Forensic Science International: Genetics 21 (2016) 1–4

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Short communication

Genetic diversity of 38 insertion–deletion polymorphisms in Jewish populations

J.F. Ferragut[a], R. Pereira[b,c], J.A. Castro[a], C. Ramon[a], I. Nogueiro[b,c], A. Amorim[b,c,d], 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] i3S—Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal
[c] IPATIMUP—Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto, Portugal
[d] Faculdade de Ciências da Universidade do Porto, Porto, Portugal

ARTICLE INFO

Article history:
Received 19 August 2015
Received in revised form 22 October 2015
Accepted 4 November 2015

Keywords:
Indels
Jews
Ashkenazim
Sephardim
Mizrahim
Chuetas
Bragança

ABSTRACT

Population genetic data of 38 non-coding biallelic autosomal indels are reported for 466 individuals, representing six populations with Jewish ancestry (Ashkenazim, Mizrahim, Sephardim, North African, Chuetas and Bragança crypto-Jews). Intra-population diversity and forensic parameters values showed that this set of indels was highly informative for forensic applications in the Jewish populations studied. Genetic distance analysis demonstrated that this set of markers efficiently separates populations from different continents, but does not seem effective for molecular anthropology studies in Mediterranean region. Finally, it is important to highlight that although the genetic distances between Jewish populations were small, significant differences were observed for Chuetas and Bragança Jews, and therefore, specific databases must be used for these populations.

1. Introduction

Indels are length polymorphisms created by small insertions or deletions of one or more nucleotides in the genome. These markers have a series of special characteristics that justify its increasing interest as a tool for a wide range of purposes, including population genetic studies, ancestry affiliation, and forensics. Indels combine many of the desirable characteristics of both SNPs and STRs, such as a widely spread distribution throughout the genome [1–3]; origin from a single rare mutation event and so unlikely to present recurrent mutations [4]; amenable to PCR design of short amplicons, improving the chances of amplification success of degraded samples [5,6]; significant differences in allele frequencies between populations, providing good results in population differentiation studies, and the ease of analysis by multiplex PCR and capillary electrophoresis [1,7,5]. In 2009, Pereira et al. [5] described a new multiplex for human identification combining 38 small non-coding biallelic autosomal indels into a single multiplex, with proven success for genotyping of degraded samples.

Jews can be traced back to populations occupying a small geographic area, in the Middle East, several thousand years ago and have maintained continuous cultural and religious traditions despite a series of Diasporas. Contemporary Jews comprise several communities that can be classified according to the location where each community developed. Among others, these include Middle Eastern Jews ("Mizrahim") (Iran and Iraq), who have always resided in the Near East; the Askenazim who lived in communities of central and eastern Europe; the Sephardim (from Sepharad, Hebrew word for Hispania) who, after their expulsion from the Iberian Peninsula in the late 15th century, lived in other Mediterranean countries (especially Bulgaria and Turkey); and the North African Jews, comprising both Sephardim and Mizrahim [8–10].

Chuetas (Majorca, Spain) and the crypto-Jewish communities in Portugal (especially in Belmonte and Bragança district) are the only current Iberian populations whose ancestors can be traced to the original Sephardic Jewish populations, because of their peculiar history that kept the memory of their Jewish origin through centuries, and their inbreeding, which has prevented their gradual assimilation into the general population, as it happened with most of converted Iberian Jews [11,12].

Geneticists have studied Jewish populations since the turn of the 20th century in an attempt to unravel what must be a

* Corresponding author. Fax: +34 971 173184.
E-mail address: picornell@uib.es (A. Picornell).

http://dx.doi.org/10.1016/j.fsigen.2015.11.003
1872-4973/

complex system of interrelationships among Jewish communities and their non-Jewish neighbours. These studies have provided evidences for shared Middle Eastern ancestry among major Jewish Diaspora groups and variable degrees of admixture with local populations [13–19]. Regarding the Jewish descendants in Iberia, genetic studies have shown that both Chuetas and Bragança Jews present a significant persistence of a Jewish heritage as well as signs of introgression from their host non-Jewish populations [20–22].

In the present study we aimed to characterize the diversity of the 38 indel markers [5] in populations with Jewish ancestry, and compare their distribution in Jewish and non-Jewish populations. Finally we evaluated the usefulness of the indel set both in forensic casework and population genetics, especially in Chuetas and in Bragança crypto-Jewish populations, due to the genetic heritage of their singular history.

2. Materials and methods

2.1. DNA samples

DNA samples from 466 unrelated individuals with known Jewish ancestry were obtained after informed consent: 136 Chueta individuals (Majorca, Spain) belonged to the collection of the Genetics Laboratory, University of the Balearic Islands; 55 Jews from Bragança (NE Portugal) collected by the Institute of Pathology and Molecular Immunology, University of Porto; and 275 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: Ashkenazi (55), Middle Eastern (28 Iranian and 26 Iraqi), North African (34 Moroccan, 13 Libyan and 13 Tunisian), and Sephardic (62 Turkish and 44 Bulgarian).

2.2. Indel genotyping

The 38 indels were genotyped using the PCR multiplex protocol originally described in Pereira et al. [5]. Amplification products were separated by capillary electrophoresis in a 3130 Genetic Analyzer (Applied Biosystems) and fragment sizes and allele calls were determined automatically using GeneMapper v3.2 ID software (Applied Biosystems). Typing quality and allele designation were warranted by the analysis of control samples of known genotype (GHEP-ISFG collaborative exercise; details available at www.gep-isfg.org/en/working-commissions/collaborative-exercise-indels-2012.html).

2.3. Data analysis

Allele frequencies, Hardy–Weinberg equilibrium analysis, AMOVA and populations pairwise genetic distances (F_{ST}) were calculated using the Arlequin v.3.5 software [23]. Gene Diversity was calculated as $1 - \Sigma_i^m (p_i)^2$. Statistical parameters of forensic interest: MP (matching probability), PD (power of discrimination), PIC (polymorphisms information content), PE (power of exclusion) and TPI (typical paternity index) were computed using Powerstats v1.2 spreadsheet [24].

To examine the relationship of the populations under study and with other published population data [5,25,26], F_{ST} genetic distances were represented in a multidimensional scaling (MDS) plot using SPSS v.15.0 (SPSS, Inc., Chicago, IL, USA). Population structure was further assessed with STRUCTURE 2.3.4 software using default settings (admixture model; correlated allele frequencies) [27,28]. All runs included a burn-in period of 50,000 iterations followed by 100,000 MCMC repetitions ($K = 1 - 9$), and were repeated ten times each in order to test the consistency of the results.

3. Results and discussion

3.1. Intra-population variability

A total of 932 chromosomes were analyzed in this study. Genotypic data and allele frequencies for each indel marker and population (Sephardic, North African, Middle Eastern and Ashkenazi Jews, Chuetas and Jews from Bragança) are shown in Supplementary Tables S1 and S2. No deviations from Hardy–Weinberg equilibrium were observed in the populations studied after Bonferroni's correction for multiple tests ($p > 0.00132$), showing that no significant levels of substructure were found.

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.11.003.

High genetic diversities (GD) were observed in all populations analyzed (Supplementary Table S3), with mean values ranging between 0.4148–0.4336 (Table 1). The highest values were found for rs2307579 (Y05) and rs34511541 (G08) and expected heterozygosities ranged between 0.4861–0.4989 and between 0.4650–0.5000, respectively, while the lowest heterozygosities were observed for rs2307839 (Y09) in Ashkenazi (0.1503) and rs2308171 (R09) in North African Jews (0.1528). These GD values were in the range found in other studies [5,25,26], although the markers showing the highest and the lowest variability differed between populations.

Supplementary material related to this article found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.11.003.

Linkage disequilibrium (LD) test showed no significant gametic association between all 38 markers after Bonferroni correction, with the exception of the pair rs16624–rs2067238 in Ashkenazi population ($p \leq 10^{-5}$), in accordance with the reported founder effect and genetic drift of this population, indicating an early bottleneck, probably corresponding to initial migrations of ancestral Ashkenazim to Europe [29]. Still, it must be said that many generations have elapsed after the founder bottleneck and therefore this LD value may be a spurious result, and since the two indels are located at different chromosomes, this association is rapidly vanished, provided random mating is established.

Table 1
Diversity and Forensic Parameters for the set of 38 indels studied in the 6 studied populations with Jewish ancestry.

	Obs Het	Exp Het	Combined MP	Combined PD	Combined PE
Sephardic Jews	0.4322 ± 0.0768	0.4244 ± 0.0738	1.9581E + 14	0.9999999999999950	99.72%
North African Jews	0.4204 ± 0.0995	0.4190 ± 0.0831	1.5530E + 14	0.9999999999999940	99.68%
Middle Eastern Jews	0.4096 ± 0.0871	0.4148 ± 0.0743	1.2510E + 14	0.9999999999999920	99.50%
Ashkenazi Jews	0.4412 ± 0.1050	0.4187 ± 0.0799	6.3431E + 13	0.9999999999999840	99.85%
Chuetas	0.4338 ± 0.0641	0.4336 ± 0.0555	3.3360E + 14	0.9999999999999970	99.70%
Bragança Jews	0.4387 ± 0.0844	0.4335 ± 0.0574	2.0471E + 14	0.9999999999999950	99.79%

Obs Het: observed heterozygosity; Exp Het: expected heterozygosity; MP: matching probability; PD: power of discrimination; PE: power of exclusion.