Forensic Science International: Genetics 15 (2015) 127–130

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

# Expanding X-chromosomal forensic haplotype frequencies database: Italian population data of four linkage groups

Carla Bini [a,*], Laura Natalia Riccardi [a], Stefania Ceccardi [a], Francesco Carano [a], Stefania Sarno [b], Donata Luiselli [b], Susi Pelotti [a]

[a] Department of Medical and Surgical Sciences, DIMEC, Institute of Legal Medicine, University of Bologna, Italy
[b] Department of Biological, Geological and Environmental Sciences, BIGEA, Laboratory of Molecular Anthropology, University of Bologna, Italy

ARTICLE INFO

*Keywords:*
X-chromosome
Investigator Argus X-12
STR markers
Haplotype frequencies
Kinship analysis

ABSTRACT

Requests for solving complex kinship casework involving at least one female are increasing and in these circumstances the analysis of X-chromosomal STR markers plays a relevant role. Actually, it is well known the superior statistical power of X-STRs compared to autosomal markers in solving relationship when two sisters or half-sisters are involved and none of parents is available, in maternity testing or in cases involving close relatives as alternative putative fathers. In addition, the possibility to amplify more loci simultaneously and the strategy based on the analysis of four linkage groups to obtain the X-haplotype provide a powerful and validated tool. Nevertheless, haplotypes frequency distribution in different populations is still needed for calculation of probabilities in relationship testing. Published haplotype frequencies from German population data are available, but in different caseworks we found unreported X-haplotypes. To enlarge the forensic X-chromosome database, we present haplotype frequencies and other parameter of forensic interest obtained from 200 anonymous DNA samples of unrelated Italian males for the four linkage groups included in the Investigator Argus X-12 kit. From the comparison of the Italian sample haplotype frequencies with other populations, significant genetic distances were found with Asian and African populations, but not with Europeans. Finally, casework examples of complex kinship analysis are presented.



## 1. Introduction

X-chromosome short tandem repeats (X-STRs) genotyping is a powerful tool in forensic genetics for solving complex kinship cases and may efficiently complement the autosomal analysis for the higher mean exclusion chance. Due to the specific inheritance pattern of the X-chromosome (ChrX), X-STRs may be more informative about an alleged relationship than autosomes [1].

Deficiency paternity cases when the disputed child is female and other complex kinship testing as incest or cases involving close blood relatives as alternative putative fathers are the main fields of ChrX marker application. Most of ChrX regions are hemizygous in males and all daughters of a fertile male inherit the same paternal X-chromosome. Knowledge of the allelic distribution, mutation rates, genetic linkage and linkage disequilibrium (LD) status in the analysis of multiple markers located on the same chromosome is required; for this purpose, genetic data that support the application of ChrX STR markers in the forensic field are increasing but still limited. Due to lacking of available Italian population data, for our caseworks usually published haplotype frequencies from German population are used [2], but in different caseworks we found unreported X-haplotypes despite the large sample size. Accordingly, a recent collaborative multi-center study on 12 forensic X-STR loci highlighted that the most important challenge is to derive sufficiently accurate haplotype frequency estimates for the four linkage groups in different world populations for correct likelihood calculation in kinship testing and proposed that scientists share their data with the forensic genetics community [3]. Up to date, 55 X-STR markers for forensic applications on the X-chromosome web site (http://www.chrx-str.org) are characterized, but the only submitted data are from German, Chinese Han, Japanese and Ghanaian populations [4].

* Corresponding author at: Department of Medical and Surgical Sciences, Institute of Legal Medicine, University of Bologna, via Irnerio 49, 40126 Bologna, Italy. Tel.: +39 0512088343.
*E-mail address:* carla.bini@unibo.it (C. Bini).

http://dx.doi.org/10.1016/j.fsigen.2014.11.008
1872-4973/

To enlarge the forensic X-chromosome reference database and generate highly informative haplotypes for forensic purpose, in the present study we analyzed 200 unrelated Italian males from different regions across the Peninsula, Sardinia and Sicily for the four linkage groups included in the multiplex PCR Investigator Argus X-12 kit that simultaneously types 12 X-STR markers (DXS10148, DXS10135, DXS8378, DXS7132, DXS10079, DXS10074, DXS10103, HPRTB, DXS10101, DXS10146, DXS10134, DXS7423) located over the entire X-chromosome. Haplotype frequencies for all X-STR loci were compared with available data for the same markers in other worldwide population samples combining multiple distance matrices [5]. Additionally, the usefulness of the X-STRs analysis was evaluated in two caseworks of blood relationships where the alleged father/mother were not available for typing.

## 2. Materials and methods

### 2.1. Samples and DNA extraction

After having obtained a written informed consent, blood samples from 200 unrelated Italian males representative of different Italian regions were collected, in particular 74 individuals from Northern Italy, 39 from Central Italy, 55 from Southern Italy as well as 12 from Sardinia and 20 from Sicily islands, according to the criteria described in Boattini et al. [6]. Sampling strategy was based on biodemographic data and considering those individuals not related from at least three generations (Supplementary Figure S1). Moreover, the participating males were previously typed for 17 Y-STRs included in the YFiler kit (Applied Biosystems) showing all different haplotypes for at least two loci. DNA extraction was carried out using QIAamp 96 DNA Blood Kit (Qiagen, Hilden, Germany) and quantified using the Quant-iT dsDNA Broad-Range Assay Kit (Invitrogen, Carlsbad, CA). Buccal swab samples from individuals involved in kinship analysis were obtained and DNA isolation was performed by a Chelex 100 extraction protocol [7].

Supplementary material related to this article can be found, in the online version, at doi:10.1016/j.fsigen.2014.11.008.

### 2.2. Amplification and data analysis

All DNA samples were amplified using the Investigator Argus X-12 kit (Qiagen, Hilden, Germany) in a final volume of 5 μl. PCR products were separated and detected on an ABI PRISM® 310 Genetic Analyzer (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions, except for the run temperature that was set at 70 °C for an improved allelic resolution [8]. Alleles were typed using GeneMapper ID v3.2 software with the corresponding provided binset (Qiagen) and the DNA 9947A was run at each electrophoresis analysis as positive control. Forensic statistical parameters were calculated using the statistic facility in the X-chromosomal database (http://www.chrx-str.org). Allele and haplotype frequencies for the 12 X-STRs and for each linkage group (LG), as well as the pairwise genetic distances (Fst) based on haplotype frequencies, were calculated by using the Arlequin software v3.5 [9] between our Italian sample and a set of both neighboring and geographically distant populations. To maximize the effectiveness of the comparisons, only populations typed for all the same X-STR loci implemented in the Argus X-12 system were considered, for a total of 4 European (Germany [2], Denmark [10], Czech Republic [11] and Sweden [12]), 4 African (Egypt [13], Somalia [10], Ivory Coast [14] and Cabo Verde [15]) and 3 Asian (China and Japan [16], Malaysia [17]) populations. Inter-population genetic distances (Fst) between the 12 considered populations for each of the four linkage groups (LG) were integrated and graphically represented by using the R software package DISTATIS [5]. This method enables to compare a set of different distance matrices and summarizes the information enclosed within them in a single multidimensional plot (*compromise space*) which best describes the similarity structure between observations. The relative contribution of each distance matrix to the compromise plot is represented by *partial factor scores*. DISTATIS computes the compromise as an optimum linear combination of the cross-product matrices associated to each distance matrix.

## 3. Result and discussion

In two hundred unrelated Italian males typed for the 12 ChrX STR markers no shared haplotypes were found. The observed number of different haplotypes was 156, 115, 114 and 142 for LG1, LG2, LG3 and LG4, respectively (Table 1). Similar to other population studies [13,15,18], the linkage group I was the most polymorphic and linkage groups II and III were least informative showing each 71 unique haplotypes. The most common haplotype was observed 10 times in the LG3, corresponding to a population frequency of 5.0%.

Compared to the number of analyzed samples, a range from 5% to 11.3% of possible haplotypes was calculated within the 4 linkage groups, confirming that the use of observed haplotypes frequencies rather than expected haplotypes frequencies is preferable as previously reported [18].

Evaluated forensic parameters are listed in Table 2. In the present study, the DXS10135 marker showed the highest forensic efficiency with PIC of 0.93, whereas DXS8378 and DXS7423 loci presented less diversity.

No locus drop-out was found. Twelve off-ladder alleles in the linkage groups I, II and IV – all reported so far [2,11,13,15,17] except for allele 26 of DXS10079 locus – were observed and designated according to their sizes (Table 3). A lower amplification efficiency at the DXS10103 locus for most of the sample tested was observed. The estimated haplotype frequencies for the four linkage group and single allele frequencies for all 12 X-STRs are available as supplementary data (Supplementary Table S1) and will be uploaded shortly on the X-chromosomal database (http://www.chrx-str.org). The forensic statistical parameters demonstrated that the 12 X-STR loci of the Investigator Argus X-12 kit are highly informative and suitable for solving complex kinship testing.

Supplementary material related to this article can be found, in the online version, at doi:10.1016/j.fsigen.2014.11.008.

Haplotype frequencies were used to compare the results between Italians and 11 populations from Europe, Africa and Asia. Inter-population genetic distances (Fst) were considered to be statistically significant for *p*-values lower than 0.00076 after Bonferroni correction for multiple testing.

Comparisons of Italians with other European populations resulted in only minor and statistically non-significant differences for each of the four LGs. Significant Fst values at different linkage groups were instead found between Italians and both Asian and African populations (the only exception being comparison with

**Table 1**
Distribution of STR haplotypes for the ChrX linkage groups ($n = 200$ males).

| Linkage group | Number of possible haplotypes | Number of different haplotypes | Number of unique haplotypes | Frequency of the most common haplotype |
|---|---|---|---|---|
| I | 3120 | 156 | 119 | 0.02 |
| II | 1296 | 115 | 71 | 0.03 |
| III | 1008 | 114 | 71 | 0.05 |
| IV | 1805 | 142 | 102 | 0.03 |