Revised Manuscript - Characterization of GlobalFiler loci in
Angolan and Guinean populations inhabiting Southern Portugal
Click here to view linked References

Characterization of GlobalFiler loci in Angolan and Guinean populations inhabiting Southern Portugal

Abstract

We analyzed the GlobalFiler STR loci for 152 and 70 unrelated individuals from Angolan and Guinean immigrant populations inhabiting Southern Portugal, respectively. After Bonferroni correction, no significant deviations from the Hardy–Weinberg equilibrium and linkage disequilibrium were detected for either population. For the Angolan population, SE33 was the most informative marker. In contrast, D5S818 and D13S317 were the least informative loci. The combined power of discrimination was 99.99999999999999999999961907%. For the Guinean population, SE33 but also D21S1 were the most informative loci, while D13S317 was the least. The combined power of discrimination was 99.99999999999999999999997915%. No significant differences were observed between Angolan, Guinean and Afro-American populations for any of the analyzed STRs. The South African population presented significant differences at D22S1045 and D10S1248 when compared to Angola, and at D22S1045 when compared to Guinea-Bissau. The MDS plot and neighbor-joining tree analysis revealed that Angolan and Guinean populations are genetically close to African-American and South African populations, and genetically different from Korean, Mexican, European (including American-Caucasian) and Middle Eastern populations.

1

Angola and Guinea-Bissau are part of the Portuguese-speaking African countries (PALOP) and are located in West Africa. The inhabitants of these countries have been under the influence of numerous migratory movements, trade networks and consecutive invasions over the years that have resulted in the mix of people from different origins we see today. The main ethnic groups were already settled in the 15th century when the Portuguese arrived to these regions bringing cultural, religious and genetic variation to the indigenous populations [1-2].

However, during the period of post-colonization and independence, Portugal ceased to be a source of net emigration to West Africa and instead became the destination for several thousands of migrants, which may, in turn have contributed to the genetic variability of the native population [3]. According to the 2015 Immigration, Frontiers and Asylum Report, from the Portuguese Foreign Affairs Service, the total number of foreign residents was 388 731. Angolan and Guinean nationals are among the most represented in Portugal with 18 247 and 17 091 immigrants respectively, the majority of whom are located in the southern region [4].

The aims of this study, were to characterize the genotypes of Angolan and Guinean immigrant populations inhabiting Southern Portugal and estimate the statistical parameters with forensic and population relevance along with comparison to other populations. Furthermore, it is the intention of the authors to investigate the genetic contribution of these immigrant populations on native Portuguese residents.

In order to satisfy the aforementioned goals, 222 bloodstain samples from unrelated Angolan (152) and Guinean (70) individuals were studied. These were obtained during paternity testing casework from 2002 to 2013 at the Portuguese National Institute of Legal Medicine and Forensic Sciences (INMLCF), and sample collection was performed under informed consent. DNA was extracted using the Chelex® 100 method [5], and PCR amplification was performed in GeneAmp® PCR system 9700 (Applied Biosystems) using GlobalFiler™ PCR Amplification Kit (Applied Biosystems). This kit includes 21 autosomal STR loci – 20 expanded CODIS loci (TPOX, D3S1358, CSF1PO, D5S818, FGA, D8S1179, D7S820, TH01, vWA, D13S317, D16S539, D18S51 and D21S11, D19S433, D2S441, D22S1045, D10S1248, D1S1656, D12S391, D2S1338) and SE33 locus – and three gender determination markers (Amelogenin, Yindel, and

DYS391) [6]. Amplified products were separated and detected by capillary electrophoresis in an Applied Biosystems® Genetic Analyzer 3130xl and allele calling was conducted using GeneMapper® ID-X 1.4 analysis software (Applied Biosystems).

Allele frequencies, observed heterozygosity (Ho), expected heterozygosity (He), Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were calculated using Arlequin v3.5 [7]. Match probability (MP), discrimination power (PD), polymorphic information content (PIC), power of exclusion (PE) and typical paternity index (TPI) were estimated using PowerStats v1.2 [8]. The exact tests of population differentiation and genetic distances were performed between Angolan and Guinean immigrants compared with other populations previously reported, such as South Portugal, Spain , Dutch , USA (Caucasian and African individuals) , South Africa, Lebanon , United Arab Emirates , Afghanistan , China , Mexico and Korea using Arlequin v3.5 [7] and PHYLIP v3.695 [9]. Results were analyzed by multidimensional scaling plot (MDS) using SPSS v20 [10] and a neighbor-joining tree was constructed using Mega 6 software [11].

For both populations the exact test of HWE within loci and LD between all pairs of STR loci showed no significant deviations after Bonferroni correction ($\rho = (0.05/21) = 0.002381$ for HWE and $\rho = (0.05/210) = 0.0002381$ for LD) (see Table S1 and S2).

Allele frequencies and forensic statistical parameters of 21 autosomal STRs can be observed in Table S1. The most informative marker for the immigrant Angolan population is SE33 which presented the highest values for Ho, PD, PIC, PE and TPI, and the lowest result for MP. In contrast, D5S818 revealed the lowest and highest values for power of discrimination and matching probability respectively, while D13S317 exhibited the lowest results for expected heterozygosity, polymorphic information content, power exclusion and typical paternity index. SE33 and D21S1 are the most informative loci for the Guinean population inhabiting Southern Portugal. SE33 presented the highest results for PD and PIC and the lowest for PM, and D21S11 displayed the highest observed heterozygosity, power of exclusion and typical paternity index. Genetic marker D13S317, which revealed the lowest values for Ho, PD, PIC, PE and TIP and the highest result for PM, is the least informative locus.

Overall match probability for the Angolan and Guinean populations was $3,8093 \times 10^{-26}$ and $2,085 \times 10^{-25}$, respectively. Combined power of discrimination and combined power of exclusion were 0.999999999999999999999999961907 and 0.999999999837329, for the Angolan immigrant population, and 0.999999999999999999999997915 and 0.999999997871199 for the Guinean immigrant population, respectively.

After Bonferroni correction ($\rho = 0.05/273 = 0.00018$), exact test of population differentiation shows no significant differences between allele frequencies of Angolan and Guinea-Bissauan populations and an African-American one. However, comparisons between Angola/South Africa and Guinea-Bissau/South Africa revealed ρ values below 0.00018 for D22S1045, D10S1248 and for D22S1045, respectively. Regarding the Angolan population, Korean, Mexican, Chinese and Dutch populations displayed meaningful differences in 18 out of 21 analyzed loci. Korean, Mexican and Dutch also exhibited considerable differences in 13, 11 and 9 loci, respectively, when compared to the Guinea-Bissauan population (see details in Table S3).

Genetic distances were estimated based on allele frequencies of 21 STR loci between the aforementioned populations. The results, presented in Table S4 and displayed in MDS plot (Fig. 1) and neighbor-joining tree (Fig. 2), show that Angola and Guinea-Bissau are genetically close to Afro-American and South African populations, and genetically different from Korean, Mexican, European (including American-Caucasian) and Middle Eastern populations.

The obtained results provide a helpful tool for forensic purposes involving individuals from the studied populations.

The laboratory is ISO/IEC 17025:2005 accredited and participates in the collaborative quality control and proficiency testing exercises of the Spanish and Portuguese Working Group (GHEP) from ISFG.

References

[1] Alvarez, M., Piedade, J., Balseiro, S., Ribas, G., & Regateiro, F. (2009). HLA-G 3'-UTR SNP and 14-bp deletion polymorphisms in Portuguese and Guinea-Bissau populations. Int J Immunogenet, 36(6): 361–366. DOI: 10.1111/j.1744-313X.2009.00875.x

[2] Santos, R. M., Silva, C. V., Costa, H. A., Gomes, P. F., Sanches, S., Espinheira, R., Santos, J. C. & Amorim, A. (2012). Genetic portrait of an immigrant population from Angola living in Lisboa. Forensic Sci Int-Gen, 6(6): e170–e173. DOI: 10.1016/j.fsigen.2012.05.010

[3] Pinheiro, T. (2008). Emigration, Immigration and Interculturality: The Meaning of the European Year of Intercultural Dialogue in Portugal. Eurolimes, 6: 63–73. ISSN: 18419259

[4] Dias, P., Machado, R. & Oliveira, A. (2015). Relatório de Imigração, Fronteiras e Asilo. Oeiras, Portugal: Serviço de Estrangeiros e Fronteiras.

[5] Walsh, P. S., Metzger, D. A., & Higuchi, R. (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques, 10(4): 506–13. PMID: 1867860

[6] Hares, D. R. (2015). Selection and Implementation of Expanded CODIS Core Loci in the United States. Forensic Sci Int-Gen., 17: 33-34. DOI: 10.1016/j.fsigen.2015.03.006

[7] Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour, 10(3): 564-567. DOI: 10.1111/j.1755-0998.2010.02847.x

[8] Tereba, A. (1999). Tools for analysis of population statistics, Profiles in DNA, 9: 14–16.

[9] Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package) version 3.6. Seattle, USA: Department of Genome Sciences, University of Washington

[10] SPSS, IBM (2011). IBM SPSS statistics for Windows, version 20.0. New York, EUA: IBM Corp.

[11] Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA 6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol, 30: 2725-2729

Figure Legends

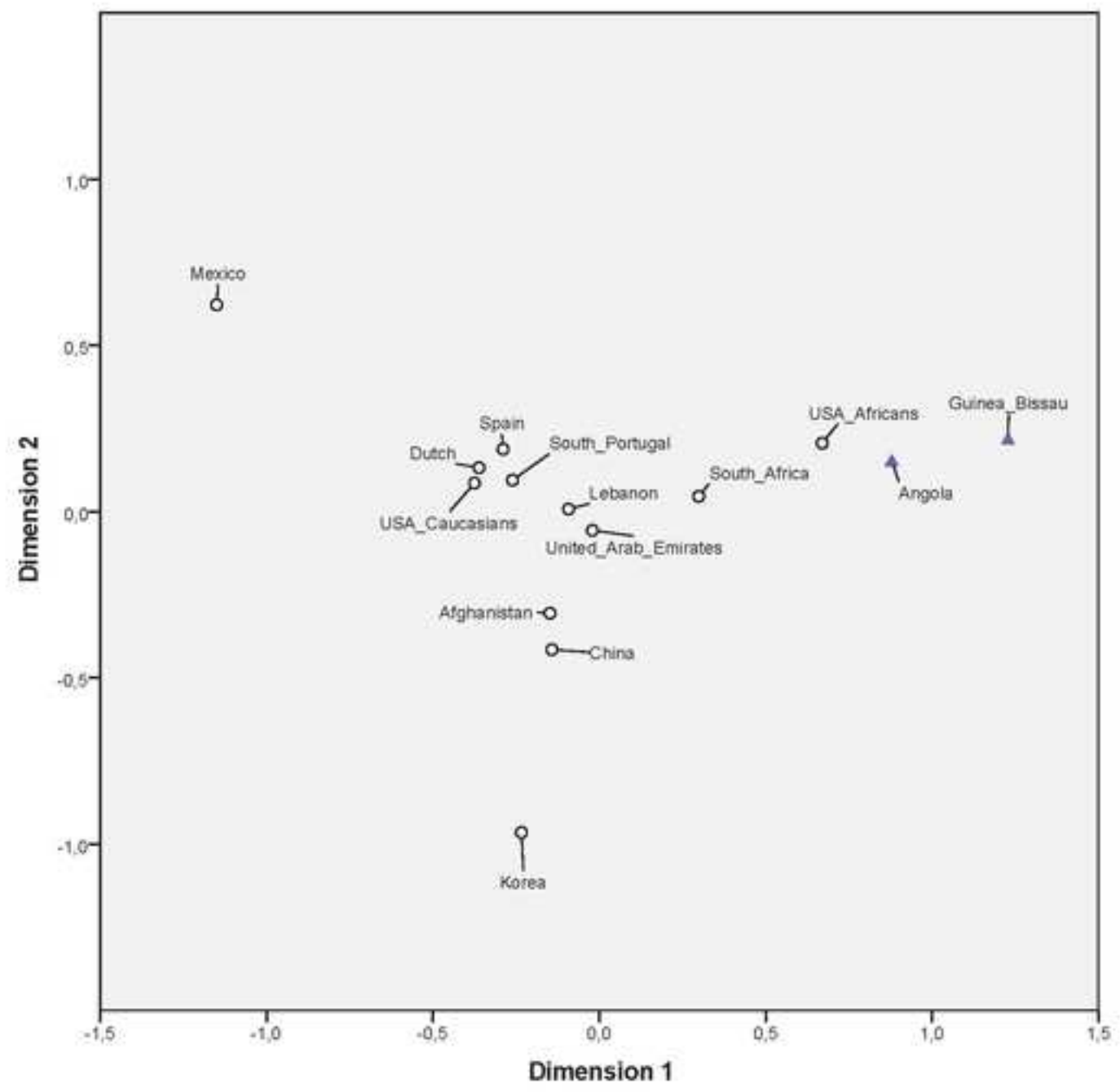**Fig. 1** Multidimensional scaling (MDS) plot based on genetic distance values between Angola and Guinea-Bissau and other populations.

Fig. 2 Neighbor-joining tree constructed with genetic distance values between Angola and Guinea-Bissau and other populations.