Article

Population genetics data for 21 autosomal STR loci for United Arab Emirates (UAE) population using next generation multiplex STR kit

Ali Alhmoudi, Osamah, Jones, Rebecca J, Tay, Guan K, Alsafar, Habiba and Hadi, Ss

Available at http://clok.uclan.ac.uk/13862/

It is advisable to refer to the publisher's version if you intend to cite from the work.
http://dx.doi.org/10.1016/j.fsigen.2015.07.009

For more information about UCLan's research in this area go to http://www.uclan.ac.uk/researchgroups/ and search for <name of research Group>.

For information about Research generally at UCLan please go to http://www.uclan.ac.uk/research/

CLoK
Central Lancashire online Knowledge
www.clok.uclan.ac.uk

1 **Forensic Population Genetics-Letter to Editor**

2 Dear JM,

3 DNA samples were analysed from 519 healthy, unrelated and consenting individuals who

4 reside in the United Arab Emirates (UAE) and were randomly chosen for this study. The

5 UAE is one of the middle-eastern countries located on the Arabian Gulf. It shares a border

6 with Iran, Saudi Arabia and Oman. The UAE was founded in 1971, and consists of seven

7 Emirates: Abu Dhabi, Dubai, Sharjah, Ajman, Ra's Al-Khaymah, Al-Fujairah and Umm Al-

8 Quwain [1]. According to the National Bureau of Statistics, (2012), the total UAE

9 population was reported to be around 8.26 million in 2010 [2]. The statistics showed that

10 some 11.5% of the total population comprised of native Arabs, with majority of the

11 population being of Indian and Pakistani ethnicities. In the early part of the twentieth

12 century, the different Arabic tribes migrated in different directions in search of suitable

13 locations to colonize. Some moved into coastal regions, while others inhabited the desert.

14 Despite the modernization throughout the union, the basic family structure and pattern of

15 native UAE Arab population has remained unchanged. Culturally, the preference for

16 consanguineous marriages remains embedded in the society [3]. However, as the

17 awareness of the social and medical impact of consanguinity increases and with

18 diversification, non-consanguineous marriages appear to be on the increase, which has

19 possibly resulted in greater genetic diversity throughout the population [4, 5]. The

20 increase in genetic diversity in the population is of interest to assess whether STR

21 markers can be used for forensic and paternity purposes. This study expands on previous

22 publications with regards to the analysis of UAE populations with the amplification of

23 additional STR markers and a larger population sample size [6].

24 The DNA samples analysed in the current study were obtained from indigenous UAE

25 nationals residing in Abu Dhabi, UAE in accordance with approval from the Ethics

26 committee of the Ministry of Health of the United Arab Emirates (2011). Informed consent

27 was received from every volunteer during this collection process and de-identified data is

28 presented. This study was also approved by the Ethics committee of the University of

29 Central Lancashire (2014) as it was carried out as part of Masters Project in DNA profiling.

30 The DNA samples provided for this study were collected and extracted using the

31 Genotek's Oragene-DNA kit (Genotek, Ottawa, Canada) in accordance with

32 manufacturer's guidelines. The quantities of extracted DNA samples were determined

33 using a NanoDrop spectrophotometer (Thermo Scientific, Wilmington DE, USA).

34 Using half volume (7.5 µl) reactions, samples were amplified using the GlobalFiler®

35 PCR amplification kit (Life Technologies, Foster City CA, USA) and alleles were called

36 using the allelic ladder provided by the manufacturer. The PCR was performed in the

37 GeneAmp® PCR System 9700 (Life Technologies). The GlobalFiler® PCR amplification

38 kit (Life Technologies) amplifies 21 autosomal STR loci, a Y-STR locus DYS 391, a Y-

39 indel marker and Amelogenin. The 21 autosomal STR loci within this amplification kit

40 were of interest for the purposes of this study. The 21 autosomal loci amplified and

41 focused on within this study were D3S1358, vWA, D16S539, CSF1PO, TPOX,

42 D8S1179, D21S11, D18S51, D2S441, D19S433, TH01, FGA, D22S1045, D5S818,

43 D13S317, D7S820, SE33, D10S1248, D1S1656, D12S391, D2S1338.

44 The PCR products   were analysed using an 8 capillary ABI 3500 DNA Genetic

45 Analyser with POP-4™ polymer (Life Technologies). GeneMapper® Software version

46 4.0 (Life Technologies) was then used for analysis.   LIZ-600 was used as internal

47 Standard (Life Technologies).

48 The alleles from all loci reported here were designated using the allelic ladder supplied

49 by the manufacturer, according to the published nomenclatures and the guidelines of

50 the International Society for Forensic Genetics (ISFG) for performing STR analyses [7].

51 The STR allele frequencies along with the parameters of population genetics: observed

52 and expected heterozygosity (Ho and He, respectively), power of discrimination (PD),

53 probability of exclusion (PE), and polymorphic information content (PIC) were estimated

54 using PowerStats version 1.2 (Promega, Madison, USA) (Supplementary Table 1).

55 Version 3.11 of the Arlequin software was used to perform an exact test to investigate

56 any departures from the Hardy-Weinberg equilibrium (HWE) [8]. The theoretical profile

57 frequency range was estimated signifying the rarest and most common heterozygous

58 genotypes. Furthermore, the number of possible genotypes was also calculated

59 (Supplementary Table 2).

60   The data generated from this study was compared to 5 published population data sets

61   for available loci [9]. Exact test comparisons were made between this current study of

62   the UAE population in this study and data from Kuwaiti, India, Saudi Arabia, Egypt, and

63   Iran.

64   Data is available upon request from shadi@uclan.ac.uk

65   Through the analysis of allele frequency data, allele 8 of TPOX was found to exhibit the

66   highest allele frequency with 49.4% in the total samples analysed for the population.

67   During analysis, two off ladder allelic variants were observed at locus SE33. These

68   variants were allele 7.3 (within 3 samples) and allele 17.3 (within 1 sample). Both of these

69   variants have been previously reported on STRBase [9]. A tri-allelic pattern (allele 6, 8,

70   10) was observed for TPOX during analysis which has also been previously reported on

71   STRBase [10]. The SE33 locus showed the largest number of different alleles (50 alleles)

72   and D13S1358, D16S539 and CSF1PO loci showed the smallest number of different

73   alleles (8 alleles). The heterozygosity (Ho) of the 21 autosomal STR loci ranged from 65%

74   (TPOX) to 92% (SE33). The power of discrimination values (PD) for all tested loci was

75   above 85%; the highest observed at SE33 with 99.3% and the least at TPOX with 85%.

76   The combined probability of exclusion (CPE), power of discrimination (CPD) and

77   matching probability (CMP) for all 21 STR loci were 0.99999999999999999999999994,

78   99.999×10^{-2} and 6.2468×10^{-27} respectively. When HWE was tested, there was no

79   statistical significance observed for 19 out of 21 autosomal STR loci. Bonferroni correction

80   was applied to the two loci (D8S1179 and D22S1045) that showed deviation from HWE

81   after which no significant departure was observed. The data for the most common STR

82   profile from the UAE population (based on this dataset) showed that even using a very

83   conservative value of 0.05 for F_{ST} leads to a discrimination power in the order of 10^{15},

84   which translates into a value which is much higher than 1 in a billion. These estimates

85   indicate that UAE might like to adopt a match probability estimates to be reported by the

86   laboratories (in case of a full match) based on statistics generated using its own

87   population allele frequency data. Further work is indicated in this area in order to develop

88   guidelines for forensic DNA laboratories in UAE.

89   Some significant differences were identified between the obtained UAE population data

90   and the other published data. The populations from Iran and Saudi Arabia showed

significant differences at fewer loci when compared with populations from Kuwait, Egypt and India ($P > 0.05$). This is also supported by low F_{ST} value for the Iranian and Saudi Arabian populations. These results support the development of population or location specific databases even when considering populations that are geographically close such as within the Middle East (Supplementary Table 3).

This current dataset establishes the characteristics of the 21 STR loci panel for the identification of individuals, in paternity testing and for crime scene analysis in the UAE.

This manuscript of population data follows the journal guidelines for publication of data described [11, 12 and 13].

Acknowledgments

Appendix A. Supplementary data

Publication\Supplementary Table 1.xlsx

Publication\Supplementary Table 2.xlsx

Publication\Supplementary Table 3.xlsx

References

[1] S.B. Miles. *The countries and tribes of the Persian Gulf,* London, Harrison and Sons (1919).

[2] National Bureau of Statistics. Demographic and Social Statistics. Internet. 19-1-2012. 19-12-2011.

[3] G.O. Tadmouri, P. Nair, T. Obeid, M.T. Al Ali, N. Al Khaja, H.A. Hamamy, Consanguinity and reproductive health among Arabs, Reprod Health. 6 (2009) 17.

[4] I. Abed, P. Hellyer. *United Arab Emirates, a New Perspective,* London, Trident Press (2001).

[5] T. Mohammad, Y. Xue, M. Evison, & C. Tyler-Smith, Genetic structure of nomadic Bedouin from Kuwait, Heredity. 103 (5) (2009) 425-433.

[6] R. Garcia-Bertrand, T.M Simms, A.M Cadenas, R.J Herrera, United Arab Emirates: Phylogenetic relationships and ancestral population s, Gene. 533 (2014) 411-419.

[7] P.M. Schneider, Scientific standards for studies in forensic genetics, Forensic Sci. Int. 165 (2007) 238-243.

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

[9] M. Al-Enizi, J. Ge, S. Ismael, H. Al-Enezi, A. Al-Awadhi, W. Al-Duaij, B. Al-Saleh, Z. Ghulloom, B. Budowle, Population genetic analyses of 15 STR loci from seven forensically-relevant populations residing in the state of Kuwait, Forensic Sci Int- Gen. 7(4) (2013) e106-e107.

[10] C.M. Ruitberg, D.J. Reeder, J.M. Butler, STRBase: A short tandem repeat DNA database for the human identity testing community, Nucl. Acids Res. 29 (1) (2001) 320-322.

[11] A. Carracedo, J.M Butler, L. Gusmão, A. Linacre, W. Parson, L. Roewer, P.M, Schneider, Update of the guidelines for the publication of genetic population data, Forensic Sci. Int. Gen. 10(2014) A1 - A2.

[12] A. Carracedo, J.M Butler, L. Gusmão, A. Linacre, W. Parson, L. Roewer, P.M, Schneider, New guidelines for the publication of genetic population data, Forensic Sci. Int. Gen. 7 (2013) 217-220.

[13] P. Lincoln, A. Carracedo, Publication of population data of human polymorphisms, Forensic Sci Int. 110 (2000) 3-5.