

Announcement of population data

## Validation of a large Italian Database of 15 STR loci

Silvano Presciuttini <sup>a,1,\*</sup>, Nicoletta Cerri <sup>b,2</sup>, Stefania Turrina <sup>c,2</sup>, Benedetto Pennato <sup>a</sup>,  
Milena Alù <sup>d</sup>, Alessio Asmundo <sup>e</sup>, Anna Barbaro <sup>f</sup>, Ilaria Boschi <sup>g</sup>, Loredana Buscemi <sup>h</sup>,  
Luciana Caenazzo <sup>i</sup>, Eugenia Carnevali <sup>j</sup>, Domenico De Leo <sup>c</sup>, Cosimo Di Nunno <sup>l</sup>,  
Ranieri Domenici <sup>m</sup>, Michela Maniscalco <sup>n</sup>, Gabriella Peloso <sup>o</sup>, Susi Pelotti <sup>p</sup>,  
Andrea Piccinini <sup>q</sup>, Daniele Podini <sup>r</sup>, Ugo Ricci <sup>s</sup>, Carlo Robino <sup>t</sup>, Luigi Saravo <sup>u</sup>,  
Andrea Verzeletti <sup>b</sup>, Marina Venturi <sup>v</sup>, Adriano Tagliabracci <sup>h,3</sup>

<sup>a</sup> Center of Statistical Genetics, SS Abetone e Brennero 2, 56127 Pisa, Italy

<sup>b</sup> Department of Surgery, Radiology and Forensic Medicine, University of Brescia, Italy

<sup>c</sup> Department of Medicine and Public Health, University of Verona, Italy

<sup>d</sup> Department of Morphological and Forensic Sciences, University of Modena, Italy

<sup>e</sup> Institute of Legal Medicine, University of Messina, Italy

<sup>f</sup> Department of Forensic Genetics SIMEF, Reggio Calabria, Italy

<sup>g</sup> Institute of Legal Medicine, Catholic University, Rome, Italy

<sup>h</sup> Chair of Legal Medicine, University of Ancona, Italy

<sup>i</sup> Department of Environmental Medicine and Public Health, University of Padova, Italy

<sup>j</sup> Department of Surgery and Forensic Sciences, University of Perugia and Section of Legal Medicine, Hospital of Terni, Italy

<sup>k</sup> Department of Internal Medicine, University of Bari, Italy

<sup>l</sup> Department of Neurosciences, University of Pisa, Italy

<sup>m</sup> Andros Center srl, Palermo, Italy

<sup>n</sup> Department of Environmental Medicine and Public Health, University of Pavia, Italy

<sup>o</sup> Department of Medicine and Public Health, Section of Legal Medicine, University of Bologna, Italy

<sup>p</sup> Institute of Legal Medicine, University of Milan, Italy

<sup>q</sup> Genoma srl, Roma, Italy

<sup>r</sup> Center of Medical and Molecular Genetics, Hospital "A. Meyer", Florence, Italy

<sup>s</sup> Department of Anatomy, Pharmacology and Legal Medicine, University of Turin, Italy

<sup>t</sup> Ra.C.I.S., Section of Biology, Messina, Italy

<sup>u</sup> Department of Biomedical Sciences, Section of Legal Medicine, University of Ferrara, Italy

Received 14 December 2004; received in revised form 2 March 2005; accepted 2 March 2005

Available online 22 April 2005

### Abstract

Results from a collaborative exercise with proficiency testing conducted by 20 Italian laboratories on the 15 loci included in the Identifier<sup>®</sup> kit were analyzed by allele sharing methods and by standard population genetics tests. The validated database,

\* Corresponding author. Tel.: +39 050 2213797; fax: +39 050 2213524.

E-mail address: [sprex@biomed.unipi.it](mailto:sprex@biomed.unipi.it) (S. Presciuttini).

<sup>1</sup> Responsible for data analysis.

<sup>2</sup> Exercise coordinator.

<sup>3</sup> President of the GeFI.

including about 1500 subjects, was merged with that of a previous exercise conducted on nine loci, and the resulting allele frequencies, subdivided by Italian region, were published on-line.

© 2005 Elsevier Ireland Ltd. All rights reserved.

**Keywords:** Identifier STR; Database validation; Population data

**Population:** Twenty laboratories scattered around Italy (16 from hospitals/universities, three from private companies, one from a national criminal justice service) typed 41–197 unrelated subjects of both sexes born in their region, totaling 1541 individuals.

**Extraction:** All labs but one extracted at least part of DNA samples from blood; saliva was also used as a source by 11 labs and three labs indicated other additional sources. Extraction methods varied by laboratory; most labs used Chelex-100, others indicated Qiagen, and three used phenol–chloroform; other commercial kits were also indicated (Amersham, Promega, Epicentre, Mac/Nag).

**PCR:** The Identifier<sup>®</sup> kit (Applied Biosystems) was used by 12 labs; the combination of ProfilerPlus<sup>®</sup> + SGM Plus<sup>®</sup> + Green<sup>™</sup> I (Applied Biosystems) was used by two labs, and one lab used a custom multiplex combination; four labs used the combination ProfilerPlus<sup>®</sup> + Cofiler<sup>®</sup> (Applied Biosystems, 13 loci).

**Typing:** Electrophoresis was carried out using five-color capillary separation by 12 labs, whereas six used a four-color separation apparatus. Two labs used vertical gels. Allele call was carried out by the Genotyper<sup>®</sup> software (Applied Biosystem) by 16 labs, whereas four used visual comparison with ladder.

**Analysis of data:** Allele sharing between pairs of individuals within local datasets was analyzed with the Excel workbook AlleleSharingSheet.xls, and allele sharing among the entire database with the program AlleleSharingMacro.xls (both are available at <http://statgen.dps.unipi.it/downloads/>). Exact tests for Hardy–Weiberg (HW) equilibrium,  $F_{ST}$  analysis and tests of population differentiation were performed by Arlequin 2000 [3]. Homozygosity test was performed by Chi-square.

**Results:** Data of laboratories from the same region were merged. Allele frequencies were available for 12 (out of 20) Italian regions from north, center and south, together representing 77% of the entire Italian population (2001 census [1], Fig. 1). Data from a previous GeFI collaborative exercise [2], which included nine of the 15 loci examined here, were crosschecked against the new database; the repeated samples were discarded, and the two databases were merged. Thus, the final published tables (<http://www.gefi-forensicdna.it>) include nine loci typed in about 2800 individuals and six loci typed in more than 1500 individuals.

**Quality control:** Blind typing of two stains provided by the organizing committee.

**Other remarks:** Allele sharing analysis allowed correcting local databases for duplicate records and presence of possibly related individuals; global allele sharing analysis highlighted two pairs of individuals typed independently by different laboratories. Analysis of outlier genotypes (those with very low HW or contingency-table expectations) allowed correcting for typos. One lab sample that remained out of HW equilibrium for a locus even after applying the Bonferroni correction was discarded. Allele frequency distributions of 9 of the 15 loci have already been compared across different Italian studies [2]. The other six loci showed frequencies consistent with those published in the following reports (from groups not participating in the present exercise): D2S1338 [4], D16S539 [5], D19S43 [6], CFS1PO–TH01–TPOX [7]. The level of genetic differentiation among regions was low at all loci, so that the overall allele frequencies can be used in general forensic analyses in Italy.

This paper follows the guidelines for publication of population data requested by the journal [8].

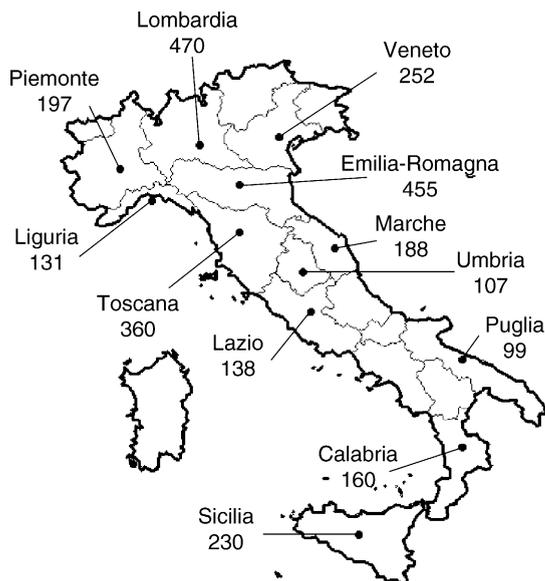


Fig. 1. Map of Italy showing regional boundaries. Numbers are sample sizes of the published database.

## References

- [1] ISTAT - 14° censimento generale della popolazione e delle abitazioni 2001. Popolazione legale (ISBN: 88-458-1069-0) Roma, 2003.

- [2] S. Presciuttini, F. Ciampini, M. Alù, N. Cerri, M. Dobosz, R. Domenici, G. Peloso, S. Pelotti, A. Piccinini, E. Ponzano, U. Ricci, Adriano Tagliabracci, J.E. Baley-Wilson, Francesco De Stefano and Vincenzo Pascali. Allele sharing in first-degree and unrelated pairs of individuals in the GeFI AmpFISTR<sup>®</sup> Profiler Plus<sup>™</sup> database, *Forensic Sci. Int.* 131 (2003) 85–89.
- [3] S. Schneider, D. Roessli, L. Excoffier, Arlequin (Version 2000): A Software for Population Genetics Data Analysis, Genetics and Biometry Laboratory, University of Geneva, Switzerland, 2000.
- [4] L. Garofano, M. Pizzamiglio, F. Donato, F. Biondi, M. Rossetti, B. Budowle, Italian population data on two new short tandem repeat loci: D2S1338 and Penta E, *Forensic Sci. Int.* 105 (1999) 131–136.
- [5] L. Garofano, M. Pizzamiglio, C. Vecchio, G. Lago, T. Floris, G. D'Errico, G. Brembilla, A. Romano, B. Budowle, Italian population data on thirteen short tandem repeat loci: HUMTH01, D21S11, D18S51, HUMVWFA31, HUMFIBRA, D8S1179, HUMTPOX, HUMCSF1PO, D16S539, D7S820, D13S317, D5S818, D3S1358, *Forensic Sci. Int.* 97 (1998) 53–60.
- [6] L. Garofano, M. Pizzamiglio, G.P. Bizzaro, F. Donato, M. Rossetti, B. Budowle, Italian population data on two new short tandem repeat loci: D6S477 and D19S433, *Forensic Sci. Int.* 101 (1999) 203–208.
- [7] R. Biondo, A. Spinella, P. Montagna, P.S. Walsh, C. Holt, B. Budowle, Regional Italian Allele frequencies at nine short tandem repeat loci, *Forensic Sci. Int.* 115 (2001) 95–98.
- [8] P. Lincoln, A. Carracedo, Publication of population data of human polymorphisms, *Forensic Sci. Int.* 110 (2000) 3–5.