots	(58)

from the Iberian Peninsula (Andalusia, Cantabria, Catalonia, Galicia, Murcia, Portugal, and Valencia) (54,62-67); Europe (Austria, Belgium, Bosnia, Germany, Greece, Hungary, Tuscany, Poland, Russia, Scotland, Serbia and Montenegro, Slovenia, and Switzerland) (68-81); North Africa (Egypt, Morocco) (82,83); the Middle East (Turkey) (84); and the Caucasus (Georgia) (85).

### Statistical analysis

Allele frequencies were estimated using the gene counting method. Expected heterozygosity under Hardy-Weinberg equilibrium was estimated by the method of Guo and Thompson (86), and locus-by-locus analysis of molecular variance (AMOVA) was performed to examine genetic substructure among the Basques, using Arlequin 3.11 software (87). Genetic differentiation between the Basques and 27 comparative populations, including 1 Caucasian and 2 North African groups, was measured using heterozygosity and gene differentiation ( $G_{ST}$ ).  $G_{ST}$  and gene diversity values for the Basques and comparative populations were calculated using DISPAN software (88). Genetic distances between populations were calculated using Shriver's distance ( $D_{SW}$ ), an adaptation of Nei's standard distance (D) weighted by the difference in number of repeats between alleles to account for the stepwise mutation pattern of tandem repeat loci. Genetic relationships between groups were examined using a multidimensional scaling (MDS) plot of the genetic distance matrix in NTSYS 2.1 software (89). MDS stress values were evaluated using the criteria of Sturrock and Rocha (90).

### RESULTS

Examination of the autosomal STR data revealed that allelic dropout occurred in 21% of the sample (101 of 478), so that not all loci were amplified for every individual. Samples that did not amplify for all loci were removed from the analysis. Allele frequencies for the autosomal STR loci in each of the Basque provinces after correction for allelic dropout are available in the [web extra material](#).

Observed heterozygosity values among the Basques ranged from 0.60526 (D5S818) to 0.92105 (vWA), with both extremes found in Navarre, likely as a result of small sample size ( $n=38$ ) (Table 2). In Alava and Vizcaya, only 2 loci had significantly lower heterozygosity values than expected. When the Bonferroni correction for multiple tests was applied, only the D8S1179 locus demonstrated an excess of homozygotes in all provinces, and bolded  $P$  values indicate those loci with lower than expected heterozygosity

after correction of the data for allelic dropout (Table 2). This suggests that for the other STR loci examined, the expectations of Hardy-Weinberg equilibrium were met (Table 2).

The results of the AMOVA (Table 3) suggested no obvious genetic structuring between provinces, as indicated by the among-groups covariance component ( $V_a = -0.095$ ). The lack of structure among provinces was confirmed by the global estimate of the fixation index among groups ( $F_{CT} = -0.0036$ ,

**TABLE 2. Exact test of Hardy-Weinberg Equilibrium for 9 autosomal loci in 4 Basque Provinces. Significant  $P$ -values are in bold\***

Locus	Alava	Vizcaya	Guipuzcoa	Navarre
	<b>n = 96</b>	<b>n = 89</b>	<b>n = 154</b>	<b>n = 38</b>
D3S1358				
$H_O$	0.77320	0.76404	0.76623	0.68421
$H_E$	0.80615	0.77433	0.77829	0.80000
$P$	<b>0.02230</b>	0.28137	0.34904	0.57454
FGA				
$H_O$	0.87500	0.85393	0.81818	0.81579
$H_E$	0.86044	0.88574	0.87578	0.87930
$P$	0.90451	<b>0.01880</b>	0.12451	0.14558
D5S818				
$H_O$	0.65625	0.79775	0.72727	0.60526
$H_E$	0.70610	0.72780	0.74159	0.73439
$P$	0.14642	0.82662	0.65212	0.42251
D7S820				
$H_O$	0.71875	0.77528	0.78571	0.73684
$H_E$	0.80928	0.81553	0.80132	0.82596
$P$	0.07874	0.32608	0.29700	0.70250
D8S1179				
$H_O$	0.77083	0.76404	0.75325	0.86842
$H_E$	0.80988	0.80956	0.81006	0.83474
$P$	<b>&lt; 0.00001</b>	<b>&lt; 0.00001</b>	<b>0.00082</b>	<b>0.04749</b>
vWA				
$H_O$	0.85417	0.82022	0.79870	0.92105
$H_E$	0.81086	0.81946	0.79864	0.82561
$P$	0.05728	0.88346	0.07616	0.43085
D13S317				
$H_O$	0.76042	0.76404	0.75325	0.71053
$H_E$	0.76167	0.79794	0.78282	0.76596
$P$	0.27417	0.11321	0.80221	0.56594
D18S51				
$H_O$	0.80208	0.79775	0.81818	0.81579
$H_E$	0.86938	0.88282	0.87571	0.87053
$P$	0.12041	0.09120	0.24056	0.86468
D21S11				
$H_O$	0.89583	0.83146	0.80519	0.86842
$H_E$	0.84004	0.84067	0.84073	0.82246
$P$	0.97004	0.05701	0.50981	0.53438

\*Abbreviations:  $H_O$  – observed heterozygosity;  $H_E$  – expected heterozygosity.

$P=0.892$ ). A small amount of subdivision was found between villages within provinces (1.309% total variation,  $F_{sc}=0.0131$ ,  $P=0.001$ ). Examination of the locus-by-locus results revealed that 3 loci made significant contributions to the differences between villages: D7S820 ( $F_{sc}=0.0332$ ,  $P=0.023$ ), vWA ( $F_{sc}=0.0185$ ,  $P=0.045$ ), and D18S51 ( $F_{sc}=0.0319$ ,  $P=0.021$ ). The majority of variation, however, was found between individuals within villages (99% total variation).

Average heterozygosity values by population ranged from 0.803 in Morocco to a high of 0.820 in Scotland (Table 4). Among the Basques, heterozygosity was lowest in Alava (0.805) and highest in Vizcaya (0.812). This was within the range of heterozygosity values seen in other modern Iberian populations (0.804-0.815). Total gene diversity between subpopulations ( $H_T$ ) was high, ranging from 0.724 for D5S818 to 0.878 for D18S51. However, most of this diversity is explained by variation between individuals within subpopulations ( $H_S$ ). The percentage of gene differentiation between subpopulations relative to the total gene differentiation ( $G_{st}$ ) ranged from a high of 0.009 for D13S317 and D21S11 to a low of 0.006 for D3S1358, FGA, D7S820, vWA, and D18S51.

Visual representation of genetic distances between populations using multidimensional scaling (Figure 2) showed that the Basque groups clustered together on the right side of the plot, near their neighbors in Cantabria. The North African and Georgian populations were found near the bottom center of plot, differentiated from the other European groups. A stress value of 0.169, well below the threshold of 0.317 for 27 populations in two dimensions, demonstrates that the

plot is an accurate representation of the genetic distance matrix. A Mantel test of matrix correlation between the original distance matrix and the MDS matrix also demonstrated that the MDS plot was an accurate represent of the genetic distances between populations (correlation coefficient:  $r=0.93498$ ,  $t$  test:  $t=5.7717$ ,  $P=1.0$ ).

## DISCUSSION

Our study of the autosomal STR variation did not support the hypotheses of a recent common ancestor between the Basques and populations either from the Caucasus or North Africa. Allelic dropout was noted for several samples in the present analysis, raising the possibility of a technical or genomic error in the typing of samples. At low sample DNA concentrations, the Profiler Kit is known to preferentially amplify short alleles and homozygotes (93). Because the samples collected in this study were from buccal swabs, and only a portion of each sample was used for STR analysis, DNA concentrations were much lower than if the samples had been from whole blood. The excess of homozygotes at D8S1179, even after correction for allelic dropout, was of particular concern. Concordance studies of autosomal STR typing across different multiplex kits have reported issues with the D8S1179 locus in certain populations, principally with alleles 15-18 using the Profiler Plus Kit (94). Reports of D8S1179 from previous studies among Basques give frequency ranges for alleles 15-17 between: 15 (0.110-0.210), 16 (0.010-0.029), and 17 (0.005-0.007) (54,56-58). D8S1179\*18 has not been previously reported in Basques. Frequencies for D8S1179 alleles 15-17 in the present study fall within the ranges previously re-

**TABLE 3.** Locus-by-locus analysis of molecular variance of 9 autosomal short tandem repeat loci. Negative values result from the manner in which the covariance components are estimated, from the mean squares and lower level variances rather than as sums of squares (91)\*

Locus	Among groups			Among populations			Within populations		
	percent variance	$F_{CT}$	$P$	percent variance	$F_{sc}$	$P$	percent variance	$F_{ST}$	$P$
D3S1358	-0.026	-0.0003	0.353	1.257	0.0126	0.142	98.768	0.0123	0.153
FGA	-0.161	-0.0016	0.677	-0.591	-0.0059	0.814	100.752	-0.0075	0.849
D5S818	0.384	0.0038	0.182	-0.292	-0.0029	0.646	99.908	0.0009	0.544
D7S820	-1.036	-0.0104	0.939	3.350	0.0332	0.023	97.686	0.0231	0.032
D8S1179	-0.564	-0.0056	0.904	0.573	0.0057	0.203	99.990	0.0001	0.322
vWA	0.011	0.0001	0.371	1.851	0.0185	0.045	98.138	0.0186	0.045
D13S317	-0.318	-0.0032	0.641	0.110	0.0011	0.497	100.208	-0.0021	0.614
D18S51	-0.546	-0.0055	0.504	3.204	0.0319	0.021	97.342	0.0266	0.010
D21S11	-0.308	-0.0031	0.602	0.767	0.0076	0.316	99.541	0.0046	0.321
Global estimates	-0.355	-0.0036	0.878	1.309	0.0131	0.011	99.045	0.0096	0.019
Covariance estimates		$V_a = -0.095$			$V_b = 0.351$			$V_c = 26.517$	

\*Abbreviations:  $F_{CT}$  – fixation index among groups;  $F_{sc}$  – fixation index among populations within groups;  $F_{ST}$  – fixation index within populations.

ported for this population (54,56-58), and we also found no individuals with allele 18. In addition, the locus-by-locus AMOVA demonstrated that the Basque provinces were homogeneous with respect to autosomal STR variation, and the single locus found to not be in HWE (D8S1179) did not significantly contribute to differences between provinces. Therefore, we do not believe that the failure to meet HWE in this case represents a technical error, and we included the D8S1179 allele in the interpopulation analyses.

The present analysis of autosomal STR variation does not support either the Caucasian or Vasco-Iberian hypothesis of Basque origins. Caucasian languages themselves are not a cohesive group, and while some linguists see similarities between Basque and some aspects of the northern

or southern Caucasian languages, these similarities have been attributed either to poor interpretation, a shared Euro-African substratum, or similarities in the evolution of language itself (11). Examination of the literature on the Basque-Caucasian hypothesis demonstrates little support from the genetic evidence (6,23,95-98). Cluster analysis of classical genetic markers showed that subpopulations sampled in Vizcaya were more genetically similar to each other than to other European populations or Caucasian groups outside Europe, such as those in Asia Minor and the Middle East (99). Comparison of Basque and populations from the Caucasus using 10 blood group and serum protein loci revealed that both non-Indo-European groups were more genetically similar to their neighbors than to each other (6). Analysis of HLA data showed that the Svani

**TABLE 4.** Gene diversity between populations based on autosomal short tandem repeat data

Population	D3S1358	FGA	D5S818	D7S820	D8S1179	vWA	D13S317	D18S51	D21S11	Average
Alava*	0.804	0.860	0.705	0.808	0.790	0.807	0.758	0.869	0.840	0.805
Vizcaya*	0.772	0.886	0.728	0.816	0.781	0.817	0.792	0.876	0.840	0.812
Guipuzcoa*	0.778	0.875	0.741	0.801	0.801	0.799	0.773	0.876	0.840	0.809
Navarre*	0.798	0.879	0.705	0.826	0.815	0.826	0.754	0.870	0.819	0.810
Andalusia	0.803	0.868	0.708	0.797	0.824	0.805	0.795	0.879	0.856	0.815
Cantabria	0.796	0.871	0.715	0.796	0.826	0.803	0.779	0.882	0.846	0.813
Catalonia	0.785	0.860	0.714	0.815	0.781	0.825	0.769	0.879	0.809	0.804
Galicia	0.786	0.855	0.712	0.796	0.817	0.822	0.795	0.880	0.830	0.810
Murcia	0.815	0.860	0.718	0.787	0.807	0.820	0.758	0.866	0.826	0.806
Valencia	0.800	0.872	0.702	0.803	0.826	0.807	0.781	0.875	0.839	0.812
Austria	0.806	0.864	0.709	0.808	0.814	0.806	0.801	0.872	0.854	0.815
Belgium	0.801	0.854	0.707	0.811	0.806	0.808	0.793	0.880	0.831	0.810
Bosnia	0.794	0.850	0.713	0.803	0.815	0.807	0.745	0.879	0.867	0.808
Germany	0.781	0.870	0.709	0.814	0.790	0.818	0.774	0.885	0.841	0.809
Greece	0.788	0.855	0.733	0.794	0.814	0.822	0.775	0.881	0.846	0.812
Hungary	0.794	0.864	0.728	0.798	0.809	0.805	0.787	0.886	0.855	0.814
Tuscany	0.788	0.865	0.722	0.796	0.832	0.793	0.745	0.868	0.852	0.807
Poland	0.802	0.863	0.717	0.812	0.797	0.804	0.759	0.873	0.866	0.810
Portugal	0.786	0.862	0.710	0.811	0.816	0.810	0.785	0.876	0.848	0.811
Russia	0.783	0.860	0.734	0.811	0.799	0.803	0.781	0.878	0.845	0.810
Scotland	0.799	0.856	0.727	0.804	0.834	0.811	0.828	0.864	0.855	0.820
Slovenia	0.797	0.876	0.718	0.810	0.780	0.809	0.785	0.879	0.855	0.812
Switzerland	0.791	0.869	0.725	0.821	0.830	0.809	0.773	0.877	0.842	0.815
Egypt	0.772	0.873	0.763	0.784	0.819	0.806	0.792	0.858	0.825	0.810
Morocco	0.779	0.851	0.727	0.772	0.824	0.822	0.748	0.878	0.831	0.803
Turkey	0.780	0.864	0.751	0.813	0.822	0.802	0.779	0.872	0.842	0.814
Georgia	0.775	0.871	0.749	0.810	0.806	0.765	0.746	0.874	0.855	0.806
H <sub>s</sub> <sup>†</sup>	0.788	0.862	0.720	0.801	0.808	0.807	0.773	0.872	0.839	0.807
H <sub>t</sub> <sup>‡</sup>	0.792	0.866	0.724	0.806	0.814	0.812	0.780	0.878	0.847	0.813
G <sub>ST</sub> <sup>§</sup>	0.006	0.006	0.007	0.006	0.008	0.006	0.009	0.006	0.009	0.007

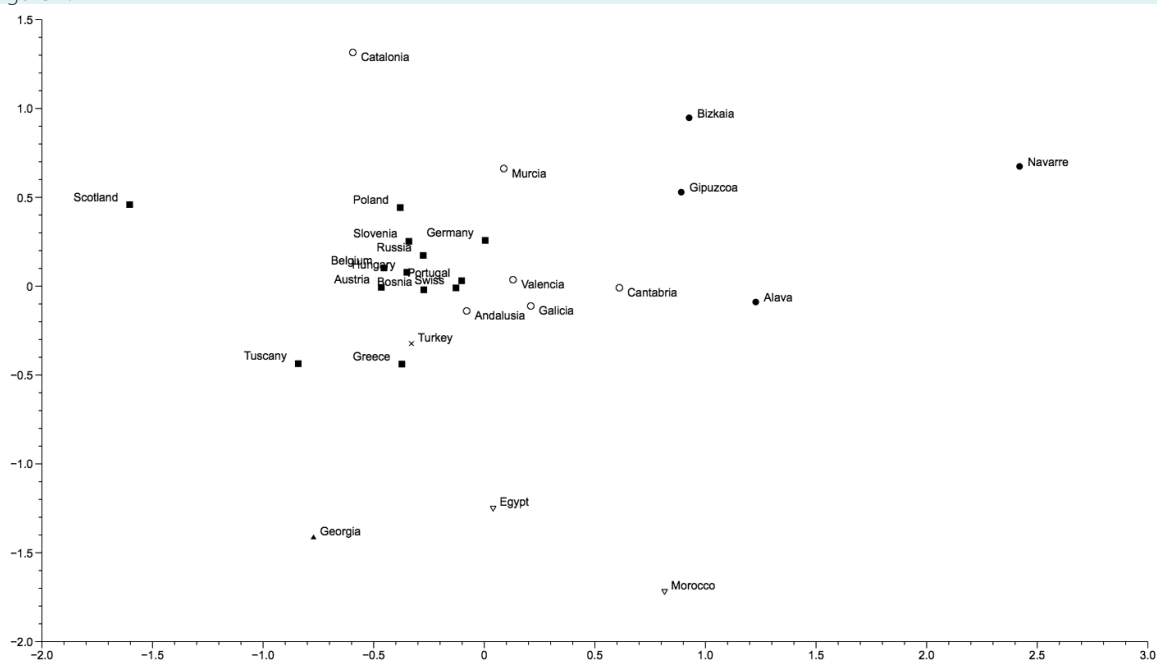
\*Present study.

†Gene diversity within subpopulations.

‡Gene diversity between subpopulations.

§Coefficient of gene differentiation (92).

Figure 2.



Multidimensional Scaling plot of genetic distance between 27 populations (Basques – closed circles, Iberia – open circles, Europe – closed squares, Middle East – cross, Caucasus – closed triangles, North Africa – open triangles). The first two axes account for 26.02% of the total genetic variation present in the sample. The stress value of 0.169, well below the threshold of 0.317 for 27 populations in 2 dimensions (90), demonstrates that the plot is an accurate representation of the genetic distance matrix. A Mantel test of matrix correlation between the original distance matrix and the MDS matrix also demonstrated that the MDS plot was an accurate represent of the genetic distances between populations (correlation coefficient:  $r=0.93498$ ,  $t$  test:  $t=5.7717$ ,  $P=1.0$ ). The Basque groups cluster together on the right side of the plot, near their neighbors in Cantabria. The North African and Georgian populations are found near the bottom center of plot, differentiated from the other European groups.

(a Kartvelian-speaking population) and the Basques were found to share only a single five-locus extended haplotype, A\*01-B\*8-DRB1\*03-DQA1\*0501-DQB1\*0201 (95). This is the most frequent HLA haplotype found in Europeans (100,101), and is present in the Svani at a frequency of 1.25% and among Basques at 2%, leading the authors to conclude that the HLA system does not support the hypothesis of a relationship between these groups.

Recent studies of molecular markers also found little similarity between Basques and populations living in the Caucasus region. Analysis of Y-SNP haplogroups found that  $F_{ST}$  values between Basques and Caucasus-dwelling groups were much greater than between Basques and surrounding Indo-European populations (96). While comparison of mtDNA sequences did reveal greater affinity between European groups and Caucasians than between West Asians and Caucasians (97), the addition of populations from Iran resulted in a genetic picture in which the Caucasus groups fell between populations from Europe

and Asia Minor with respect to mtDNA sequence variation (98). As with Y-SNPs, genetic distances based on mtDNA sequences were greater between Basques and Caucasians than between Basques and Indo-Europeans, lending credence to the hypothesis of no genetic relationship between Basques and Caucasian populations. The results of the present study agree with those previously published using other genetic markers, as genetic distances based on autosomal data place the Basque groups in a different quadrant of the MDS plot than the population from Georgia.

The Vasco-Iberian hypothesis is based partly on craniometrics, the anthropometry of head shape (102). Broca suggested, based on a sample of 60 skulls from Guipuzcoa, that the Basques were similar to populations in North Africa (103,104). However, a reanalysis of Broca's sample supplemented by the addition of 19 skulls noted no greater similarity between Basques and North African groups than any other European populations with regards to head shape (105). A more recent multivariate analysis of 20 craniomet-



ric variables in 13 Iberian populations demonstrates the unique position of Basques in the Iberian Peninsula (106). Regardless of sex, Basques were distinct in every analysis performed. The differences between Basques and other Iberian populations could not be accounted for solely by geographic distance and were instead attributed to greater age of the Basque population relative to the others.

The majority of genetic studies supporting a relationship between Basques and North African populations have been based on HLA data (8,107-111). Other genetic systems do not support a relationship between Basques and North African groups (112-114), and additional HLA analyses also found no evidence of a relationship between the two populations (29,115-119). Preliminary investigation of autosomal STRs in Vizcaya Province indicated similarity with the Basque province of Guipuzcoa, and distinction from North African groups in the Maghreb (58). The present study demonstrates the lack of relationship between Basques and populations of North Africa, as the Basque populations do not cluster near either North African population included in the MDS plot (Morocco and Egypt), but rather are found near neighboring Cantabria. Our results instead lend support to the hypothesis that the Basques are a distinct European population, with no detectable prehistoric connection to (or recent gene flow from) populations in the Caucasus or North Africa.

Dataset available from the first author upon request. The authors thank the Basque participants, Dr Arantza Apraiz for collection of the samples, and Dr Rohina Rubicz for her assistance with DNA extraction.

**Funding** This work is supported in part by National Geographic Society Grant (Project 6935-00) to the University of Kansas Laboratory of Biological Anthropology.

**Ethical approval** Received from the University of Kansas Human Subjects Committee (HSCL #11955).

**Declaration of authorship** KLY performed all of the statistical analyses and prepared and edited the manuscript. GS directed and performed all genotyping analyses included in this manuscript. RD directed genotyping and participated in drafting the manuscript. MHC directed and performed all genotyping analyses included in this manuscript.

**Competing interests** All authors have completed the Unified Competing Interest form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

## References

- Calafell F, Bertranpetit J. Principal component analysis of gene frequencies and the origin of Basques. *Am J Phys Anthropol.* 1994;93:201-15. [Medline:8147436](#) [doi:10.1002/ajpa.1330930205](#)
- Trask RL. *The history of Basque.* London, New York: Routledge; 1997.
- Blaud HC. *The Basques.* San Francisco (CA): R and E Research Associates; 1974.
- Goedde HW, Hirth L, Benkmann HG, Pellicer A, Pellicer T, Stahn M, et al. Population genetic studies of red cell enzyme polymorphisms in four Spanish populations. *Hum Hered.* 1972;22:552-60. [Medline:4362124](#) [doi:10.1159/000152537](#)
- Piazza A, Cappello N, Olivetti E, Rendine S. The Basques in Europe: a genetic analysis. *Munibe (Antropologia y Arqueologia) Suplemento.* 1988;6:169-77.
- Bertorelle G, Bertranpetit J, Calafell F, Nasidze IS, Barbujani G. Do Basque- and Caucasian-speaking populations share non-Indo-European ancestors? *Eur J Hum Genet.* 1995;3:256-63. [Medline:8528674](#)
- Cavalli-Sforza LL. The Basque population and ancient migrations in Europe. *Munibe (Antropologia y Arqueologia) Suplemento.* 1988;6:129-37.
- Arnaiz-Villena A, Martinez-Laso J, Gomez-Casado E, Diaz-Campos N, Santos P, Martinho A, et al. Relatedness among Basques, Portuguese, Spaniards, and Algerians studied by HLA allelic frequencies and haplotypes. *Immunogenetics.* 1997;47:37-43. [Medline:9382919](#) [doi:10.1007/s002510050324](#)
- Cavalli-Sforza LL, Piazza A. Human genomic diversity in Europe: a summary of recent research and prospects for the future. *Eur J Hum Genet.* 1993;1:3-18. [Medline:7520820](#)
- Calderon R, Wentzel J, Roberts DF. HLA frequencies in Basques in Spain and in neighbouring populations. *Ann Hum Biol.* 1993;20:109-20. [Medline:8447656](#) [doi:10.1080/03014469300002562](#)
- Jiménez EF. *Struggle and survival of the pre-Roman languages of the Iberian Peninsula.* Lewiston (NY): E. Mellen Press; 2001.
- Boyd WC, Boyd LG. New data on blood groups and other inherited factors in Europe and Egypt. *Am J Phys Anthropol.* 1937;23:49-70. [doi:10.1002/ajpa.1330230106](#)
- Chalmers JN, Ikin EW, Mourant AE. The ABO, MN and Rh blood groups of the Basque people. *Am J Phys Anthropol.* 1949;7:529-44. [Medline:15404766](#) [doi:10.1002/ajpa.1330070405](#)
- Chalmers JN, Ikin EW, Mourant AE. Basque blood groups. *Nature.* 1948;162:27. [Medline:18869685](#) [doi:10.1038/162027a0](#)
- Etcheverry MA. The Rhesus Factor, its genetics and clinical importance [in Spanish]. *Dia Med.* 1945;17:1237-51. [Medline:21014621](#)
- Dausset J, Legrand L, Levine MH, Qulici JC, Colombani M, Ruffie J. Genetic structure and distribution of HL-A antigens in a Basque village. In: Dausset J, Colombani, J, eds. *Histocompatibility testing.* Copenhagen: Munksgaard; 1972. p. 99-105.
- Levine MH. A hematological approach to Basque isolation in two French Basque Villages. *Ann N Y Acad Sci.* 1977;293:185-93. [Medline:411412](#) [doi:10.1111/j.1749-6632.1977.tb41812.x](#)
- Levine MH, von Hagen V, Qulici JC, Salmon D. Anthropology of a Basque village: A new hematypological study. *Cahiers*



- d'anthropologie et d'ecologie humaine. 1974;11:159-71.
- 19 MacClancy J. Biological Basques, sociologically speaking. In: M Chapman, ed. *Social and biological aspects of ethnicity*. Oxford (UK): Oxford University Press; 1993.
  - 20 Mourant AE. The blood groups of the basques. *Nature*. 1947;160:505-6. [Medline:20340597](#) [doi:10.1038/160505b0](#)
  - 21 Aguirre AI, Vicario A, Mazon LI, De Pancorbo MM, Arizti P, Estomba A, et al. AK1, PGD, GC and HP frequencies in the Basque population: a review. *Gene Geogr*. 1989;3:41-51. [Medline:2487055](#)
  - 22 Manzano C, Aguirre AI, Iriondo M, Martin M, Osaba L, de la Rúa C. Genetic polymorphisms of the Basques from Gipuzkoa: genetic heterogeneity of the Basque population. *Ann Hum Biol*. 1996;23:285-96. [Medline:8830916](#) [doi:10.1080/03014469600004522](#)
  - 23 Aguirre AI, Vicario A, Mazon LI, de Pancorbo MM, Estomba A, Lostao C. Acid phosphatase, adenosine deaminase and esterase D polymorphisms in the Spanish Basque population. *Hum Hered*. 1991;41:93-102. [Medline:1855788](#) [doi:10.1159/000153985](#)
  - 24 Manzano C, Moral P, De la Rúa C, Moreno P. Serum protein polymorphisms (GC, TF, and PI subtypes) in the Basque population of Alava. *Hum Hered*. 1993;43:121-5. [Medline:8359814](#) [doi:10.1159/000154128](#)
  - 25 Calderon R, Vidales C, Pena JA, Perez-Miranda A, Dugoujon JM. Immunoglobulin allotypes (GM and KM) in Basques from Spain: approach to the origin of the Basque population. *Hum Biol*. 1998;70:667-98. [Medline:9686480](#)
  - 26 Dugoujon JM, Clayton J, Sevin A, Constans J, Loirat F, Hazout S. Immunoglobulin (Gm and Km) allotypes in some Pyrenean populations of France. *Coll Antropol*. 1989;13:43-50.
  - 27 Esteban E, Dugoujon JM, Guitard E, Senegas MT, Manzano C, de la Rúa C, et al. Genetic diversity in northern Spain (Basque Country and Cantabria): GM and KM variation related to demographic histories. *Eur J Hum Genet*. 1998;6:315-24. [Medline:9781037](#) [doi:10.1038/sj.ejhg.5200186](#)
  - 28 Hazout S, Dugoujon JM, Loirat F, Constans J. Genetic similarity maps and immunoglobulin allotypes of eleven populations from the Pyrenees (France). *Ann Hum Genet*. 1991;55:161-74. [Medline:1952808](#) [doi:10.1111/j.1469-1809.1991.tb00409.x](#)
  - 29 Garcia Fernandez E, Arrieta A, Rinon M, Maruri N, Arranz MC, Pena JA, et al. Genetic polymorphisms of HLA class I and class II system in the Basque population. *Transplant Proc*. 1997;29:3707-9. [Medline:9414894](#) [doi:10.1016/S0041-1345\(97\)01078-6](#)
  - 30 Iriondo M, Manzano C, de la Rúa C. HLA-DQA1 in autochthonous Basques: description of a genocline for the DQA1\*0201 allele in Europe. *Int J Legal Med*. 1996;109:181-5. [Medline:9007633](#) [doi:10.1007/BF01225516](#)
  - 31 Bosch E, Calafell F, Santos FR, Perez-Lezaun A, Comas D, Benchemsi N, et al. Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. *Am J Hum Genet*. 1999;65:1623-38. [Medline:10577916](#) [doi:10.1086/302676](#)
  - 32 Calderon R, Carrion M, Perez-Miranda A, Pena JA, Dugoujon JM, Crouau-Roy B. 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*. 2003;75:117-27. [Medline:12713152](#) [doi:10.1353/hub.2003.0018](#)
  - 33 Garcia O, Martin P, Gusmao L, Albarran C, Alonso S, de la Rúa C, et al. A Basque Country autochthonous population study of 11 Y-chromosome STR loci. *Forensic Sci Int*. 2004;145:65-8. [Medline:15374597](#)
  - 34 Gonzalez-Neira A, Gusmao L, Brion M, Lareu MV, Amorim A, Carracedo A. Distribution of Y-chromosome STR defined haplotypes in Iberia. *Forensic Sci Int*. 2000;110:117-26. [Medline:10808100](#) [doi:10.1016/S0379-0738\(00\)00156-0](#)
  - 35 Quintana-Murci L, Semino O, Bandelt HJ, Passarino G, McElreavey K, Santachiara-Benerecetti AS. Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. *Nat Genet*. 1999;23:437-41. [Medline:10581031](#) [doi:10.1038/70550](#)
  - 36 Rosser ZH, Zerjal T, Hurler ME, Adojaan M, Alavantic D, Amorim A, et al. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet*. 2000;67:1526-43. [Medline:11078479](#) [doi:10.1086/316890](#)
  - 37 Adams SM, Bosch E, Balaesque PL, Ballereau SJ, Lee AC, Arroyo E, et al. The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am J Hum Genet*. 2008;83:725-36. [Medline:19061982](#) [doi:10.1016/j.ajhg.2008.11.007](#)
  - 38 Alzualde A, Izagirre N, Alonso S, Alonso A, Albarran C, Azkarate A, et al. Insights into the "isolation" of the Basques: mtDNA lineages from the historical site of Aldaieta (6th-7th centuries AD). *Am J Phys Anthropol*. 2006;130:394-404. [Medline:16425179](#) [doi:10.1002/ajpa.20375](#)
  - 39 Alzualde A, Izagirre N, Alonso S, Alonso A, de la Rúa C. Temporal mitochondrial DNA variation in the Basque Country: influence of post-neolithic events. *Ann Hum Genet*. 2005;69:665-79. [Medline:16266406](#) [doi:10.1046/j.1529-8817.2005.00170.x](#)
  - 40 Bertranpetit J, Sala J, Calafell F, Underhill PA, Moral P, Comas D. Human mitochondrial DNA variation and the origin of Basques. *Ann Hum Genet*. 1995;59:63-81. [Medline:7762985](#) [doi:10.1111/j.1469-1809.1995.tb01606.x](#)
  - 41 Corte-Real HB, Macaulay VA, Richards MB, Hariti G, Issad MS, Cambon-Thomsen A, et al. Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. *Ann Hum Genet*. 1996;60:331-50. [Medline:8865993](#) [doi:10.1111/j.1469-1809.1996.tb01196.x](#)
  - 42 Gonzalez AM, Brehm A, Perez JA, Maca-Meyer N, Flores C, Cabrera VM. Mitochondrial DNA affinities at the Atlantic fringe of Europe. *Am J Phys Anthropol*. 2003;120:391-404. [Medline:12627534](#) [doi:10.1002/ajpa.10168](#)
  - 43 Gonzalez AM, Garcia O, Larruga JM, Cabrera VM. The mitochondrial lineage U8a reveals a Paleolithic settlement in the Basque country.

- BMC Genomics. 2006;7:124. [Medline:16719915](#) [doi:10.1186/1471-2164-7-124](#)
- 44 Izagirre N, de la Rúa C. 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.* 1999;65:199-207. [Medline:10364533](#) [doi:10.1086/302442](#)
- 45 Laayouni H, Bertranpetit J. From the detection of population structure to the reconstruction of population history: the historical reading of the human genome. *Heredity.* 2009;103:362-3. [Medline:19623206](#) [doi:10.1038/hdy.2009.76](#)
- 46 Nothnagel M, Lu TT, Kayser M, Krawczak M. Genomic and geographic distribution of SNP-defined runs of homozygosity in Europeans. *Hum Mol Genet.* 2010;19:2927-35. [Medline:20462934](#) [doi:10.1093/hmg/ddq198](#)
- 47 Rodriguez-Ezpeleta N, Alvarez-Busto J, Imaz L, Regueiro M, Azcarate MN, Bilbao R, et al. High-density SNP genotyping detects homogeneity of Spanish and French Basques, and confirms their genomic distinctiveness from other European populations. *Hum Genet.* 2010;128:113-7. [Medline:20443121](#) [doi:10.1007/s00439-010-0833-4](#)
- 48 Alonso S, Castro A, Fernandez I, Gomez de Cedron M, Garcia-Orad A, Meyer E, et al. Population study of 3 STR loci in the Basque Country (northern Spain). *Int J Legal Med.* 1995;107:239-45. [Medline:7632600](#) [doi:10.1007/BF01245481](#)
- 49 Arrieta MI, Martinez B, Millan JM, Gil A, Monros E, Nunez T, et al. Study of a trimeric tandem repeat locus (SBMA) in the Basque population: comparison with other populations. *Gene Geogr.* 1997;11:61-72. [Medline:9615216](#)
- 50 Garcia O, Martin P, Budowle B, Uriarte J, Albarran C, Alonso A. Basque Country autochthonous population data on 7 short tandem repeat loci. *Int J Legal Med.* 1998;111:162-4. [Medline:9587802](#) [doi:10.1007/s004140050140](#)
- 51 Garcia O, Uriarte I, Martin P, Albarran C, Alonso A. STR data from Basque country autochthonous population. *Forensic Sci Int.* 2001;115:111-2. [Medline:11056278](#) [doi:10.1016/S0379-0738\(00\)00313-3](#)
- 52 Iriondo M, Barbero MC, Izagirre N, Manzano C. Data on six short-tandem repeat polymorphisms in an autochthonous Basque population. *Hum Hered.* 1997;47:131-7. [Medline:9156324](#) [doi:10.1159/000154401](#)
- 53 Iriondo M, de la Rúa C, Barbero MC, Aguirre A, Manzano C. Analysis of 6 short tandem repeat loci in Navarre (northern Spain). *Hum Biol.* 1999;71:43-54. [Medline:9972097](#)
- 54 Perez-Lezaun A, Calafell F, Clarimon J, Bosch E, Mateu E, Gusmao L, et al. Allele frequencies of 13 short tandem repeats in population samples from the Iberian Peninsula and northern Africa. *Int J Legal Med.* 2000;113:208-14. [Medline:10929236](#) [doi:10.1007/s004149900091](#)
- 55 Perez-Miranda AM, Alfonso-Sanchez MA, Kalantar A, Garcia-Obregon S, de Pancorbo MM, Pena JA, et al. Microsatellite data support subpopulation structuring among Basques. *J Hum Genet.* 2005;50:403-14. [Medline:16133660](#) [doi:10.1007/s10038-005-0268-2](#)
- 56 Perez-Miranda AM, Alfonso-Sanchez MA, Kalantar A, Pena JA, Pancorbo MM, Herrera RJ. Allelic frequencies of 13 STR loci in autochthonous Basques from the province of Vizcaya (Spain). *Forensic Sci Int.* 2005;152:259-62. [Medline:15978353](#) [doi:10.1016/j.forsciint.2004.09.118](#)
- 57 Perez-Miranda AM, Alfonso-Sanchez MA, Pena JA, de Pancorbo MM, Herrera RJ. Genetic polymorphisms at 13 STR loci in autochthonous Basques from the province of Alava (Spain). *Leg med (Tokyo).* 2005;7:58-61. [Medline:15556017](#)
- 58 Zlojutro M, Roy R, Palikij J, Crawford MH. Autosomal STR variation in a Basque population: Vizcaya Province. *Hum Biol.* 2006;78:599-618. [Medline:17506289](#) [doi:10.1353/hub.2007.0007](#)
- 59 Pena JA, Calderon R, Perez-Miranda A, Vidales C, Dugoujon JM, Carrion M, et al. Microsatellite DNA markers from HLA region (D6S105, D6S265 and TNFa) in autochthonous Basques from Northern Navarre (Spain). *Ann Hum Biol.* 2002;29:176-91. [Medline:11874622](#) [doi:10.1080/03014460110075675](#)
- 60 de Pancorbo MM, Castro A, Fernandez-Fernandez I, Garcia-Orad A. Population genetics and forensic applications using multiplex PCR (CSF1PO, TPOX, and TH01) loci in the Basque Country. *J Forensic Sci.* 1998;43:1181-7. [Medline:9846395](#)
- 61 Sun G, McGarvey ST, Bayoumi R, Mulligan CJ, Barrantes R, Raskin S, et al. Global genetic variation at nine short tandem repeat loci and implications on forensic genetics. *Eur J Hum Genet.* 2003;11:39-49. [Medline:12529704](#) [doi:10.1038/sj.ejhg.5200902](#)
- 62 Sanz P, Prieto V, Flores I, Torres Y, Lopez-Soto M, Farfan MJ. Population data of 13 STRs in southern Spain (Andalusia). *Forensic Sci Int.* 2001;119:113-5. [Medline:11348802](#) [doi:10.1016/S0379-0738\(00\)00422-9](#)
- 63 Zarrabeitia MT, Riancho JA. Population data on nine STRs from Cantabria, a mountainous region in northern Spain. *Forensic Sci Int.* 2001;122:175-7. [Medline:11672975](#) [doi:10.1016/S0379-0738\(01\)00496-0](#)
- 64 Gusmao L, Sanchez-Diz P, Alves C, Lareu MV, Carracedo A, Amorim A. Genetic diversity of nine STRs in two northwest Iberian populations: Galicia and northern Portugal. *Int J Legal Med.* 2000;114:109-13. [Medline:11197615](#) [doi:10.1007/s004149900106](#)
- 65 Torio A, Moya-Quiles MR, Luna A, Alvarez-Lopez MR. STR data for the AmpFISTR Profiler Plus loci from Murcia region (SE Spain). *Forensic Sci Int.* 2002;129:140-1. [Medline:12243886](#) [doi:10.1016/S0379-0738\(02\)00226-8](#)
- 66 Amorim A, Gusmao L, Alves C. STR data (AmpFISTR profiler plus) from north Portugal. *Forensic Sci Int.* 2001;115:119-21. [Medline:11056281](#) [doi:10.1016/S0379-0738\(00\)00320-0](#)
- 67 Tomas C, Picornell A, Castro JA, Misericordia M. STR data for the AmpFISTR profiler plus loci from Majorcan, Minorcan and Valencian populations (Eastern Spain). *Forensic Sci Int.* 2001;121:201-4.

- Medline:11566426 doi:10.1016/S0379-0738(01)00402-9
- 68 Zarrabeitia MT, Riancho JA, Sanchez-Diz P, Sanchez-Velasco P. 7-Locus Y chromosome haplotype profiling in a northern Spain population. *Forensic Sci Int.* 2001;123:78-80. Medline:11731203 doi:10.1016/S0379-0738(01)00521-7
- 69 Decorte R, Engelen M, Larno L, Nelissen K, Gilissen A, Cassiman JJ. Belgian population data for 15 STR loci (AmpFISTR SGM Plus and AmpFISTR profiler PCR amplification kit). *Forensic Sci Int.* 2004;139:211-3. Medline:15040919 doi:10.1016/j.forsciint.2003.11.007
- 70 Konjhodzic R, Kubat M, Skavic J. Bosnian population data for the 15 STR loci in the Power Plex 16 kit. *Int J Legal Med.* 2004;118:119-21. Medline:14991368 doi:10.1007/s00414-004-0431-x
- 71 Anslinger K, Rolf B, Keil W. Evaluation and application of the AmpF/STR profiler plus PCR amplification kit in a Bavarian population sample. *Int J Legal Med.* 2001;114:278-80. Medline:11355411 doi:10.1007/s004149900116
- 72 Sanchez-Diz P, Lareu MV, Brion M, Skitsa I, Carracedo A. STR data for the AmpFISTR Profiler Plus loci from Greece. *Forensic Sci Int.* 2002;126:265-6. Medline:12380573
- 73 Egyed B, Furedi S, Angyal M, Boutrand L, Vandenbergh A, Woller J, et al. Analysis of eight STR loci in two Hungarian populations. *Int J Legal Med.* 2000;113:272-5. Medline:11009062 doi:10.1007/s004149900112
- 74 Furedi S, Budowle B, Woller J, Padar Z. Hungarian population data on six STR loci – HUMVWFA31, HUMTH01, HUMCSF1PO, HUMFES/FPS, HUMTPOX, and HUMHPR1B – derived using multiplex PCR amplification and manual typing. *Int J Legal Med.* 1996;109:100-1. Medline:8912057 doi:10.1007/BF01355526
- 75 Ricci U, Sani I, Giunti L, Guarducci S, Coviello S, Giovannucci Uzielli ML. Analysis of 13 tetrameric short tandem repeat loci in a population of Tuscany (Central Italy) performed by means of an automated infrared sequencer. *Forensic Sci Int.* 2002;125:83-5. Medline:11852206 doi:10.1016/S0379-0738(01)00604-1
- 76 Pepinski W, Janica J, Skawronska M, Niemcunowicz-Janica A, Soltyszewski I. Population genetics of 15 STR loci in the population of Podlasie (NE Poland). *Forensic Sci Int.* 2001;124:226-7. Medline:11792517 doi:10.1016/S0379-0738(01)00603-X
- 77 Kornienko IV, Vodolazhsky DI, Ivanov PL. Genetic variation of the nine Profiler Plus loci in Russians. *Int J Legal Med.* 2002;116:309-11. Medline:12376845
- 78 Goodwin W, Scouler C, Linacre A. 13 STR loci frequency data from a Scottish population. *Forensic Sci Int.* 2001;116:187-8. Medline:11182270 doi:10.1016/S0379-0738(00)00317-0
- 79 Zupanic I, Balazic J, Komel R. Analysis of nine short tandem repeat (STR) loci in the Slovenian population. *Int J Legal Med.* 1998;111:248-50. Medline:9728751 doi:10.1007/s004140050162
- 80 Zupanic Pajnic I, Sterlinko H, Balaic J, Komel R. Parentage testing with 14 STR loci and population data for 5 STRs in the Slovenian population. *Int J Legal Med.* 2001;114:178-80. Medline:11296891 doi:10.1007/s004140000179
- 81 Gehrig C, Hochmeister M, Borer UV, Dirnhofer R, Budowle B. Swiss Caucasian population data for 13 STR loci using AmpFISTR profiler plus and coiler PCR amplification kits. *J Forensic Sci.* 1999;44:1035-8. Medline:10486953
- 82 Klitschar M, al-Hammadi N, Reichenpfader B. Population genetic studies on the tetrameric short tandem repeat loci D3S1358, VWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 and D7S820 in Egypt. *Forensic Sci Int.* 1999;104:23-31. Medline:10533274 doi:10.1016/S0379-0738(99)00092-4
- 83 Jauffrit A, El Amri H, Airaud F, Andre MT, Herbert O, Landeau-Trottier G, et al. DNA short tandem repeat profiling of Morocco. *J Forensic Sci.* 2003;48:458-9. Medline:12665020
- 84 Ascioglu F, Akyuz F, Cetinkaya U, Yilmaz S, Koluacik S, Vural B, et al. Turkish population data on nine short tandem repeat loci: HumCSF1PO, HumTH01, HumTPOX, HumFES/FPS, HumF13B, HumVWA, D3S1358, D7S820, D16S539. *Forensic Sci Int.* 2002;126:252-3. Medline:12378651
- 85 Shepard EM, Herrera RJ. Genetic encapsulation among Near Eastern populations. *J Hum Genet.* 2006;51:467-76. Medline:16557342 doi:10.1007/s10038-006-0387-4
- 86 Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics.* 1992;48:361-72. Medline:1637966 doi:10.2307/2532296
- 87 Excoffier L, Laval G, Schneider JA. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online.* 2005;1:47-50. Medline:19325852
- 88 Ota T. DISPAN: Genetic distance and phylogenetic analysis. University Park (PA): Institute of Molecular Evolutionary Genetics, Pennsylvania State University; 1993.
- 89 Rohlf FJ. NTSYS-PC: numerical taxonomy and multivariate analysis system, 2.1 ed. New York (NY): Exeter Software; 2005.
- 90 Sturrock K, Rocha J. A multidimensional scaling stress evaluation table. *Field Methods.* 2000;12:49-60. doi:10.1177/1525822X0001200104
- 91 Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 1992;131:479-91. Medline:1644282
- 92 Nei M. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci U S A.* 1973;70:3321-3. Medline:4519626 doi:10.1073/pnas.70.12.3321
- 93 Pawlowski R, Maciejewska A. Forensic validation of a multiplex containing nine STRs – population genetics in northern Poland. *Int J Legal Med.* 2000;114:45-9. Medline:11197627 doi:10.1007/s004140000139
- 94 Budowle B, Masibay A, Anderson SJ, Barna C, Biega L, Brenneke S, et al. STR primer concordance study. *Forensic Sci Int.* 2001;124:47-54. Medline:11741760 doi:10.1016/S0379-0738(01)00563-1
- 95 Sanchez-Velasco P, Leyva-Cobian F. The HLA class I and class II allele

- frequencies studied at the DNA level in the Svanetian population (Upper Caucasus) and their relationships to Western European populations. *Tissue Antigens*. 2001;58:223-33. [Medline:11782273](#) [doi:10.1034/j.1399-0039.2001.580402.x](#)
- 96 Nasidze I, Sarkisian T, Kerimov A, Stoneking M. Testing hypotheses of language replacement in the Caucasus: evidence from the Y-chromosome. *Hum Genet*. 2003;112:255-61. [Medline:12596050](#)
- 97 Nasidze I, Stoneking M. Mitochondrial DNA variation and language replacements in the Caucasus. *Proc Biol Sci*. 2001;268:1197-206. [Medline:11375109](#) [doi:10.1098/rspb.2001.1610](#)
- 98 Nasidze I, Ling EY, Quinque D, Dupanloup I, Cordaux R, Rychkov S, et al. Mitochondrial DNA and Y-chromosome variation in the Caucasus. *Ann Hum Genet*. 2004;68:205-21. [Medline:15180701](#) [doi:10.1046/j.1529-8817.2004.00092.x](#)
- 99 Aguirre A, Vicario A, Mazon LI, Estomba A, Martinez de Pancorbo M, Arrieta Pico V, et al. Are the Basques a single and a unique population? *Am J Hum Genet*. 1991;49:450-8. [Medline:1867201](#)
- 100 Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T. Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: K Tsuji, M Aizawa, T Sasazuki, eds. *Proceedings of the Eleventh International Histocompatibility Workshop and Conference*. Oxford (UK): Oxford University Press; 1991. p.1065-220.
- 101 Spinola H, Middleton D, Brehm A. HLA genes in Portugal inferred from sequence-based typing: in the crossroad between Europe and Africa. *Tissue Antigens*. 2005;66:26-36. [Medline:15982254](#) [doi:10.1111/j.1399-0039.2005.00430.x](#)
- 102 Closson CC. The hierarchy of European races. *Am J Sociol*. 1897;3:314-27. [doi:10.1086/210710](#)
- 103 Broca P. Translation of the Greater Part of the Address Delivered by M. Broca, President, at the Opening Meeting of the French Association for the Advancement of the Sciences, at the Havre Congress, 1877. *The Journal of the Anthropological Institute of Great Britain and Ireland*. 1878;7:187-200. [doi:10.2307/2841388](#)
- 104 Broca P. Letter to the Society regarding Basque crania. *Journal for the Anthropological Society of London*. 1864;2:cclxvii-cclxxiii.
- 105 Morant GM. A contribution to Basque craniometry. *Biometrika*. 1929;21:67-84.
- 106 Fox CL, Gonzalez Martin A, Vives Civit S. Cranial variation in the Iberian Peninsula and the Balearic Islands: inferences about the history of the population. *Am J Phys Anthropol*. 1996;99:413-28. [Medline:8850182](#) [doi:10.1002/\(SICI\)1096-8644\(199603\)99:3<413::AID-AJPA4>3.0.CO;2-X](#)
- 107 Arnaiz-Villena A, Rodríguez de Córdoba S, Vela F, Pascual JC, Cervero J, Bootello A. HLA antigens in a sample of the Spanish population: common features among Spaniards, Basques, and Sardinians. *Hum Genet*. 1981;58:344-8. [Medline:6948767](#) [doi:10.1007/BF00294935](#)
- 108 Arnaiz-Villena A, Benmamar D, Alvarez M, Diaz-Campos N, Varela P, Gomez-Casado E, et al. HLA allele and haplotype frequencies in Algerians. Relatedness to Spaniards and Basques. *Hum Immunol*. 1995;43:259-68. [Medline:7499173](#) [doi:10.1016/0198-8859\(95\)00024-X](#)
- 109 Martinez-Laso J, De Juan D, Martinez-Quiles N, Gomez-Casado E, Cuadrado E, Arnaiz-Villena A. The contribution of the HLA-A, -B, -C and -DR, -DQ DNA typing to the study of the origins of Spaniards and Basques. *Tissue Antigens*. 1995;45:237-45. [Medline:7638859](#) [doi:10.1111/j.1399-0039.1995.tb02446.x](#)
- 110 Arnaiz-Villena A, Martinez-Laso J, Alonso-Garcia J. The correlation between languages and genes: the Usko-Mediterranean peoples. *Hum Immunol*. 2001;62:1051-61. [Medline:11543906](#) [doi:10.1016/S0198-8859\(01\)00300-7](#)
- 111 Arnaiz-Villena A, Gomez-Casado E, Martinez-Laso J. Population genetic relationships between Mediterranean populations determined by HLA allele distribution and a historic perspective. *Tissue Antigens*. 2002;60:111-21. [Medline:12392505](#) [doi:10.1034/j.1399-0039.2002.600201.x](#)
- 112 Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J. High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. *Am J Hum Genet*. 2001;68:1019-29. [Medline:11254456](#) [doi:10.1086/319521](#)
- 113 Bosch E, Calafell F, Perez-Lezaun A, Comas D, Mateu E, Bertranpetit J. Population history of north Africa: evidence from classical genetic markers. *Hum Biol*. 1997;69:295-311. [Medline:9164042](#)
- 114 Brion M, Salas A, Gonzalez-Neira A, Lareu MV, Carracedo A. 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*. 2003;122:147-61. [Medline:12949835](#) [doi:10.1002/ajpa.10231](#)
- 115 Arnaiz-Villena A, Martinez-Laso J, Alonso-Garcia J. Iberia: population genetics, anthropology, and linguistics. *Hum Biol*. 1999;71:725-43. [Medline:10510567](#)
- 116 Brown RJ, Rowold D, Tahir M, Barna C, Duncan G, Herrer RJ. Distribution of the HLA-DQA1 and polymarker alleles in the Basque population of Spain. *Forensic Sci Int*. 2000;108:145-51. [Medline:10722200](#) [doi:10.1016/S0379-0738\(99\)00202-9](#)
- 117 Comas D, Mateu E, Calafell F, Perez-Lezaun A, Bosch E, Martinez-Arias R, et al. HLA class I and class II DNA typing and the origin of Basques. *Tissue Antigens*. 1998;51:30-40. [Medline:9459501](#) [doi:10.1111/j.1399-0039.1998.tb02944.x](#)
- 118 Comas D, Calafell F, Benchemsi N, Helal A, Lefranc G, Stoneking M, et al. Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Straits. *Hum Genet*. 2000;107:312-9. [Medline:11129330](#) [doi:10.1007/s004390000370](#)
- 119 Perez-Miranda AM, Alfonso-Sanchez MA, Pena JA, Calderon R. HLA-DQA1 polymorphism in autochthonous Basques from Navarre (Spain): genetic position within European and Mediterranean scopes. *Tissue Antigens*. 2003;61:465-74. [Medline:12823770](#) [doi:10.1034/j.1399-0039.2003.00060.x](#)