

**Group membership, geography and shared ancestry: genetic variation in the Basotho of Lesotho**

Francesco Montinaro, Joseph Davies and Cristian Capelli

Affiliation: Department of Zoology, University of Oxford, South Parks Road, OX1 3PS, Oxford UK

Number of text pages: 17

Number of figures: 2

Number of tables: 1

**Running title:** Cultural affiliation and genetic variation in Lesotho

**Corresponding author:**

Cristian Capelli

Tel: +44 (0)1865 271261

Fax: +44 (0) 1865 281 843

Email address: [cristian.capelli@zoo.ox.ac.uk](mailto:cristian.capelli@zoo.ox.ac.uk)

## **ABSTRACT**

### **Objectives**

The investigation of the evolution of cultural and genetic traits and how they interact represents a vibrant area of research in evolutionary genetics, whose dynamics are particularly relevant for our species. One of the key assumptions of the “gene-culture coevolution” framework is the co-inheritance of cultural and genetic traits. A corollary of the model is that culturally defined groups with a unique (or a limited number of) common origin(s) whose membership is inherited only through the male or female line are expected to show a relatively low intragroup variation for genetic markers similarly transmitted. Across human societies this is expected to be the case for cultural toponymies and family names within patrilineal and matrilineal groups considered in association with the non-recombining region of the Y chromosome (NRY) and the mitochondrial DNA (mtDNA) portion of the genome, respectively. This study aims at exploring the degree of correlation between culture and genetics by investigating the genetic variation of culturally and geographically defined groups.

### **Methods**

We analysed the genetic variation at NRY and mtDNA in 181 individuals from the Basotho, a Southern African patrilineal population from Lesotho, in combination with information about group membership and geographic origin.

### **Results**

Our results show that: a) the genetic distance between individuals belonging to the same culturally-defined group is lower than the population as a whole when NRY markers are considered; b) cultural traits have a bigger impact than geography for the within-group variation of Y chromosome, but not mtDNA; c) within-group genetic variation is compatible with a more homogeneous origin for less common groups.

## **Conclusions**

Our results provided additional evidence for the relevance of the dual inheritance model (culture and genetics) in understanding the patterns of human genetic variation, as implied by gene-culture co-evolution theory.

## INTRODUCTION

The dynamics through which an individual is recognised as a member of a group represents a critical part of the mechanism which defines the identity of a person. In many human societies shared origins are the basis for granting group membership: a group of individuals all claiming a direct connection, along the male or female line, to a common ancestor, often a relevant figure in the history and/or mythology of the community. These “descent groups” can be organised in various layers and can play a significant role in shaping the social dynamics within a population. The clan, here defined as a group of households reporting a shared ancestry, refers to an intermediate level within the hierarchical structuring of a given society, distributed between lineages (many of which are combined together to form a clan) and tribes (which are formed by multiple clans, (Murphy, 1986). Clans are often named after *totems* (Lévi-Strauss, 1963), symbolic animals particularly relevant within the cosmology of the group. In most human societies clan membership is patrilineal, that is children inherit the lineage of the father (Parkin and Stone, 2007).

Given their relevance in the everyday life of group members, it is relevant to ask the degree through which clan affiliation is a social construction. Do people who belong to the same clan also share a common origin? Chaix et al. (2004) addressed this question in patrilineal Central Asian populations and found that clans, and lineages within, did show evidence of common origins. However the same pattern was not observed for tribes, which in turn resulted in a combination of different clans whose members claimed a mythical common ancestor possibly to reinforce their unity. Support for a correlation between genetics and claimed ancestry was also provided for the origin of the Cameroonian Nso’ dynasty, whose Y chromosome haplogroup composition was shown to be compatible with one of the oral histories about its foundation (Veeramah et al., 2008). More recently, a correlation between Y chromosome variation and clan affiliation within a single village in Ghana has been observed (Sanchez-Faddeev et al., 2013). Common origin was also shown for individuals sharing the same surnames, the western societies version of patrilineally inherited social markers (Solé-Morata et al., 2015;

Martinez-Cadenas et al., 2015). Spanish and British (but not Irish) surnames, showed a correlation between surname occurrence and genetic diversity: the more common a surname was, the less genetically homogenous its bearers were (McEvoy and Bradley, 2006; King and Jobling, 2009; Martinez-Cadenas et al., 2015). One aspect that has not been properly explored in previous investigations has been the role played by geography in shaping the observed within-group genetic similarity. A number of reports have highlighted the correlation between geographic and genetic distance between populations and individuals (Ramachandran et al., 2005; Prugnolle et al., 2005; Novembre et al., 2008). While the study of Chaix et al (2004) was based on semi-nomadic groups, both the Cameroonians and the western European populations investigated for surnames belong to patrilocal groups. Both traits being patrilineally inherited, the distinct roles played by group membership and residency are more complex to disentangle. The increased diversity observed within common surnames could be due simply to their polygenic origin but geographic distribution might still play a relevant role in shaping such diversity. These observations leave open the question of what the strongest predictor for genetic similarity is: geography or group affiliation. In order to further explore the way cultural and biological elements are connected to each other in their intergenerational transmission, we investigated the genetic variation as shown by Y chromosome STRs and mtDNA sequences within a sample of Basotho from Lesotho for which cultural and geographic information was available (Marks et al., 2012; Marks et al., 2014). In doing so, we asked three questions:

- i) Are individuals belonging to the same Basotho clan genetically more similar to each other than a random sample?
- ii) What influences genetic similarity the most: geographic proximity or group membership?
- iii) How does the combination of geography and group affiliation affect genetic diversity among Basotho?

By addressing these questions we provided new insights into the relationships between cultural practises and genetic diversity in human populations.

## **MATERIALS AND METHODS**

### **Samples**

Detailed information of the samples here investigated is available in Marks et al. 2012 and Marks et al. 2014. Briefly, we used a dataset composed by male individuals that claimed having all four grandparents belonging to the Basotho population. Information on their clan and subclan affiliation, their place of birth and village of residency was provided during sampling. Furthermore, the corresponding information (except the village of residency) for their parents (F: Father, M: Mother) and grandparents (GF: Grandfather, GM: Grandmother) were also provided when possible.

All the samples were previously analysed for 17 NRY STRs and for mitochondrial DNA hyper variable region I (from positions 16040 to 16383). For the Y chromosome, the analyses were performed on 15 STR loci haplotypes after the removal of the bi-allelic marker DYS385. The final dataset comprised 181 individuals with no missing molecular data (Supplementary table 1).

### **Statistical Analysis**

*Metadata Analysis.* We explored the relationship between the clan and the place of birth of the donors versus those of the parents and grandfathers (GFM: grandfather (mother), GFF: grandfather (father) and grandmothers (GMM: grandmother (mother) GMF: grandmother (father)).

We graphically represented these associations via circos plots (Krzywinski et al., 2009) using the R package circlize (Gu et al., 2014; Ihaka and Gentleman, 1996), where for each donor the association between the two variables is reported as a ribbon linking the two elements. The width of the ribbon is proportional to the number of individuals showing such a link.

We tested for non-random association between all the variables (considering parents and grandparents separately) via a  $\chi^2$  test and the p-value has been computed through a Monte Carlo test with 1,000 replications. In details, the expected probability of observing

a match (mismatch) for a given variable has been estimated as  $(1 - p_C)$ , where  $p_C$  is the frequency of the variable in the dataset. The expected number of total matches and mismatches were estimated on the number of pairs reported for that variable and compared to the observed values. We explored in this way the degree of non-random association between donors and parents in relation to their clan, place of birth and village of residence. We similarly explored the degree of correlation between the three variables for the donors, the number of expected matches based on the sum of products of the frequency of the status of each variable in the two datasets compared.

*Genetic analysis.* We estimated the genetic distance between all the analysed individuals using their NRY and mtDNA haplotypes. For the Y chromosome haplotypes we used the Manhattan distance (Krause, 1986), which is the total number of mutational steps linking two haplotypes across the 15 STRs here analysed. For the mitochondrial DNA, we estimated the total number of sites which differed between each pair of haplotypes. All the calculations were performed using the R software and the functions `dist` and `dist.dna`, implemented in the packages `STATS` and `APE` (Paradis et al., 2004). Samples were pooled according to different variables (clan, sub-clan, village of residency [village], place of birth [PoB] for donors, their parents and grandparents) and related distributions were constructed by estimating the distance metrics between donors sharing the same character within a given variable (e.g. clan: it comprises all the comparisons between individuals belonging to the same clan, estimated over all the clans). We also estimated the same distances for pools of samples intersecting the clan with village or PoB of the donors (same village  $\cap$  same clan; same PoB  $\cap$  same clan). A Wilcoxon test (Wilcoxon, 1945) was performed to test for significant differences in the distribution of the genetic distances when different variables were considered. A Bonferroni correction was introduced to correct for multiple comparisons. We extended this analysis to take into account the occurrence of a given clan in the population. Given the small number of individuals often found in the less common clans in our database (Table 1), we divided the dataset in two groups (common vs rare) according to the size of each clan as reported in Ellenberger (1912), or considering the

observation of their occurrence in our dataset (number of donors belonging to a given clan). In details, we considered a clan as rare, when its effective or sample size was lower than 20,000 and 5, respectively. For the first approach we considered only those clans for which the effective size was available (Table 1). The two criteria generated similar but not identical groups (see Table 1, values in bold). We calculated the total number of pairs of haplotypes having less than 3 differences (0-2) within each clan, grouped the results in common and rare clans, and performed a  $X^2$  test between the two groups, as defined using the two clustering criteria, and the p-value have been estimated trough 1,000 Monte Carlo simulations. The range of mutational steps was chosen in accordance to the average number of differences observed in clusters within rare, very rare, and medium frequency surnames in Spanish, non-R1b Y chromosomes analysed for the same set of STRs, which better represent chromosomes with a recent single common origin (Martinez-Cadenas et al., 2015).



## RESULTS

### Metadata analysis

We analysed a dataset composed of 181 Basotho individuals from Lesotho, whose cultural and geographic information had been recorded. The percentage of donors, parents and grandparents for whom the information here shown is available is reported in Supporting Table 1. With the exception of the sub-clan (available in 16-72% of cases), other records (place of birth, clan and sub-clan affiliation for donor, parents and grandparents) were recorded in at least 62% of the cases (Supporting Table 1).

The degree of correspondence for clan affiliation and place of birth (PoB) between the donors and their relatives has been explored by circos plots, as described in the Material and Methods section (Figure 1). As expected in a group where membership is paternally determined, there is a full correspondence between the clan of the donors and their paternal male ancestors, confirming that, at least in the last generations, a full patrilineal inheritance for clan affiliation has been implemented (Figure 1A). We noted that the association was also statistically significant for the clan of the donor and of their grandmother (Father: GMF,  $p = 9e-04$ ) and grandfather (Mother: GFM,  $p = 9e-04$ ). When tested for clan-based assortative mating (mother-father pairs,  $n=181$ ) a non-random association was found (Figure 1A; note that given a full match between the clans of fathers and donors, the test between clan donor and clan mother is equivalent to the one between clan father and clan mother). However, clans resulted distributed in a non-homogenous way across the geographic regions and villages from which samples were collected in Lesotho (Supporting Figure 1, Table S1b; Marks et al, 2012; 2014), a result that possibly explains the non-random pattern of unions between clans. The importance of geography in affecting a spouse's clan is further supported by the observation that all the comparisons along the male and female line involving the place of birth resulted in statistical significance ( $p = 0.0009$ , Figure 1B), in line with the fact that the vast majority of marriages occurred between people residing in the same village (number of couples from same village = 117 out of 179; Marks et al, 2012).

## Genetic Analysis

We assessed the existence of significant differences in the genetic distance between individuals within the whole population when compared to a subset of samples defined according to their geographical or cultural affiliation.

When compared to the whole Lesotho sample, the Y chromosome distribution of genetic distances estimated within clans, subclans, same village  $\cap$  same clan, same place of birth  $\cap$  same clan, all resulted as significantly smaller ( $p < 0.05$ ; after Bonferroni correction, Figure 2A). On the other hand, none of the geographical variables considered singularly showed a significant deviation from the distribution of the whole Basotho sample (with the exclusion of the place of birth of the grandmother (Mother: GMM) (Figure 2A)

Similar comparisons based on the mtDNA distribution of genetic distances gave no significant differences from the reference Lesotho sample, except when individuals born in the same village were considered (Figure 2B).

Analysis of European surname diversity has provided evidence for a correlation between surname frequency and genetic heterogeneity (King and Jobling, 2009; Martinez-Cadenas et al., 2015). To assess the possibility that a similar pattern might also be present in relation to group membership, we pooled the individuals into two groups according to how common a clan was. We used two sets of criteria to determine the occurrence of a given clan: one based on historical data collected in previous ethnographic investigations; the other mirroring the clan distribution in our sample (Table 1). In both cases, the number of pairs having a genetic distance lower than 3 mutational steps is statistically lower in the “rare clan” group than the “common clan” group ( $p = 0.038$ ,  $X^2=6.011$  and  $p=0.017$ ,  $X^2=7.457$ ).

## DISCUSSION

Group membership represents an important cultural trait in most human societies. An individual is often recognized as a member of a group by the claim of an ancestral link with a relevant/mythological founder. The transmission of this membership can occur along the female or the male line (Radcliffe-Brown and Forde, 1987). In this context, clans, a group of families reporting a shared ancestry, are in most case patrilineally inherited, even if exceptions exist (Barnard, 2008).

Our analyses of group affiliation in Lesotho provided useful insights into the dynamics of cultural transmission in the Basotho populations. In concordance with a patrilineal inheritance pattern, we found a complete correlation for clan affiliation along the male line, suggesting that the shift of affiliation has been an extremely rare event, at least in the few last generations.

The observation of statistically significant matches between the clan of the donor and GMF and GMM, could be interpreted as a preference of marriage between individuals of the same clan, even if the presence of differential clan distribution in the village cannot be excluded. Furthermore, we found that in all the comparisons, people tend to marry more often than not individuals from the same location, confirming that geography is an important factor shaping mating patterns (Wijsman and Cavalli-Sforza, 1984; Cavalli-Sforza and Feldman, 1981).

The pattern of genetic variation observed for the paternally transmitted Y chromosome markers showed a strong correlation with patrilineal cultural traits. We observed that the Y STRs haplotypes genetic distances are statistically lower between individuals belonging to the same clan and sub-clan than the variation observed within the whole population, providing a positive answer to question i) (*are individuals belonging to the same Basotho clan genetically more similar than a random sample?*). On the other hand, with the exception of GMM ( $p = 0.001$ ), none of the genetic distances estimated considering geographical information (village and place of birth) were statistically more different than the Lesotho population. This could be interpreted as greater influence of cultural affiliation on the genetic structure of Basotho population. (ii. *What does*

*influence the most genetic similarity: geographic proximity or group membership?*). Interestingly, when the geographical information is considered together with group membership (same clan and same PoB or same village), the distribution is significantly smaller than the Basotho grouped by clan ( $p=0.004$  and  $p=0.007$  for village and place of birth, respectively) or considered as a whole ( $p=0.0005$  and  $p=0.0019$  for village and place of birth, respectively). This is consistent with a scenario where cultural affiliation plays a major role in shaping the distribution of genetic variation, with a secondary contribution by geography (*iii. How does the combination of geography and ancestry affect genetic diversity among Basotho?*), in line with the metadata analysis where the comparisons between donors and their parents were statistically significant for non-random association of their PoB.

The results for the mitochondrial DNA were in line with the expectations for a patrilineal society. No significant reduction of mtDNA diversity was observed for any variable taken in consideration. Only people born in the same place showed to be more similar to each other for mtDNA haplotype than the Basotho as a whole, possibly due to the observed behaviour of unions occurring mostly between people born in the same village.

The predictive value of genetic similarity on the basis of shared cultural traits here reported is paralleled by surnames in western societies. Such parallel extends to the correlation between genetic variation and frequency of clan. We in fact observed that less common clans, as for surnames, are genetically more homogeneous than the more common ones. As suggested for surnames, this is consistent with a less heterogeneous origin for rare groups; more common ones have probably experienced multiple origins, cuckoldry events and/or membership shift. The latter appears to have been less relevant over the last generations if we considered information provided by donors, who reported a perfect match between the clan of the father and offspring.

While the importance of migration rates and migration distances in relation to residential norms has been previously investigated (Marks et al, 2012), the results we presented here showed that geography, here in the form of co-residency as the result of reduction in male mobility, increases NRY similarity (but not mtDNA) only when other cultural aspects (e.g. group membership) are considered. The reported cumulative effect

underlines the complexity of disentangling intertwined traits when addressing the impact of culture on the genetics of human populations and suggests once again the need for considering the wider context when genetic data is analysed in relation to culture (Kumar et al., 2006).

## **CONCLUSIONS**

The impact of culture on human genetic variation has been analysed in relation to the modification of selective pressures and degree of population structure (Laland et al., 2010); Heyer et al., 2012). According to gene-culture co-evolution theory, individuals inherit both genetic and cultural information, the two elements being mutually interconnected. The feedback loop implied by this dual transmission model is shaped by the fidelity of the process and the degree of co-inheritance of the two systems (Laland et al, 2010). Our results provided additional evidence for the relevance of the dual inheritance model in understanding the patterns of human genetic variation, as implied by gene-culture co-evolution theory (Cavalli-Sforza and Feldman, 1981, Scott-Phillips et al., 2014). Additional work specifically focusing on the degree of fidelity for different cultural traits (as for example in food preferences) is expected to shed further light on these co-evolutionary dynamics.

## **ACKNOWLEDGMENTS**

We would like to thank N. Matlanyane, N. Qokolo, C. Batini and G. Busby for their contribution to the collection of the Lesotho samples and Sarah Marks and Hila Levy for the DNA analysis. We are grateful to the Lesotho Ministries of Health and Social Welfare, Local Government, Tourism, Environment and Culture for their support. We would also like to thank all the people who made this work possible by contributing with their DNA samples and all the individuals who facilitated the organization of the field work and the collection of the samples. This project was funded by the Boise Fund Trust. We also would like to thank Konstatina Isidoros for helpful suggestions. FM and CC designed the research; FM and JD performed the analysis; FM and CC wrote the

paper.

## LEGENDS

### Figures

*Figure 1: Distribution of donors' and their ancestors' cultural (A) and geographic (B) affiliation.* Each circos plot depicts the relationship between the clan (A) or place of birth: PoB (B)) of the donors and their ancestors. The width of sector and ribbons is proportional to the frequency of groups or their combination, respectively.

*Figure 2: Distributions of genetic distance of Basotho subsamples based on their cultural/geographic affiliations.* We estimated the genetic distance between each pair of individuals in subgroups defined by donors' self-affiliation and compared with the whole Basotho sample by the meaning of Wilcoxon test. Boxes in light and dark grey represent subsamples for which the distribution is significantly lower than the whole Basotho population (leftmost plot) with  $p < 0.05$  and  $p < 0.01$  (after Bonferroni correction), respectively.

*Figure S1: Clan and sampling region.* Distribution of the analysed clans in the 6 geographical regions where samples were originally collected (Marks et al, 2012, 2014).

### Tables

*Table 1:* Number of people belonging to observed clans as reported by participants (n) or previously reported (Ellenberger, 1912). Values in bold indicate sample size/census size  $< 5/25\,000$ , and therefore, were considered as rare in the analysis (excluding clans with only 1 individual; see methods).

*Table S1:* Dataset used for the analysis described in the manuscript. a) Ethnographic data analysed in this study. Each variable has been coded by a random number. b) Distribution of clans in different villages c) 15 STRs Y chromosome haplotypes. d) Mitochondrial DNA sequences.

*Table S2:* Meta-data available for the 181 individuals in the dataset (in percentages).

## REFERENCES

- Barnard A. 2008. The Co-Evolution of Language and Kinship. In: Reader NJA, Director HC, Director RD, Fellow WJE, editors. Early Human Kinship. Blackwell Publishing Ltd. p 232–243. Available from:  
<http://onlinelibrary.wiley.com/doi/10.1002/9781444302714.ch13/summary>
- Cavalli-Sforza LL, Feldman MW. 1981. Cultural Transmission and Evolution: A Quantitative Approach. Princeton University Press.
- Chaix R, Austerlitz F, Khegay T, Jacquesson S, Hammer MF, Heyer E, Quintana-Murci L. 2004. The Genetic or Mythical Ancestry of Descent Groups: Lessons from the Y Chromosome. *Am J Hum Genet* 75:1113–1116.
- Ellenberger DF. 1912. History of the Basuto, ancient and modern. Negro Universities Press.
- Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. circlize Implements and enhances circular visualization in R. *Bioinforma Oxf Engl* 30:2811–2812.
- Heyer E, Chaix R, Pavard S, Austerlitz F. 2012. Sex-specific demographic behaviours that shape human genomic variation. *Mol Ecol* 21:597–612.
- Ihaka R, Gentleman R. 1996. R: A Language for Data Analysis and Graphics. *J Comput Graph Stat* 5:299–314.
- King TE, Jobling MA. 2009. Founders, drift, and infidelity: the relationship between Y chromosome diversity and patrilineal surnames. *Mol Biol Evol* 26:1093–1102.
- Krause EF. 1986. Taxicab Geometry: An Adventure in Non-Euclidean Geometry. Courier Corporation.
- Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra



- MA. 2009. Circos: An information aesthetic for comparative genomics. *Genome Res* [Internet]. Available from: <http://genome.cshlp.org/content/early/2009/06/15/gr.092759.109>
- Kumar V, Langstieh BT, Madhavi KV, Naidu VM, Singh HP, Biswas S, Thangaraj K, Singh L, Reddy BM. 2006. Global patterns in human mitochondrial DNA and Y-chromosome variation caused by spatial instability of the local cultural processes. *PLoS Genet* 2:e53.
- Laland KN, Odling-Smee J, Myles S. 2010. How culture shaped the human genome: bringing genetics and the human sciences together. *Nat Rev Genet* 11:137–148.
- Lévi-Strauss C. 1963. *Totemism*. Beacon Press.
- Marks SJ, Levy H, Martinez-Cadenas C, Montinaro F, Capelli C. 2012. Migration distance rather than migration rate explains genetic diversity in human patrilocal groups. *Mol Ecol* 21:4958–4969.
- Marks SJ, Montinaro F, Levy H, Brisighelli F, Ferri G, Bertoncini S, Batini C, Busby GBJ, Arthur C, Mitchell P, Stewart BA, Oosthuizen O, Oosthuizen E, D'Amato ME, Davison S, Pascali V, Capelli C. 2014. Static and Moving Frontiers: The Genetic Landscape of Southern African Bantu-Speaking Populations. *Mol Biol Evol*.
- Martinez-Cadenas C, Blanco-Verea A, Hernando B, Busby GB, Brion M, Carracedo A, Salas A, Capelli C. 2015. The relationship between surname frequency and Y chromosome variation in Spain. *Eur J Hum Genet EJHG*.
- McEvoy B, Bradley DG. 2006. Y-chromosomes and the extent of patrilineal ancestry in Irish surnames. *Hum Genet* 119:212–219.
- Murphy RF. 1986. *Cultural and social anthropology: an overture*. Englewood Cliffs, N.J.: Prentice-Hall. Available from: <http://catalog.hathitrust.org/Record/000556095>
- Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, Bustamante CD. 2008. Genes mirror

- geography within Europe. *Nature* 456:98–101.
- Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinforma Oxf Engl* 20:289–290.
- Parkin R, Stone L. 2007. *Kinship and family: an anthropological reader*. Blackwell Publ.
- Prugnolle F, Manica A, Balloux F. 2005. Geography predicts neutral genetic diversity of human populations. *Curr Biol* 15:R159–R160.
- Radcliffe-Brown AR, Forde CD. 1987. *African Systems of Kinship and Marriage*. KPI.
- Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL. 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc Natl Acad Sci U S A* 102:15942–15947.
- Sanchez-Faddeev H, Pijpe J, van der Hulle T, Meij HJ, van der Gaag KJ, Slagboom PE, Westendorp RGJ, de Knijff P. 2013. The influence of clan structure on the genetic variation in a single Ghanaian village. *Eur J Hum Genet EJHG* 21:1134–1139.
- Scott-Phillips TC, Laland KN, Shuker DM, Dickins TE, West SA. 2014. The Niche Construction Perspective: A Critical Appraisal. *Evolution* 68:1231–1243.
- Solé-Morata N, Bertranpetit J, Comas D, Calafell F. 2015. Y-chromosome diversity in Catalan surname samples: insights into surname origin and frequency. *Eur J Hum Genet* [Internet]. Available from: <http://www.nature.com/ejhg/journal/vaop/ncurrent/full/ejhg201514a.html>
- Veeramah KR, Zeitlyn D, Fanzo VG, Mendell NR, Connell BA, Weale ME, Bradman N, Thomas MG. 2008. Sex-Specific Genetic Data Support One of Two Alternative Versions of the Foundation of the Ruling Dynasty of the Nso' in Cameroon. *Curr Anthropol* 49:707–714.
- Wijsman EM, Cavalli-Sforza LL. 1984. Migration and Genetic Population Structure with

Special Reference to Humans. *Annu Rev Ecol Syst* 15:279–301.

Wilcoxon F. 1945. Individual Comparisons by Ranking Methods. *Biom Bull* 1:80–83.

