International Congress Series 1239 (2003) 307–310

High-resolution analysis of male genomes by the addition of nine biallelic polymorphisms to the classic 8-STR forensic haplotype

A. Caglià [a], I. Boschi [a], F. Scarnicci [a], M. Dobosz [a], P. Underhill [b], V.L. Pascali [a], C. Capelli [a,c,*]

[a] *Department of Legal Medicine, Immunohematology Laboratory, Catholic University, Largo F. Vito, 1-00168 Rome, Italy*
[b] *Department of Genetics, Stanford University, Stanford, CA, USA*
[c] *Department of Biology, University College, London, UK*

Abstract

DNA typing of male-specific polymorphisms is a well-established procedure of molecular analysis. A haplotype of eight different human male Y-specific short tandem repeats (STRs) has been intensively used for forensic casework. This haplotype has also been effectively used to address specific problems of population genetics. A collection of 50 male genomes from our laboratory previously genotyped for 8-Y-STR has been reinvestigated with a battery of eight single nucleotide polymorphisms (SNPs) mapping to the Y-chromosome. The addition of these biallelic markers provided additional identification power. Population investigation revealed genetic structure in Italy, with notably implications in Forensic Genetics.
© 2003 Elsevier Science B.V. All rights reserved.

Keywords: Y-chromosome; SNP: single nucleotide polymorphism; STR: short tandem repeat; Population

1. Introduction

DNA typing of male-specific polymorphisms is a well-established procedure of molecular analysis. A haplotype of eight different human male Y-specific short tandem repeats (STRs) (DYS19, DYS389 I and II, DYS390, DYS391, DYS392, DYS393 and DYS385) selected in 1997 by the seminal work of a group of laboratories [1], has been

* Corresponding author. Department of Legal Medicine, Immunohematology Laboratory, Catholic University, Largo F. Vito, 1-00168 Rome, Italy. Tel.: +39-06-3550-7031; fax: +39-06-3550-7033.
E-mail address: capelli.c@iol.it (C. Capelli).

0531-5131/03 © 2003 Elsevier Science B.V. All rights reserved.
PII: S0531-5131(02)00321-7

intensively used for forensic studies and is now applied to real forensic casework. This haplotype has also been effectively used to address specific problems of population genetics [2]. One of the largest archives of population data is currently maintained for the 8-Y-STR haplotype, which qualifies to become the starting point for future activity on male identity profiles. We [3] and others [4] have used this haplotype to investigate the genetic variability of a large sample of Italian males.

Following this study, a collection of male genomes typed for the 8-Y-STR haplotype has been established at our laboratory for future typing activity.

We have now reinvestigated 50 of these genomes with a battery of nine single nucleotide polymorphisms (SNPs) mapping to the Y-chromosome (M170, M172, M173, M17, TAT, M89, M9, 12f2 and 92R7) [5].

SNPs mutation rate is really low, so mutations can be considered as unique. The non-recombining part of the Y-chromosome (NRY) is transmitted from father to son with mutations being the only changes. The genotyping of those mutations can permit to unequivocally reconstruct the genealogical relationship of the different Y-chromosomes. Using the eight above indicated biallelic markers, a total of nine different groups could be identified. In the 50 samples we genotyped, only eight of the nine markers turned out to be polymorphic, with tat [6] always showing the ancestral state. This is not surprising as this marker has been found, so far, only in northern Europe and central Asia [5,7].

2. Increase in Y-STRs haplotype resolution by inclusion of nine SNPs

The 50 samples have been previously typed also for eight Y-linked STRs. The combination of the scored alleles at the different loci is used to construct haplotypes. For identification purposes, haplotypes are collected in a database and the significance of a match is evaluated according to the correspondent frequency in the collection [8]. The resolution of the identification is related to the ability to separate haplotypes, with the ultimate goal of being able to identify all the chromosomes as unique. Increasing the number of types of STRs can increase the resolution power [9], but matching types can be still present even after. For the 50 samples we analysed, using the seven STRs haplotype, four couples resulted in being still identical. The addition of the highly polymorphic DYS385 microsatellite permitted the resolution of two of the four couples. The typing of the eight biallelic markers separated the remaining two couples, clearly indicating the increase in identification that biallelic markers can provide.

3. Population differentiation in Italy

The ability to identify genealogically related chromosomes is a key feature of slowly evolving markers. This permits the comparison of groups of chromosome known to be related, avoiding problems due to recurrent mutations occurring in STR-based haplotypes. In a previous work [10], a sample of Central North Italians was analysed using the same set of markers. Our collection of individuals could be indicated as Central South Italians, reflecting the geographical position of our laboratory. Using the genealogical groups

identified by the biallelic markers, it is possible to investigate the differences and similarities of the populations. The two groups, despite being drawn possibly from the same population (Italians) are clearly shown different frequencies of types, in particular, the frequencies of the m172 and m173 types. Using a Fisher Exact test of population differentiation (implemented by Arlequin, 1.1) [11], the two populations clearly resulted in being highly significantly different ($p < 0.001$). This result has striking implications on the evaluation of frequencies of types in different parts of Italy, suggesting a non-random distribution of types across the peninsula. This indicates that care has to be taken when considering the frequency of a certain type in a specific database, especially when considering types with geographic origin different from the reference sample.

4. Relationship with other European populations

We finally investigated the genetic relationship of our sample with other European populations [10] using Principal Component analysis (compiled by the use of the POPSTR software, Henry Harpending, personal communication). Fig. 1 shows the spatial representation of the first two components summarizing almost 80% of the total variance. The plot clearly separates Europe into at least two main groups; one continental and one related to the Mediterranean, as previously shown [10]. Intriguingly, the two Italian samples are grouped separately. The closer relationship of our sample with the Mediterranean group might reflect the dispersion of the Neolithic agricultural technology [12], but other more recent historical events, as the wide Greek colonization in the south, are possible causes.

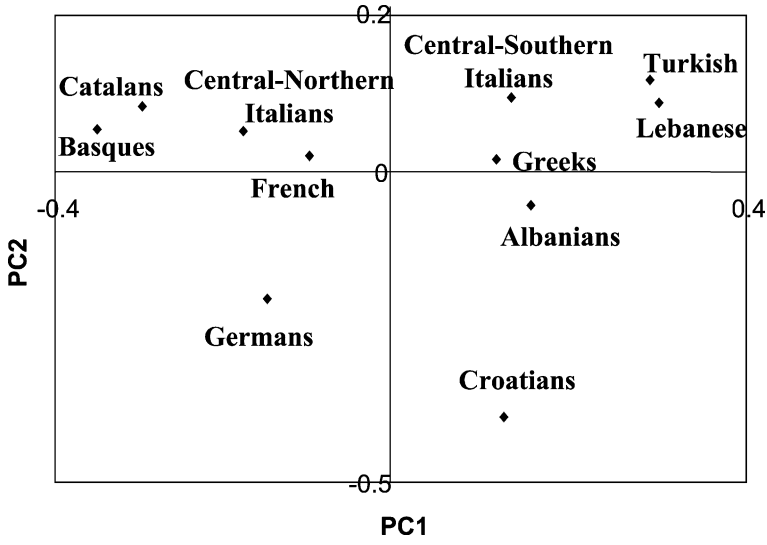

Fig. 1. Graphic representation of the principal–component analysis. Data for European populations from Ref. [10].

We are, at the moment, investigating these issues using multiple population samples from the Italian peninsula.

We have clearly shown the extreme degree of information that SNPs genotyping can provide. Increase in the power of identification is the prime impact on forensic genetics, but the geographical population structure shown by SNPs comparison is an additional issue that requires consideration. The information gathered by biallelic polymorphisms will be useful to investigate the demographic events involved.

References

[1] M. Kayser, A. Cagliá, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, F. Heidorn, S. Herrmann, B. Herzog, M. Honda, K. Honda, M.A. Jobling, M. Krawczak, K. Leim, S. Meuser, E. Meyer, W. Oesterreich, A. Pandya, W. Parson, A. Piccinini, A. Perez-Lezaun, M. Prinz, C. Schmitt, P.M. Schneider, R. Szibor, J. Teifel-Greding, G. Weichhold, P. de Knijff, L. Roewer, Evaluation of Y-chromosomal STRs: a multicenter study, Int. J. Leg. Med. 110 (1997) 125–133.

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

[3] A. Caglià, M. Dobosz, I. Boschi, E. d'Aloja, V.L. Pascali, Increased forensic efficiency of a STR-based Y-specific haplotype by addition of the highly polymorphic DYS385 locus, Int. J. Leg. Med. 111 (1998) 142–146.

[4] S. Presciuttini, A. Caglià, M. Alù, A. Asmundo, L. Buscemi, L. Caenazzo, E. Carnevali, E. Carra, Z. De Battisti, F. De Stefano, R. Domenici, A. Piccinini, N. Resta, U. Ricci, V.L. Pascali, Y-chromosome haplotypes in Italy: the GE.F.I. collaborative database. Forensic Science International, in press.

[5] P.A. Underhill, P. Shen, A.A. Lin, L. Jin, G. Passarino, W.H. Yang, E. Kauffman, B. Bonné-Tamir, J. Bertranpetit, P. Francalacci, M. Ibrahim, T. Jenkins, J.R. Kidd, S.Q. Mehdi, M.T. Seielstad, R.S. Wells, A. Piazza, R.W. Davis, M.W. Feldman, L.L. Cavalli-Sforza, P.J. Oefner, Y chromosome sequence variation and the history of human populations, Nat. Genet. 26 (2000) 358–361.

[6] T. Zerjal, B. Dashnyam, A. Pandya, M. Kayser, L. Roewer, F.R. Santos, W. Schiefenhovel, N. Fretwell, M.A. Jobling, S. Harihara, K. Shimizu, D. Semjidmaa, A. Sajantila, P. Salo, M.H. Crawford, E.K. Ginter, O.V. Evgrafov, C. Tyler-Smith, Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis, Am. J. Hum. Genet. 60 (1997) 1174–1183.

[7] T. Zerjal, L. Beckman, G. Beckman, A.V. Mikelsaar, A. Krumina, V. Kucinskas, M.E. Hurles, C. Tyler-Smith, Geographical, linguistic, and cultural influences on genetic diversity: Y-chromosomal distribution in Northern European populations, Mol. Biol. Evol. 18 (2001) 1077–1087.

[8] V.L. Pascali, M. Dobosz, B. Brinkmann, Coordinating Y-chromosomal STR research for the courts, Int. J. Leg. Med. 112 (1999) 1.

[9] A. Caglia, M. Dobosz, I. Boschi, E. d'Aloja, V.L. Pascali, Increased forensic efficiency of a STR-based Y-specific haplotype by addition of the highly polymorphic DYS385 locus, Int. J. Leg. Med. 111 (1998) 142–146.

[10] O. Semino, G. Passarino, P.J. Oefner, A.A. Lin, S. Arbuzova, L.E. Beckman, G. De Benedictis, P. Francalacci, A. Kouvatsi, S. Limborska, M. Marcikiae, A. Mika, B. Mika, D. Primorac, A.S. Santachiara-Benerecetti, L.L. Cavalli-Sforza, P.A. Underhill, The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective, Science 290 (2000) 1155–1159.

[11] S. Schneider, J.M. Kueffer, D. Roessli, L. Excoffier, Arlequin version 1.1: a Software for Population Genetic Data Analysis. Genetics and Biometry, University of Geneva, Switzerland, 1997.

[12] L.L. Cavalli-Sfrorza, P. Menozzi, A. Piazza, The History and the Geography of Human Genes, Princeton Univ. Press, Princeton, 1994.