International Journal of Legal Medicine

STUDY OF INDEL GENETIC MARKERS WITH FORENSIC AND ANCESTRY INFORMATIVE INTEREST IN PALOP'S IMMIGRANT POPULATIONS IN LISBOA

--Manuscript Draft--

Manuscript Number:	IJLM-D-16-00270R2
Full Title:	STUDY OF INDEL GENETIC MARKERS WITH FORENSIC AND ANCESTRY INFORMATIVE INTEREST IN PALOP'S IMMIGRANT POPULATIONS IN LISBOA
Article Type:	Population Data
Corresponding Author:	António Amorim Instituto Nacional de Medicina Legal PORTUGAL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Instituto Nacional de Medicina Legal
Corresponding Author's Secondary Institution:	
First Author:	Ana Inácio
First Author Secondary Information:	
Order of Authors:	Ana Inácio
	Heloísa Afonso Costa
	Cláudia Vieira da Silva
	Teresa Ribeiro
	Maria João Porto
	Jorge Costa Santos
	Gilberto Igrejas
	António Amorim
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	The migratory phenomenon in Portugal has become one of the main factors for the genetic variability. In the last few years, a new class of autosomal insertion/deletion markers - InDel - has attracted interest in forensic genetics. Since there is no data for InDel markers of PALOP immigrants living in Lisboa, our aim is the characterization of those groups of individuals by typing them with at least 30 InDel markers and to compare different groups of individuals/populations. We studied 454 bloodstain samples belonging to immigrant individuals from Angola, Guinea-Bissau and Mozambique. DNA extraction was performed with the Chelex® 100 method. After extraction all samples were typed with the Investigator® DIPplex method. Through the obtained results, allelic frequencies show that all markers are at Hardy-Weinberg equilibrium and we can confirm that those populations show significant genetic distances between themselves, between them and the host Lisboa population. Because of this they introduce genetic variability in Lisboa population.
Author Comments:	Dear Editor, The Reviewer #2 comment: ready for publication. Our Response: Nothing to declare.

	The Reviewer #1 comment: Acceptable with MINOR revisions that do not require further scrutiny. Our Response: Almost all the proposed corrections were introduced. About the comment " I also feel 38 references for a simple population study ", we decide to maintain the 38 references because all of them were used to develop our study and to write this paper. Best Regards
Response to Reviewers:	Dear Editor, The Reviewer #2 comment: ready for publication. Our Response: Nothing to declare. The Reviewer #1 comment: Acceptable with MINOR revisions that do not require further scrutiny. Our Response: Almost all the proposed corrections were introduced. About the comment “ I also feel 38 references for a simple population study ”, we decide to maintain the 38 references because all of them were used to develop our study and to write this paper. Best Regards

Dear Editor,

The Reviewer #2 comment: ready for publication.

Our Response:

Nothing to declare.

The Reviewer #1 comment: Acceptable with MINOR revisions that do not require further scrutiny.

Our Response:

Almost all the proposed corrections were introduced.

About the comment “ I also feel 38 references for a simple population study ”, we decide to maintain the 38 references because all of them were used to develop our study and to write this paper.

Best Regards

STUDY OF INDEL GENETIC MARKERS WITH FORENSIC AND ANCESTRY INFORMATIVE INTEREST IN PALOP'S IMMIGRANT POPULATIONS IN LISBOA

Ana Inácio[1,2], Heloísa Afonso Costa[1,6], Cláudia Vieira da Silva[1,6], Teresa Ribeiro[1,6], Maria João Porto[1,6], Jorge Costa Santos[3,4,6],Gilberto Igrejas[2], António Amorim[1,4,5,6]

1 - Instituto Nacional de Medicina Legal e Ciências Forenses, Portugal

2 - Departamento de Genética e Biotecnologia, Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal

3 - Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal

4 - Instituto Superior de Ciências da Saúde Egas Moniz, Monte de Caparica, Portugal

5 - Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

6 - CENCIFOR - Centro de Ciências Forenses, Coimbra, Portugal

Introduction

Since the early 1970's, in Portugal, the flow of immigrants from African countries has increased. We can relate this reality with the post-colonial and post-independence period of Portuguese-speaking African countries.

According to the Portugal Contemporary Base - PORDATA -, by the end of 2014, the total number of immigrants from PALOP (Portuguese-speaking African countries) in Portugal was about 91 000, and from those, 75 000 are part of Lisboa population.

The number of immigrants in Portugal is relevant and the migratory phenomenon in Portugal, particularly in Lisboa, can become one of the main factors for the genetic variability [1-7].

In the last few years a new class of autosomal insertion/deletion markers, - InDel -, has attracted interest in forensic genetics, mainly by their abundance in the human genome and by the simple analysis methodologies associated to their study [8]. They are characterized by the presence or absence of a specific sequence of nucleotides [9-12]. Significative differences in allele frequencies of InDel markers, between different groups or populations, can be used as ancestry and eventually, evolutionary indicators [13].

InDels represent approximately 15.6% of all the polymorphisms in human genome, suggesting that in human population there is a minimum of 1.56 million InDels [14,15]. Furthermore, InDels combine characteristics of STRs and SNPs, filling some gaps of each one of these markers and introducing a new class of genetic markers for forensic purposes. Advances in forensic genetics are noticed over the past few years and show us the interest for the development of biallelic markers, such as SNPs and InDels [9,16].

Until today only one commercial kit is available for the study of this kind of polymorphisms - Investigator® DIPplex - developed by Qiagen. This kit allows the study of 30 InDel markers and the homologous gene of amelogenin as informative of the individual's sexual gender [12,16,17]. This multiplex has the advantage that amplified products can be separated and analyzed by capillary electrophoresis using an automatic sequencer, with similar technology to STR analysis.

The Investigator® DIPplex kit (Qiagen, Hilden, Germany) has already been employed in some population studies, specially in populations of European, American and Asian origin [8-10,12,16-24]. Demonstrating high discrimination power and high power of exclusion, has shown to be proper for forensic cases, specially, in individual identification that require analysis of degraded samples. In relation to kinship investigations, this kit can only be a complement to the STR analysis.

Since there is a reduced amount of data for InDel markers of PALOP immigrants living in Lisboa, our aim is to characterize those groups of individuals by typing them with 30 markers and compare different groups of individuals/populations.

Material and methods

454 bloodstain samples were studied, 258 from Angola, 124 from Guinea-Bissau and 72 from Mozambique. These bloodstain samples were collected from immigrant individuals, inhabitants of Lisboa metropolitan area, undergoing forensic investigations in Instituto Nacional de Medicina Legal e Ciências Forenses (INMLCF). The number of studied individuals from each of the different African countries - Angola, Guinea-Bissau, Mozambique - represent each immigrant group within the Lisboa population.

An interview was conducted in order to register personal data of the studied individuals, particularly the name, the age, the birthplace, the individual and the parental ethnicity. According to Portuguese legal regulations, samples from routine forensic cases ongoing at INMLCF can be used for investigation purposes, which, naturally

include genetic studies. In our study all samples are used with special codification without any connection to personal or judicial data related to the donor.

DNA extraction was carried out using the Chelex®100 resin extraction method [25]. InDel typing was accomplished with Investigator® DIPplex PCR Amplification Kit (Qiagen, Hilden, Germany). DNA fragments separation, detection and identification was achieved with capillary electrophoresis using an ABI PRISM® Genetic Analyser 3130 xl sequencer (Applied Biosystems, Foster City, USA).

Arlequin software ver.3.5 [26] was used for the calculations of allele frequencies for each locus, as well as expected and observed heterozigosities. Beyond this, Arlequin software ver.3.5 [26] was also used to calculate Hardy-Weinberg equilibrium (HWE) and to estimate *p*-values, which was considered to be significant at $p < 0.0001$, after Bonferroni correction [27]. Concerning forensic parameters, power of exclusion (PE), discrimination power (DP), polymorphic information content (PIC), typical paternity index (TPI) and matching probability (MP) were calculated using PowerStats software, ver.12 (Promega, UK).

Phylogenetic comparison between the three populations, - Angola immigrants, Guinea-Bissau immigrants and Mozambique immigrants - , as well as others, South of Portugal [28] and Cape Verde immigrants [29] were accomplished with Arlequin software ver.3.5 and with PHYLIP - Phylogeny Inference Package ver.3.2 [30].

Results

A sample of 454 immigrant individuals from different populations, Angola, Guinea-Bissau and Mozambique was studied aiming to characterize them with InDel genetic markers for future application in forensic casework samples. Supplementary tables S1, S2 and S3 in ESM 1 show the frequencies for insertions and deletion alleles for all studied populations/groups. All studied InDel followed Hardy-Weinberg expectations ($p < 0.0001$) except the *locus* HLD97 from Guinea-Bissau immigrant population.

Expected heterozigosity (He), Observed heterozigosity (Ho), PE, DP, PIC, TPI and MP values, for all populations/groups are also presented in supplementary tables S1, S2 and S3 (ESM 1). The combined matching probability (CMP), the combined power of discrimination (CPD) and the combined power of exclusion (CPE) are presented in supplementary table S4 (ESM 1). The achieved values for forensic parameters allow a satisfactory level of discrimination in forensic cases.

By the analysis of the obtained electropherograms, we detected microvariants in loci HLD92, HLD99 and HLD84 which influenced the interpretation of the genotypes (see ESM 2, Figures S1, S2, S3). The detection and confirmation of microvariants particularly in loci HLD92, HLD99 and HLD84 is an advantage of using this panel of InDel. It is important to note that microvariant of HLD92 locus is not described in the literature.

Discussion and conclusion

Through the allelic frequencies obtained for the three immigrant populations in the study, it was possible to verify the existence of genetic differences between them. Figure 1 presents the phylogenetic tree representative of those genetic distances, presented in supplementary table S5 (ESM 1).

The populations of Angola, Guinea-Bissau and Mozambique, before the major migration of Bantu, was inhabited by regional tribes and by Khoisan tribes [21,22]. Thus, the great evolutionary phenomenon of the African population was due to the expansion of Bantu tribes along two main different regions - East, where Mozambique is included, and West of Africa where Angola and Guinea-Bissau are included [23,24].

Concerning Cape Verde, it was discovered uninhabited and was later colonized mainly by people of the African West coast due to the slave trade practiced by the Portuguese [25,26]. So it is expected that these three immigrants populations, - Angola, Guinea-Bissau and Cape Verde -, are closer and Mozambique immigrants has a greater genetic distance in relation to the other three African populations [37].

The results show us those immigrants populations of Angola, Guinea-Bissau and Cape Verde are more approximated between them and much more separated of the Portuguese population. Mozambique immigrants are closer with Portuguese population than the other ones, and can participate in an individual cluster. Angola immigrants, Guinea-Bissau immigrants and Cape Verde immigrants are included in the other cluster, and as expected taking into account the expansion of Bantu tribes along two main different regions as mentioned before.

Through the obtained results it’s possible to confirm that the studied African populations show significant genetic distances between them, between them and the Lisboa population and so we can conclude that they introduce genetic variability in the Lisboa population.

Furthermore, the Investigator DIPplex® PCR amplification kit (Qiagen, Hilden, Germany), due to it's characteristics such as short amplicon size, absence of stutters and simplicity to implement in forensic genetic laboratories, can be very interesting as a supplement of any study of STR in forensic caseworks with the studied Portuguese and African populations.

This paper follows the guidelines for publication data request by the journal [38].

References

[1] P. Morais, A. Amorim, C. Vieira da Silva, T. Ribeiro, J. Costa Santos, H. Afonso Costa, Genetic portrait of Lisboa immigrant population from Cabo Verde with mitochondrial DNA analysis, J. Genet. (2015).

[2] F. Simão, H. Afonso Costa, C. Vieira da Silva, T. Ribeiro, M.J. Porto, J. Costa Santos, A. Amorim, Genetic portrait of Lisboa immigrant population from Angola with mitochondrial DNA, Forensic Sci. Int. Genet. 15 (2014) 33–38. doi:10.1016/j.fsigen.2014.09.013.

[3] H. Afonso Costa, P. Morais, C. Vieira da Silva, S. Matos, R. Marques Santos, R. Espinheira, J. Costa Santos, A. Amorim, X-Chromosome STR markers data in a Cabo Verde immigrant population of Lisboa, Mol. Biol. Rep. 41 (2014) 2559–2569. doi:10.1007/s11033-014-3114-9.

[4] R. Marques dos Santos, C. Vieira da Silva, H. Afonso Costa, P. Ferreira Gomes, S. Sanches, R. Espinheira, J. Costa Santos, A. Amorim, Genetic portrait of an immigrant population from Angola living in Lisboa, Forensic Sci. Int. Genet. 6 (2012) 170–173. doi:10.1016/j.fsigen.2012.05.010.

[5] A. Amorim, R. Marques dos Santos, C. Vieira da Silva, H. Afonso Costa, R. Espinheira, P. Ferreira Gomes, J. Costa Santos, Genetic portrait of a native population of Cabo Verde living in Lisboa, Forensic Sci. Int. Genet. 6 (2012) 166–169. doi:10.1016/j.fsigen.2012.05.006.

[6] A. Amorim, R. Marques dos Santos, C. Vieira da Silva, H. Afonso Costa, R. Espinheira, J. Costa Santos, Genetic portrait of Brazilian immigrant population living in Lisboa, Forensic Sci. Int. Genet. 6 (2012) 121–124. doi:10.1016/j.fsigen.2012.01.008.

[7] A. Amorim, C. Vieira da Silva, H. Afonso Costa, S. Sanches, R. Espinheira, J.

Costa Santos, Allele frequencies of all CODIS and four non-CODIS STR loci of an immigrant Brazilian population living in Lisbon – Preliminary results, Forensic Sci. Int. Genet. Suppl. Ser. 3 (2011) 127–128. doi:10.1016/j.fsigss.2011.08.063.

[8] A. Carvalho, M.F. Pinheiro, Population data of 30 insertion/delection polymorphisms from a sample taken in the North of Portugal, Int. J. Legal Med. 127 (2013) 65–67. doi:10.1007/s00414-012-0693-7.

[9] H.-T. Meng, Y.-D. Zhang, C.-M. Shen, G.-L. Yuan, C.-H. Yang, R. Jin, J.-W. Yan, H.-D. Wang, W.-J. Liu, H. Jing, B.-F. Zhu, Genetic polymorphism analyses of 30 InDels in Chinese Xibe ethnic group and its population genetic differentiations with other groups, Sci. Rep. 5 (2015) 8260. doi:10.1038/srep08260.

[10] B.L. LaRue, R. Lagacé, C.W. Chang, A. Holt, L. Hennessy, J. Ge, J.L. King, R. Chakraborty, B. Budowle, Characterization of 114 insertion/deletion (INDEL) polymorphisms, and selection for a global INDEL panel for human identification, Leg. Med. 16 (2014) 26–32. doi:10.1016/j.legalmed.2013.10.006.

[11] A. Zidkova, A. Horinek, V. Kebrdlova, M. Korabecna, Application of the new insertion-deletion polymorphism kit for forensic identification and parentage testing on the Czech population, Int. J. Legal Med. 127 (2013) 7–10. doi:10.1007/s00414-011-0649-3.

[12] W. Pepinski, M. Abreu-Glowacka, M. Koralewska-Kordel, E. Michalak, K. Kordel, A. Niemcunowicz-Janica, M. Szeremeta, M. Konarzewska, Population genetics of 30 INDELs in populations of Poland and Taiwan, Mol. Biol. Rep. 40 (2013) 4333–4338. doi:10.1007/s11033-013-2521-7.

[13] R. Pereira, C. Phillips, N. Pinto, C. Santos, S.E.B. dos Santos, A. Amorim, Á. Carracedo, L. Gusmão, Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing, PLoS One. 7 (2012) 1–10. doi:10.1371/journal.pone.0029684.

[14] R. Klein, C. Neumann, R. Roy, Detection of insertion/deletion polymorphisms from challenged samples using the Investigator DIPplex® Kit, Forensic Sci. Int. Genet. 16 (2015) 29–37. doi:10.1016/j.fsigen.2014.11.020.

[15] R.E. Mills, C.T. Luttig, C.E. Larkins, A. Beauchamp, C. Tsui, W.S. Pittard, S.E.

Devine, An initial map of insertion and deletion (INDEL) variation in the human genome, Genome Res. 16 (2006) 1182–1190. doi:10.1101/gr.4565806.

[16] T.D.J.B. Ferreira Palha, E.M. Ribeiro Rodrigues, G.C. Cavalcante, A. Marrero, I.R. de Souza, C.J. Seki Uehara, C.H.A. Silveira da Motta, D. Koshikene, D.A. da Silva, E.F. de Carvalho, G. Chemale, J.M. Freitas, L. Alexandre, R.T.F. Paranaiba, M.P. Soler, S. Santos, Population genetic analysis of insertion–deletion polymorphisms in a Brazilian population using the Investigator DIPplex kit, Forensic Sci. Int. Genet. 19 (2015) 10–14. doi:10.1016/j.fsigen.2015.03.015.

[17] K.M. Seong, J.H. Park, Y.S. Hyun, P.W. Kang, D.H. Choi, M.S. Han, K.W. Park, K.W. Chung, Population genetics of insertion-deletion polymorphisms in South Koreans using Investigator DIPplex kit, Forensic Sci. Int. Genet. 8 (2014) 80–83. doi:10.1016/j.fsigen.2013.06.013.

[18] B.L. LaRue, J. Ge, J.L. King, B. Budowle, A validation study of the Qiagen Investigator DIPplex® kit; An INDEL-based assay for human identification, Int. J. Legal Med. 126 (2012) 533–540. doi:10.1007/s00414-012-0667-9.

[19] C.T. Li, S.H. Zhang, S.M. Zhao, Genetic analysis of 30 InDel markers for forensic use in five different Chinese populations., Genet. Mol. Res. 10 (2011) 964–979. doi:10.4238/vol10-2gmr1082.

[20] A.M. Neuvonen, J.U. Palo, M. Hedman, A. Sajantila, Discrimination power of Investigator DIPplex loci in Finnish and Somali populations, Forensic Sci. Int. Genet. 6 (2012) 99–102. doi:10.1016/j.fsigen.2011.09.005.

[21] Y.L. Wei, C.J. Qin, H. Dong, J. Jia, C.X. Li, A validation study of a multiplex INDEL assay for forensic use in four Chinese populations, Forensic Sci. Int. Genet. 9 (2014) e22–e25. doi:10.1016/j.fsigen.2013.09.002.

[22] S.L. Friis, C. Børsting, E. Rockenbauer, L. Poulsen, S.F. Fredslund, C. Tomas, N. Morling, Typing of 30 insertion/deletions in Danes using the first commercial indel kit - Mentype® DIPplex, Forensic Sci. Int. Genet. 6 (2012) 72–74. doi:10.1016/j.fsigen.2011.08.002.

[23] Z. Kis, A. Zalán, A. Volgyi, Z. Kozma, L. Domján, H. Pamjav, Genome deletion

and insertion polymorphisms (DIPs) in the Hungarian population, Forensic Sci. Int. Genet. 6 (2012) 125–126.

[24] C. Tomas, I. Skitsa, E. Steinmeier, L. Poulsen, a. Ampati, C. Børsting, N. Morling, Results for five sets of forensic genetic markers studied in a Greek population sample, Forensic Sci. Int. Genet. 16 (2015) 132–137. doi:10.1016/j.fsigen.2015.01.001.

[25] P. Walsh, D. Metzger, R. Higuchi, Chelex 100 as a médium for simple extraction of DNA for PCR-based typing from forensic material, Biotechniques. 10 (1991) 506–513.

[26] L. Excoffier, G. Laval, S. Schneider, Arlequin (version 3.0): An integrated software package for population genetics data analysis, Evol. Bioinform. Online. 1 (2005) 47–50.

[27] R.A. Armstrong, When to use the bonferroni correction, Ophthalmic Physiol. Opt. 34 (2014) 502–508. doi:10.1111/opo.12131.

[28] C. Vieira da Silva, S. Matos, H. Afonso Costa, P. Morais, R.M. Dos Santos, R. Espinheira, J. Costa Santos, A. Amorim, Genetic portrait of south Portugal population with InDel markers, Forensic Sci. Int. Genet. 7 (2013) 10–12. doi:10.1016/j.fsigen.2013.03.009.

[29] R.A. Reis, InDel e SNP: Caracterização Genética das Populações Imigrantes de Cabo Verde e Guiné-Bissau a residir na Região de Lisboa, Escola Superior de Saúde Egas Moniz, 2014.

[30] J. Felsenstein, PHYLIP - Phylogeny Inference Package (Version 3.2), Cladistics. 5 (1989) 164–166.

[31] J. Redinha, Etnias e Culturas de Angola, Associação das Universidades de Língua Portuguesa, 2009.

[32] G. Firmino, A Situação Do Português no Contexto Multilingue De Moçambique, Univ. Eduardo Mondlane. (2005).

[33] G. Berniell-Lee, F. Calafell, E. Bosch, E. Heyer, L. Sica, P. Mouguiama-daouda, L. Van Der Veen, J. Hombert, L. Quintana-Murci, D. Comas, Genetic and

Demographic Implications of the Bantu Expansion : Insights from Human Paternal Lineages, Mol. Biol. Evol. 26(7)1581–1589. 26 (2008) 1581–1589. doi:10.1093/molbev/msp069.

[34] S. Beleza, L. Gusmão, A. Amorim, A. Carracedo, A. Salas, The genetic legacy of western Bantu migrations, Hum. Genet. 117 (2005) 366–375. doi:10.1007/s00439-005-1290-3.

[35] A. Carreira, Migrações nas Ilhas de Cabo Verde, 2ª ed., Instituto Caboverdeano do Livro, 1983.

[36] A. Brehm, L. Pereira, H.-J. Bandelt, M.J. Prata, A. Amorim, Mitochondrial portrait of the Cabo Verde archipelago: the Senegambian outpost of Atlantic slave trade., Ann. Hum. Genet. 66 (2002) 49–60. doi:doi:10.1017/S0003480001001002.

[37] L. Pereira, L. Gusmão, C. Alves, A. Amorim, M.J. Prata, Bantu and European Y-lineages in Sub-Saharan Africa, Ann. Hum. Genet. 66 (2002) 369–378.

[38] A. Carracedo, J.M. Butler, L. Gusmão, A. Linacre, W. Parson, L. Roewer, P.M. Schneider, Update of guidelines for the publication of genetic population data, Forensic Sci. Int. Genet. (2014) A1–A2.

Fig. 1: Phylogenetic tree representing the genetic distances among populations under study.

Figure

Click here to download Figure Figure 1.jpg