Forensic Science International: Genetics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Short communication

Genetic portrait of Jewish populations based on three sets of X-chromosome markers: Indels, Alu insertions and STRs

J.F. Ferragut[a], K. Bentayebi[b], R. Pereira[c,d], J.A. Castro[a], A. Amorim[c,d,e], C. Ramon[a], A. Picornell[a,*]

[a] Institut Universitari d'Investigació en Ciències de la Salut (IUNICS) i Laboratori de Genètica, Departament de Biologia, Universitat de les Illes Balears, Carretera de Valldemossa, km 7.5, 07122, Palma de Mallorca, Illes Balears, Spain
[b] Biochemistry and Immunology Laboratory-Forensic Unit, Mohammed V Agdal University, Faculty of Sciences, Avenue Ibn Batouta Agdal, BP 1014 Rabat, Morocco
[c] i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Rua Alfredo Allen, 208, 4200-135 Porto, Portugal
[d] IPATIMUP – Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Rua Júlio Amaral de Carvalho, 45, 4200-135 Porto, Portugal
[e] Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, s/n, 4169-007 Porto, Portugal

ARTICLE INFO

Keywords:
X-chromosome
Alu insertions
X-STR
X-indel
Jews
Chuetas
Majorca
Investigator Argus X-12 kit

ABSTRACT

Population genetic data for 53 X-chromosome markers (32 X-indels, 9 X-Alu insertions and 12 X-STRs) are reported for five populations with Jewish ancestry (Sephardim, North African Jews, Middle Eastern Jews, Ashkenazim, and Chuetas) and Majorca, as the host population of Chuetas.

Genetic distances between these populations demonstrated significant differences, except between Sephardic and North African Jews, with the Chuetas as the most differentiated group, in accordance with the particular demographic history of this population. X-chromosome analysis and a comparison with autosomal data suggest a generally sex-biased demographic history in Jewish populations. Asymmetry was found between female and male effective population sizes both in the admixture processes between Jewish communities, and between them and their respective non-Jewish host populations.

Results further show that these X-linked markers are highly informative for forensic purposes, and highlight the need for specific databases for differentiated Jewish populations.

1. Introduction

Modern Jews comprise an aggregate of ethno-religious communities that can be traced back to a national and religious group originating several thousand years ago, and maintaining continuous cultural, historical, and religious traditions since that time. Historical evidence suggests a Middle Eastern origin, followed by a series of migrations leading to the establishment of communities of Jews worldwide, in what is termed the Jewish Diaspora. These communities can be classified on the basis of their main regions of residence: (i) the Sephardim who, after their expulsion from the Iberian Peninsula in the late 15th century, migrated to other Mediterranean countries where they mixed with local Jewish communities [1,2]; (ii) the North African Jews, of whom there is evidence in North Africa as early as the first centuries AD. These communities were augmented as a consequence of the Spanish expulsion [1,3,4]; (iii) Middle Eastern Jews (Iran and Iraq), who originated from Babylonian or Persian communities in the fourth to sixth centuries BC [5,6]; and (iv) the Ashkenazim, who have lived since the first millennium of the common era in central and Eastern Europe, but whose origins remain controversial to this day [7,8].

Chuetas are a group of descendants of the Jewish population living in Majorca (Balearic Islands, Spain) who, despite their official conversion to Christianity (1391–1435), were discriminated against and isolated from the old-Christian Majorcan population until the middle of the 20th century. Chuetas, together with some Crypto-Jewish communities in Portugal [9], are the only current Iberian populations whose ancestors can be traced back to the original Sephardic Jewish populations, given their peculiar history that kept the memory of their Jewish origin over the centuries. Unlike what happened with most of the converted Iberian Jews, their inbreeding has restricted their gradual assimilation into the general population [10].

Many historical and demographic events have shaped the genetic portrait of these groups, such as religious conversion, assimilation, bottlenecks, and intermarriage with different populations as a consequence of their various migrations. This complex demographic history imposes special challenges to better understanding the origins and genetic structure of these groups. For this reason, they have been the focus of genetic studies since the turn of the 20th century. These studies have provided evidence for shared Middle Eastern ancestry between major Jewish Diaspora groups, and variable degrees of admixture with

* Corresponding author at: Laboratori de Genètica, Departament de Biologia, Universitat de les Illes Balears, Cra. Valldemossa, km 7.5, 07122 – Palma de Mallorca, Illes Balears, Spain.
E-mail address: apicornell@uib.es (A. Picornell).

http://dx.doi.org/10.1016/j.fsigen.2017.09.008
Received 25 May 2017; Received in revised form 25 August 2017; Accepted 12 September 2017

local populations [e.g. 11–16]. Regarding Chuetas, genetic studies have shown that this population presents a significant persistence of Jewish heritage as well as signs of introgression from their non-Jewish host population [e.g. 17,18].

The choice of the X-chromosome comes from the many features it affords, making it a good source of information for population genetics and anthropology, and an important tool in forensic cases [19–21]. Compared with autosomes, the X-chromosome has a lower recombination rate, lower mutation rate, and a smaller effective population size (Ne), resulting in faster genetic drift. Consequently, both linkage disequilibrium (LD) and population structure in the X-chromosome are expected to be stronger than in the autosomes. On the other hand, since two thirds of X-chromosome history has been spent in females, X-chromosome polymorphisms mainly reflect the history of females [22]. Finally, in kinship analysis, such as in cases of father-daughter, mother-son, grandmother-granddaughter, or putative sisters testing, X-chromosome markers are an extremely useful source of information [23].

Population data on X-chromosome markers can be considered very scarce for Jewish groups [24–27]. The present work focuses on the comprehensive analysis of 53 X-chromosomal markers of different types – 32 insertion-deletion polymorphisms (Indels), 9 Alu insertions, and 12 STRs – aiming to evaluate their usefulness in a forensic context and to contribute to refining knowledge regarding the complex system of interrelationships between Jewish communities and their non-Jewish neighbours.

2. Material and methods

2.1. DNA samples

DNA samples from 500 unrelated individuals (276 males and 224 females) were obtained after informed consent: 402 with known Jewish ancestry, and 98 from Majorcan individuals, included in the study as the host population of Chuetas.

Samples from Jewish populations were 281 individuals of the National Laboratory for the Genetics of Israeli Populations at Tel-Aviv University. Following classical criteria, these samples were categorized into four groups: Sephardic (65 Turkish and 44 Bulgarian), North African (35 Moroccan, 13 Libyan, and 12 Tunisian), Middle Eastern (30 Iranian and 27 Iraqi) and Ashkenazi (55). The 121 Chueta individuals and 98 individuals from Majorca (Balearic Islands, Spain) belonged to the collection of the Genetics Laboratory, University of the Balearic Islands.

2.2. Genetic markers and genotyping

Samples were typed for three sets of X-chromosome genetic markers: (i) 32 X-indels previously reported by Pereira et al. [28]; (ii) a set of 9 X-chromosome Alu insertions (Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634, Ya5DP77, Ya5NBC491, Yb8NBC578, Ya5DP4 and Ya5DP13) described by Callinan et al. [29]; and (iii) 12 X-STRs included in the Investigator Argus X-12 kit (Qiagen GmbH, Hilden, Germany). X-STRs of Majorcan individuals have been published elsewhere [30].

Amplification was performed according to the manufacturer's instructions (X-STRs) or using previously described protocols (X-indels according to Pereira et al. [28], and X-Alu insertions as in González-Pérez et al. [31]).

To genotype Alu sequences, 15 µl of the PCR products were run in 2% 1 x TBE agarose gels containing ethidium bromide, and reaction products were directly visualized using ultraviolet fluorescence. For STRs and Indels, PCR products were separated by capillary electrophoresis on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA) and analysed with GeneMapper ID v3.2 (Applied Biosystems).

2.3. Statistical analysis

Allele frequencies, exact test of Hardy-Weinberg equilibrium (HWE) for female samples, pairwise exact test of linkage disequilibrium (LD), and haplotype diversity (HD) for male samples were estimated using Arlequin v.3.5.1.2 software [32]. Statistical parameters of forensic interest were computed using Genoproof3 theory manual formulae [33] through the Forensic X-chromosome STR homepage (http://www.chrx-str.org).

In order to examine the relationship between the populations studied and with other published data, pairwise F_{ST} genetic distances were calculated with POPTREE2 software [34], while the Analysis of Molecular Variance (AMOVA) and corresponding non-differentiation *p*-values were assessed using Arlequin v3.5.1.2. For easier visualization of the observed genetic distances, a multidimensional scaling (MDS) plot of the pairwise F_{ST} matrix was represented using SPSS v.15.0 (SPSS, Inc., Chicago, IL, USA).

3. Results and discussion

3.1. Genetic diversity

Allele frequency data for the different types of X-chromosome markers studied in five populations with Jewish ancestry (Sephardic, North African, Middle Eastern, Ashkenazi and Chuetas) and in Majorca are included in Supplementary Tables 1–3.

Values of average gene diversity across loci in the studied populations are summarized in Supplementary Tables 4–6. Statistically significant deviations were found in several Alu insertion polymorphisms in Sephardim (Supplementary Table 5) when assuming the conventional level of significance ($p < 0.05$). However, after Bonferroni correction for multiple tests, all of the deviations lost statistical significance.

Most X-indels showed diversities in the highest range of the possible values for biallelic markers (≤ 0.5), except MID2637, MID3753, MID3692 and MID3727, with average values below 0.25. Average gene diversity was similar in all populations, ranging between 0.368 and 0.384, in line with other studies [35,36].

Five X-Alu insertions (Ya5DP62, Yb8DP49, Yd3JX437, Yb8NBC634 and Ya5DP13) were revealed to be polymorphic in all populations, whereas the others appeared as monomorphic in at least one studied population (Ya5DP4 insertion in Middle Eastern Jews, Ya5DP77 absence in Sephardic, and Ya5NBC491 and Yb8NBC578 absence in Ashkenazi Jews). Most Alu elements showed moderate to low diversity. Ya5DP62 displayed the highest heterozygosity (0.374) and Ya5NBC491 the lowest (0.063). Average gene diversity in the Jewish populations for this set of X-chromosome Alu insertions ranged from 0.156 to 0.171, in accordance with heterozygosity ranges found in other Mediterranean populations [29,37,38].

All X-STRs were highly polymorphic in all populations (average gene diversity ranged from 0.787 in Chuetas to 0.805 in North African Jews). Locus-by-locus analyses revealed that DXS10135 had the highest diversity (with 28 alleles and average heterozygosity of 0.926), whilst DXS10103 and DXS7423 were the least diverse markers (mean heterozygosity of 0.681), as described elsewhere [30,35,39].

Typing of the 149 males from populations with Jewish ancestry resulted in 148 different haplotypes when all 12 X-STRs were included (Table 1); one haplotype was shared between two Sephardic individuals who are, to our knowledge, unrelated. Linkage groups (LG) 1–4 revealed 124, 96, 93, and 108 haplotypes, respectively (Supplementary Table 7). Out of all the observed haplotypes, 96% showed frequencies < 0.030, and the most common haplotype was observed in seven Chueta individuals in LG3, displaying a frequency of 0.047. In these populations, LG1 proved to be the most polymorphic group and LG3 the least variable, similarly to other studies [e.g. 30,40]. The Sephardic population had the lowest haplotype diversity value (0.995).