**Population genetic data for 17** **non-CODIS STR loci for the Saudi Arabian population using the SureID®23comp Human Identification Kit.**

Hussain M Alsafiah 1 & William H Goodwin 2

1. Forensic Genetics Laboratory, General Administration of Criminal Evidences, Public Security, Ministry of Interior, Saudi Arabia.
2. School of Natural Sciences, University of Central Lancashire, Preston, United Kingdom.

|  |  |
| --- | --- |
| Corresponding Author: | Hussain M Alsafiah. |

**Abstract**

Our previous work focused on validation the SureID 23comp Human Identification Kit (Health Gene Technologies, China), following the minimum criteria for validation recommended by the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM) using 500 samples from the population of Saudi Arabia. The kit genotypes 22 STRs, 17 of which are non-CODIS, and Amelogenin. The validation tests showed that it has the potential to increase the level of confidence in conclusions in complex cases especially in a country like Saudi Arabia where consanguineous marriage is more common. Previous studies on the population of Saudi Arabia, either by questionnaires or by genetic testing using 21 STRs included in the GlobalFiler kit, suggested an elevated rate of consanguinity in the population of Saudi Arabia.

In this paper, the allele frequency data, common forensic parameters for the 17 non-CODIS STR loci are presented. We examined whether the loci would exhibit excess of homozygosity in the data set.

**Keywords:**

Kinship testing - Non-CODIS STRs – Saudi Arabia

**Introduction:**

The number of STR markers required for the CODIS and for the ESS are sufficient in most kinship cases; however, it is still possible to have inconclusive results in complex kinship cases. They can be further complicated in case of deficient pedigrees or in case of testing distant relationships. In more complex kinship cases, the resolution of kinship testing can be increased by testing additional STRs using complementary STR kits.

Consanguineous marriage, which is more common in Saudi Arabia, can make kinship testing more challenging. Previous studies conducted by questionnaires (for example [1]) on 3212 families found that the percentages of consanguinity were 56.8%. In addition, Khubrani *et al.* [2] studied 523 male samples from the population of Saudi Arabia using the GlobalFiler kit and highlighted excess of homozygosity in 20/21 autosomal STRs (aSTRs). This study estimated the inbreeding coefficient (FIS) in Saudi Arabia by 0.0476, suggesting higher level of consanguinity.

SureID® 23 comp Human Identification kit (Health Gene Technologies, China), multiplexes 17 non-CODIS STRs (D3S1744, D4S2366, D5S2800, D6S474, D7S3048, D8S1132, D9S1122, D11S2368, D13S325, D14S1434, D15S659, D17S1301, D18S1346, D19S253, D20S482, D21S2055, and D22GATA198B05) and D1S1656, D2S441, D10S1248, D12S391, D16S539. Twelve out of the 17 non-CODIS STRs are not included in other available supplementary kits, such as PowerPlex® CS7 System (Promega Corporation) and Investigator® HDplex Kit (Qiagen).

The kit has been internally validated following the minimum criteria for validation recommended by the ENFSI and by the SWGDAM in [3] and has met the criteria commonly used in forensic genetics laboratories which showed that the kit has the potential to increase the level of confidence in conclusions in complex cases.

Here, the allele frequency data for the 17 non-CODIS loci generated using the 500 samples from the population of Saudi Arabia is presented and the commonly forensic statistical parameters were assessed for the population of Saudi Arabia. In addition, we asked whether the 17 non-CODIS loci imitate the elevated level of homozygosity observed in other common aSTRs.

**Materials and Methods:**

Five-hundred blood samples from the population of Saudi Arabia was used in the study. The extraction, quantification procedures are fully described in [4], amplification and analysis in [3]

The forensic parameters were calculated using PowerStat v 1.2 (Promega Promega Corporation). Arlequin v3.5.2 Software was used to calculate the expected heterozygosity (He), to carry out the exact test for detecting deviation from the Hardy-Weinberg equilibrium (HWE) and to estimate the average FIS for the 17 non-CODIS STRs included in the kit.

The data of the 17 non-CODIS STRs was sent to STRidER for quality control check (STR000178).

**Results and Discussion:**

A set of 17 non-CODIS STRs were genotyped using SureID® 23 comp Human Identification kit. Table 1 shows the allele frequency data for the STRs, common forensic parameters and the *P* values of the Exact test for HWE. Assuming no deviation from HWE and all STRs are independent, they showed 1.2E-20 CMP, 0.9999747848 CPE, and 0.999999999999999999988164 CPD. D21S2055 was the most informative locus, with a MP of 0.016, and D17S1301 was the least informative locus, with a MP of 0.162. Heterozygosity ranged from 0.624 (D20S482) to 0.89 (D21S2055). The number of observed alleles per locus varied from 7 alleles in D17S1301 to 20 alleles in D21S2055. The parameters show the usefulness of using the kit as a complementary STR kit.

Fourteen STRs showed fewer than expected heterozygotes (D9S1122, D4S2366 and D8S1132 were the exception). D20S482 was the only locus that showed significant deviation due to excess of homozygosity (*P* value = 0) (Table 1). As null alleles may contribute to this deviation, SNP variants with > 1% frequency at the flanking regions (100 bp each side) of the locus using the 1000 Genome browser [5] were examined. Two SNPs in the 5´flanking region: rs151133985 (all populations C: 99%, G: 1%; Africans C: 98%, G: 2%) and rs77560248 (all populations C: 94%, T: 6%; Europeans and South Asians C: 91%, T: 9%); and one SNP at the 3´flanking region: rs551422781 (Africans G: 99%, A: 1%), were found. These SNPs may cause null alleles if any of them was at a critical annealing region of the primer. However, none of the populations of European, South Asian and African that were studied using the same kit has shown deviation from HWE at this locus [6]. Sequencing homozygotes samples or using a different kit to genotype the locus may reveal more information about the possible contribution of null alleles.

**Conclusion**

The allele frequency data and the common forensic parameters were estimated and are presented in this paper. Due to the D20S482’s deviation, the locus cannot be included in the product rule to calculate the probability of a DNA profile for the population of Saudi Arabia.

Although, the study suggests that the excess of homozygosity observed in Saudi population may be, in part, due to the elevated rate of consanguinity, more studies are needed to find out the other possible reasons of this deviation.

**Acknowledgments**

This work was supported by the Saudi Arabian cultural bureau in London (UKSACB).

**Conflict of interest statement**

The authors declare no Conflict of interests.

**References**

[1] M. El-Hazmi; A. Al-Swailem; A. Warsy; R. Sulaimani and a. Al-Meshari, Consanguinity among the Saudi Arabian Population, Journal of Medical Genetics 32(8) (1995) pp. 623-626.

[2] Y. Khubrani; J. Wetton and M. Jobling, Analysis of 21 Autosomal STRs in Saudi Arabia Reveals Population Structure and the Influence of Consanguinity, Forensic Science International: Genetics, 39, (2019) pp. 97-102.

[3] H. Alsafiah; A. Aljanabi; S. Hadi; S. Alturayeif and W. Goodwin, An Evaluation of the SureID 23comp Human Identification Kit for Kinship Testing, Scientific Reports, 9 (1) (2019) pp. 16859.

[4] H. Alsafiah; W. Goodwin; S. Hadi; M. Alshaikhi and P. Wepeba, Population Genetic Data for 21 Autosomal STR Loci for the Saudi Arabian Population using the GlobalFiler® PCR Amplification Kit, Forensic Science International: Genetics, 31 (Supplement C) (2017) pp. e59-e61.

[5] A. Auton; G. R. Abecasis; D. M. Altshuler; *et al.*, A Global Reference for Human Genetic Variation, Nature, 526 (7571) (2015) pp. 68-74.

[6] S. Iyavoo; O. Afolabi; B. Boggi; A. Bernotaite and T. Haizel, Population Genetics Data for 22 Autosomal STR Loci in European, South Asian and African Populations using SureID® 23comp Human DNA Identification Kit, Forensic Science International, 301 (2019) pp. 174-181.