International Congress Series 1288 (2006) 91–93

www.ics-elsevier.com

Y-chromosomal and mitochondrial markers: A comparison between four population groups of Italy

F. Brisighelli [a,b,*], C. Capelli [a,c], V. Álvarez-Iglesias [b], B. Arredi [a], L. Baldassarri [a], I. Boschi [a], M. Dobosz [a], F. Scarnicci [a], A. Salas [b], A. Carracedo [b], V.L. Pascali [a]

[a] *Forensic Genetics Laboratory, Istituto di Medicina Legale, Università Cattolica del S. Cuore, Roma, Italy*
[b] *Unidad de Genética, Instituto de Medicina Legal, Facultad de Medicina, Universidad de Santiago de Compostela, Galicia, Spain*
[c] *Promega Genetic Identity Europe, Promega Corporation, Madison WI, USA*

Abstract. The investigation into the genetic diversity of humans has become fundamental to the understanding of the pre-history and history of populations, and it is presently addressing crucial issues concerning human evolution that intersect with demographic, cultural and linguistic events. Numerous studies have been recently focused on the Italian Peninsula, and the current set of data regarding this country can now fit into a general frame in which local differences seems to emerge and be interpreted in the context of other cultural and historical knowledge. However, a comprehensive study based on multiple genetic systems and on extensive sampling is still missing. Here we report new data on the Y chromosome and mitochondrial DNA (mtDNA) over a significantly larger Italian sample. In particular we address four geographic sites that in the past have been the theatre of significant events in the framework of Italy's peopling: Latium (central-west), Piceno (central-east), Calabria (south-west) and Messapia (south-east). Concerning the Y polymorphisms, we based our study on STR and SNP markers in order to tackle population events positioned at various stages of the evolutionary history of Italy on the male side, and to account for local differences. In a similar way, mtDNA has been analyzed for the control region and for a selection of informative mtDNA coding region SNPs. The availability of both sets of loci including slow- and fast-evolving markers has enabled us to undertake multiple-level comparisons. We paid special interest to the distribution of genetic variability across our populations and we aimed to compare the mainframe emerging from the haploid male and female

* Corresponding author. Forensic Genetics Laboratory, Istituto di Medicina Legale, Università Cattolica del S. Cuore, Roma, Italy. Tel.: +39 635507031; fax: +39 635507033.
E-mail address: francesca.brisighelli@rm.unicatt.it (F. Brisighelli).

doi:10.1016/j.ics.2005.10.061

inherited loci. Preliminary results provided us with intriguing inferences regarding the prehistory and history of Italy.

Keywords: Y-chromosome; mtDNA; SNPs; Haplotypes; Italian Peninsula; Population substructure

1. Introduction

The identification of differences within Italian national borders is critical in understanding population history, setting up forensic reference databases and correctly interpreting association studies. The past migrations linked to specific historical events seem to have been the major cause of the current distribution of genetic variation.

The genetic structure of the present Italian population seems to reflect the ethnic stratification of pre-Roman times. In this study we report new data on the Y chromosome and mtDNA over a significant larger Italian sample. In particular we address four geographic sites: Latium (central-west, LAT), Piceno (central-east, PIC), Calabria (south-west, CAL) and Messapia (south-east, MES). Individuals were paternally unrelated, they had different surnames and they defined themselves as belonging to paternal and maternal lineages residing in the area from at least three generations.

2. Material and methods

Samples were collected by mouth swabs and blood drawing. Extraction was performed with a modified salting-out method [1]. Genetic variation was investigated by both fast (Y microsatellites/mtDNA hypervariable region I and II, HVS-I/II), and slow (SNPs) evolving markers. Nine Y-microsatellites were genotyped: DYS388, DYS393, DYS392, DYS19, DYS390, DYS391, DYS389 I and II and DYS385. PCR amplification was performed in two different multiplexes as previously described [2]. The following Y-SNPs were genotyped for all the male samples: M9, M17, M26, M35, M78, M81, M89, M173, M170, M172, M201, 92R7, 12f2, sry10831, YAP, RPS4Y and tat. We have followed the phylogenetic relationship between Y-markers and nomenclature as indicated by the Y chromosome consortium [3]. MtDNA has been characterized for the complete sequence of the control region (that includes the two hypervariable regions HVS-I/II) following conditions described by Salas et al. [4], and a selection of informative mtDNA coding region SNPs, genotyped as described in Quintáns et al. [5]. We here show some preliminary results based on the analysis of the HVS-I and SNPs. Analysis were undertaken on an ABI3100 Genetic Analyser (AB, Applied Biosystem).

3. Results

Standard statistics were calculated using Arlequin software [6]. A total of 239 individuals were genotyped for nine Y-chromosome linked STRs. Number of haplotypes per area were 44 (LAT), 52 (PIC), 51 (CAL) and 67 (MES). Unique types were 37, 46, 45 and 63, in LAT, PIC, CAL and MES, respectively. AMOVA analysis revealed significant genetic variation across populations (Fst = 0.025, $p < 0.01$). The 239 investigated chromosomes were clustered in 13 different haplogroups. We did not detect any member of the following Y lineages: A, C, N3 and P*(xR). Four Y haplogroups represented 84.55% of the total chromosomes, namely, R1*(xR1a1), J2 , G and E3b1. A total of 200 individuals were sequenced for the HVS-I: 45 LAT, 53 PIC, 50 CAL and 52 MES. All the polymorphisms observed were nucleotide substitutions. Uniquely found haplotypes were 13 in the LAT sample, 32 in PIC, 25 and 29 in the CAL and MES samples, respectively. Total sequence diversity is 0.979, while nucleotide diversity is 0.015. Mean nucleotide pairwise difference is 4.196 ± 2.093. AMOVA analysis revealed a between population variation of 1.81% (Fst = 0.018, $p < 0.01$). The 200 investigated sequences were clustered in 19 different haplogroups. We did not detect representatives of haplogroups H4, H7 and L3*. Four haplogroups represented 67% of the total chromosomes: the paraphyletic

H*, J, T, pre-V and U*. AMOVA was additionally calculated using two different geographic schemes: Central vs. South (LAT/PIC vs. CAL/MES) and West vs. East (LAT/CAL vs. PIC/MES). Y chromosome STR variation between groups was positive only for Central vs. South (3.5%, $p < 0.01$). East–West grouping revealed higher within groups variation (3.73% vs. 0.23%). The same analysis was done also for HVS-I, showing the same level of variation independently of the clustering scheme used (0.40% and 1.50%, between and within groups, respectively).

4. Discussion

The AMOVA data suggested a higher degree of structuring for Y chromosome than mtDNA. Lower male than female gene flow across populations has been suggested as a possible explanation for this observation [7], although we can not rule out differences inherent to these markers. In addition, Y chromosome data suggest a certain differentiation between central and southern Italian samples, a difference not detected in the mtDNA side. The total number of identified haplotypes and the number of uniquely found types with respect to the sample size was higher for Y chromosome than mtDNA (85% vs. 57%, 87% vs. 74%, respectively). The population sub-structure detected (especially for the Y chromosome) raises the issue of development of local reference databases to be used for forensic purposes. We have included in our analysis the Y chromosome STR locus DYS388 [8,9]. This locus has been previously evaluated for possible forensic application, but discarded due to apparent low level of informativeness. In order to evaluate the level of resolution offered by this locus, we have firstly calculated the variance of the number of repeats for each locus and population: DYS388 showed a value of 3.013 that is in the upper range of the values obtained from the other loci (from 0.4 to 1.5, not including DYS385). The exclusion of this locus resulted in a 3% reduction of the number of haplotypes. This result indicates that, for the Italian population, the inclusion of the DYS388 locus in the established panel of Y chromosome STRs would increase the power of discrimination offered by this genetic system.This work shows, at this level of resolution, substantial differences for mtDNA and Y chromosome genetic variation. Differential sex-related gene flow across geographic areas has been suggested as the major force shaping the observed results. However, additional sampling on a larger set of populations and the inclusion of other genetic systems (i.e. autosomal markers) will help to clarify this issue. We are currently extending both the genotyping and population set.

References

[1] S.A. Miller, D.D. Dykes, H.F. Polesky, Nucleic Acids Res. 16 (1988) 1215.
[2] C. Capelli, et al., Forensic Sci. Int. (in press) (2005 July 2; Epub ahead of print).
[3] Y chromosome consortium, Genome Res. 12 (2002) 339–348.
[4] A. Salas, et al., Eur. J. Hum. Genet. 6 (4) (1998) 365–375.
[5] B. Quintáns, et al., Forensic Sci. Int. 140 (2–3) (2004) 251–257.
[6] S. Schneider, D. Roessli, L. Excoffier, Genetics and Biometery Laboratory, University of Geneva, Geneva, 2000.
[7] M.T. Seilstad, E. Munch, L. Cavalli-Sforza, Nat. Genet. 20 (1998) 278–280.
[8] M. Kayser, et al., Int. J. Leg. Med. 110 (1997) 125–133.
[9] P. de Knijff, et al., Int. J. Leg. Med. 110 (3) (1997) 134–149.