High-density SNP genotyping detects homogeneity of Spanish and French Basques, and confirms their genomic distinctiveness from other European populations

Researchers from the Co-operative Research Center, CIC bioGUNE, from the Basque Foundation for Health Innovation and Research and from the Department of Genetics, Physical Anthropology and Animal Physiology of the University of the Basque Country have presented the first high-throughput population genetic analysis including Basques from Spanish and French provinces. The study, published on-line in Human Genetics, shows that all Basques constitute a homogeneous group that can be clearly differentiated from other European populations.

Summary:

Basques are a linguistically isolated population who speak the only non Indo-European language in Western Europe, and are considered one of the European genetic outliers based on the analysis of classical markers (*i.e.*, HLA, Rh and ABO), of microsatellites data, of mitochondrial DNA (mtDNA) sequences, and Y-chromosome polymorphisms. Yet, complementary analyses based on *Alu* insertions as well as on mtDNA variation were not able to detect differences between Basques and other European populations.

Livening the debate on the genetic distinctiveness of Basques, a recent study based on 280,862 SNPs typed on pools of 30 individuals and on 109 SNPs typed individually showed no difference between Spanish-Basques and other Spanish populations and reveals that Spanish-Basques cluster with other Spanish regions rather than with French Basques, who come together with the remaining European populations included in the study. These conclusions disagree with numerous previous studies, and are not consistent with the historical and linguistic evidence that supports the distinctiveness of Basques.

In order to resolve this controversy, the researchers from CIC bioGUNE and collaborators, sampled and genotyped 83 Spanish Basque individuals that were selected under very strict criteria, and used these data to infer population structure based on more than 60,000 SNPs of several European populations. The study, which represents the first high-throughput analysis including Basques from Spanish and French provinces, shows that all Basques constitute a homogeneous group that can be clearly differentiated from other European populations (see Figs. 1 and 2).
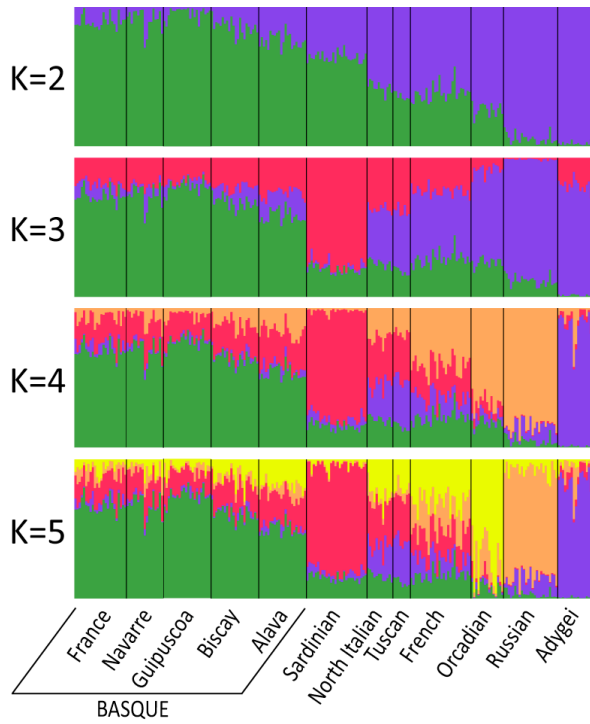
Fig1. Population structure inferred by Bayesian clustering

Graphical representation of the individual ancestry where each bar represents an individual and each color, its inferred membership in each of the K genetic clusters. Clustering was inferred with STRUCTURE v 2.2 and based on 38,228 SNPs. Analysis was performed without any prior population assignment, based on the admixture model and a burn-in period of 20,000 iterations followed by 40,000 iterations from which estimates were obtained. For each value of K, ten replicate analyses were performed and analyzed with CLUMPP (Jakobsson and Rosenberg 2007) to identify common modes as in Jakobsson et *al.* (2008). Results were plotted using DISTRUCT (Rosenberg 2004).
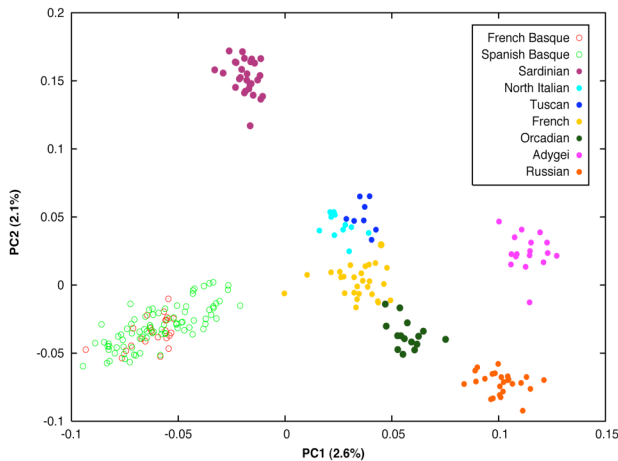

Fig. 2: Population stratification of European populations.

Graphic representation of the two principal components obtained using EIGENSOFT (Price et al. 2006) and based on 61,504 autosomal SNPs. The first and second Principal Components explain the 2.6 and 2.1% of the variation, respectively. 'Spanish Basque' label (red open circles) designates individuals form Alava, Biscay, Guipuscoa and Navarre and 'French Basque' label (green open circles), individuals from the French Basque region.