FIRST-DEGREE FAMILIAL RELATIONSHIPS COINCIDENCES IN A POPULATION DATA BASE OF JUJUY (ARGENTINA) COMPARED WITH SIMULATED POPULATIONS

Juan José Martínez 1, Marcos Felipe Tula Molina 2, Luis Adrián Sánchez 2, María Inés Deus 2, María Isabel Ramella 2, Gladys Beatriz Maurín 2, Gustavo Martínez 3, Ariel Chernomoretz 4, María Soledad Escobar 5, Gustavo Sibilla 5, María Cecilia Miozzo 2*

Introduction

Jujuy is an Andean province from Northwestern Argentina. It has a population with a great input of the native South American gene pool, together with European and African contributions [1, 2]. For forensic purposes, a frequency population database was built by selecting 500 nonrelated individuals from Jujuy province and by genotyping them with 21 autosomal STRs. To corroborate the presence of repeated individuals, the genetic profiles were incorporated in GENis (http://www.fundacionsadosky.org.ar/genis), an open-source software developed in Argentina for storage

and comparison of DNA profiles. After this confirmation (no matches were found in high stringency), we performed a searching strategy with low stringency, often used to find parent-child relationships. As opposed to what was expected, four (4) adventitious matches were found (Table 2, Fig. 1). To investigate if this result was caused by a specific genetic

regiones Andinas, CONICET and Universidad Nacional de Jujuy, San

Salvador de Jujuy, Jujuy, Argentina 2 Laboratorio Regional de Genética Forense del NOA, Departamento Médico, Poder Judicial

2 Laboratorio Regional de Genética Forense del NOA, Departamento Médico, Poder Judicial de Julya, San Salvador de Julya, Julya, Argentina
3 Servicio de Genética Forense y Registro Provincial de Datos Genéticos, Superior Tribunal de Justicia de la Provincia de Estre Rios, Paraná, Entre Rios, Argentina
4 Departamento de Fisica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires - Laboratorio de Biología de Sistemas Integrativa Fundación Instituto Leloir, Buenos Aires, Argentina
Fundación Instituto Leloir, Buenos Aires, Argentina
*Corresponding author E-mail address: cmiozzo⊛justiciajuju,gov.ar

DNA from 500 buccal swab samples were extracted with DNA IQ™ Casework Pro Kit for Maxwell® 16 (Promega). Samples were genotyped for 21 autosomal STR with GlobalFiler M PCR Amplification kit (Applied Biosystems). The bias from Hardy-Weinberg equilibrium (HWE) from the Jujuy province database was estimated by means of the exact test in Arlequin 3.5 [3] and allele frequencies were estimated in GenAlEx 6.502 [4, 5]. Forensic population parameters were estimated with the software PowerStats v 1.2 [6]. One hundred (100) populations were simulated with 500 random genetic profiles each one using the observed Jujuy province allele frequencies. For this purpose, the forensim package [7] for R [8] was used and every population was analyzed with GENis with a low stringency search strategy. The results from GENis were recorded and its distribution was described. Finally, this distribution was compared with the observed value obtained in the Jujuy province database.

characteristic of the Jujuy population (e. g., endogamy), we performed a simulation-based approach to test the hypothesis that adventitious matches arose at random.

Results and Discussion

There were no deviations from HWE in the Jujuy population after Bonferroni correction (p-value 0.00238). The forensic parameters demonstrated that these 21 STR have a good performance for parentage testing and forensic casework (Table 1). Low values of likelihood ratios (LRs), for parent-child relationships, were obtained for the matches of the Jujuy population. There were inconsistencies (exclusions) in all matches when more STRs were assayed, including Y STRs in three of the four matches (Table 2). This proved that the coincidences were adventitious in the Jujuy population.

The adventitious matches for the simulated populations were on average 2.07 (mean 95% CI: 1.77-2.36), even though 4 matches were found in 12 populations, 5 matches in Bonferroni correction: 0.00238

Table 1 Hardy-Weinberg equilibrium and forensic parameters for Jujuy Population

Locus	Observed heterozygisity	Expected heterozygosity	P-value	Power of discrimination	Power of exclusion	Polymorphic information content	Typical paternity index
D3S1358	0.65200	0.65690	0.57797	0.8268	0.3580	0.6035	1.44
vWA	0.69200	0.72033	0.46858	0.8752	0.4160	0.6749	1.62
D16S539	0.81600	0.79504	0.76572	0.9232	0.6291	0.7624	2.72
CSF1PO	0.75400	0.70434	0.46608	0.8533	0.5166	0.6511	2.03
трох	0.64800	0.60895	0.01116	0.7840	0.3525	0.5552	1.42
D8S1179	0.80400	0.79873	0.79155	0.9289	0.6065	0.7704	2.55
D21S11	0.85200	0.83453	0.39702	0.9477	0.6988	0.8127	3.38
D18S51	0.82600	0.85417	0.32329	0.9612	0.6482	0.8362	2.87
D2S441	0.60000	0.59458	0.78973	0.7890	0.2909	0.5517	1.25
D19S433	0.77600	0.81052	0.00800	0.9397	0.5553	0.7864	2.23
TH01	0.63800	0.66486	0.56982	0.8355	0.3390	0.6121	1.38
FGA	0.86600	0.87040	0.81838	0.9685	0.7266	0.8559	3.73
D22S1045	0.61200	0.60951	0.25914	0.7733	0.3055	0.5315	1.29
D5S818	0.65400	0.67839	0.04594	0.8577	0.3607	0.6431	1.45
D13S317	0.80400	0.80944	0.66389	0.9372	0.6065	0.7836	2.55
D7S820	0.70400	0.72837	0.72990	0.8834	0.4345	0.6821	1.69
ACTBP2	0.90600	0.92575	0.04092	0.9875	0.8077	0.9200	5.32
D10S1248	0.66600	0.68691	0.96923	0.8531	0.3776	0.6347	1.50
D1S1656	0.85200	0.85093	0.03375	0.9591	0.6988	0.8337	3.38
D12S391	0.77000	0.79854	0.69007	0.9354	0.5446	0.7709	2.17
D2S1338	0.84000	0.84226	0.39013	0.9552	0.6753	0.8233	3.13

ower of exclusion: 0.99999995

5 and there were 2 populations with 6 matches. As there were no significant differences between Jujuy and the simulated populations (Fig. 2), the cause of adventitious matches due to a specific Jujuy population characteristic was disregarded. The LRs of 207 matches obtained from the 100 simulated populations, ranged mostly from 4 to 100,000, but there was one value above 2,000,000 (Fig. 3). It is known that the number of adventitious matches increases when you go down with stringency in DNA databases, but analyzing a high number of autosomal STRs (21), the 4 matches in the Jujuy population were unexpected. Thus, care should be taken with first-degree relationships searches, even with a high number of markers. LR calculations and complementary studies (like more autosomal markers, the Y chromosome, and mitochondrial DNA) should be performed in this kind of search.

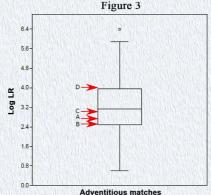
Table 2

JUJUY POPULATION MATCHES

Match	Sex of the individuals	LR	Exclusions with additional STRs
Α	♀-♂	512	1 Automomal STR
В	ð-ð	302	3 Automomal STR + Y Haplotype
C	ð - ð	905	3 Automomal STR + Y Haplotype
D	ð - ð	13.496	1 Automomal STR + Y Haplotype

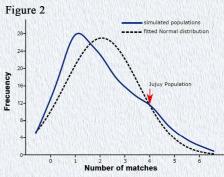
Conclusions

DNA databases are one of the most powerful tools in the investigation of crimes, linking DNA profiles from crime-related biological trace material to each other and to the possible donors. They are vital for the resolution of serial crimes. Adventitious matches could arise when forensic genetic databases became larger, when the number of loci are insufficient, and also when you go down with search stringency. This may occur with the searching strategy called "familial searching", allowed in some countries, where a search for matches with possible relatives of the donor of a crime scene associated DNA profile can be conducted [9]. Here we described a case study from a forensic database from Argentina. The next step is to shed light on the prediction of the number of adventitious coincidences in a low stringency search strategy.



Box-and-whisker plot showing the range of variation of LRs of adventitious matches from simulated populations. The box delimits the 25th and 75th quartiles, the horizontal bar indicates the median, and the whiskers extend to the most extreme data points. Red arrows indicate the values of log LR found in the Jujuy





Histogram based on the number of matches found by GENis under low stringency search from simulated populations and the number of coincidences from the Jujuy database.

References

- References

 1. S. Cardoso, L. Palencia-Madrid, L. Valverde, M.A. Alfonso-Sánchez, L. Gómez-Pérez, E. Alfaro, C.M. Bravi, J.E. Dipieri, J.A. Peña, M.M. Pancorbo, Mitochondrial DNA control region data reveal high prevalence of Native American lineages in Jujuy province, DWA Tagentina, Forensio Sci. Int. Genet. 7 (2013) e52-e55.

 2. G. Martinez, A. Borosky, D. Corach, C. Llull, L. Locarno, M. Lojo, M. Marino, M.C. Miozzo, N. Modesti, C. Pacharoni, J.P. Pilili, M.I. Ramella, A. Sala, C. Schaller, C. Vullo, U. Toscamini, Investigator® HDJlex (Qiagen) reference population database for forensic use in Argentina, Forensic Sci. Int. Genet. 26 (2017) 91-95.

 3. L. Excoffier, G. Laval, S. Schneider, Arlequin ver. 3.0: An integrated software package for population genetics data analysis, Evol. Bioinform. 1 (2005) 47-50.

 4. R. Peakall, P.E. Smouse, GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research, Mol. Ecol. Notes 6 (2006) 288-295.

 5. R. Peakall, P.E. Smouse, GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an Undate. Bioinformatics 28 (2012) 2537-2539.

 6. A. Tereba, Tools for analysis of population statistics, Profiles DNA 3 (1999) 14-16.

 7. H. Haned, Porensim: an open-source initiative for the evaluation of statistical methods in forensic genetics, Forensic Sci. Int. Genet. 5 (2011) 265-268.

 8. R Development Core Team. 2014. R: a language and environment for statistical computing.

 9. ENFSI DNA working group. 2017. DNA database management review and recommendations.