Legal Medicine 26 (2017) 14–17

Contents lists available at ScienceDirect

Legal Medicine

journal homepage: www.elsevier.com/locate/legalmed

Study of autosomal STR loci with IBS method in full sibling identification

Li Yuan [a,b,*], Xu Xu [b], Dong Zhao [a,b], He Ren [c], Chaohui Hu [d], Wen Chen [b], Shicheng Hao [b], Di Lu [a,b], Lin Zhang [e]

[a] Collaborative Innovation Center of Judicial Civilization, Beijing 100088, PR China
[b] Key Laboratory of Evidence Science (China University of Political Science and Law), Ministry of Education, Beijing 100088, PR China
[c] Beijing Police College, Beijing 102022, PR China
[d] College of Mathematics and Computer Science, Hunan Normal University, Changsha 410081, PR China
[e] West China School of Preclinical and Forensic Medicine, Sichuan University, Chengdu 610041, PR China

ARTICLE INFO

Article history:
Received 3 May 2016
Received in revised form 25 January 2017
Accepted 31 January 2017
Available online 2 February 2017

Keywords:
Forensic biological evidence
Full-sib relation
Short tandem repeat
Discriminant analysis
Allele identity by state

ABSTRACT

Objective: We investigated the application of 51 autosomal short tandem repeat (STR) loci with the identity by state (IBS) method and a discriminant function algorithm in full-sib identification. *Methods:* A total of 342 pairs of full sibs (FSs) and 3900 pairs of unrelated individuals (UIs) were genotyped for 51 STR loci. Groups were formed in accordance with discrimination power (DP) values and the number of loci, and IBS scores of FSs and UIs were analyzed and compared. The discriminant functions of FS-UI were determined by using the Fisher discriminant with SPSS software. *Results:* All IBS in FSs and UIs groups showed normal distributions and there were significant differences between FS-UI. Receiver operating characteristic curves revealed that the detection efficiency of full-sib identification was affected by both the locus polymorphism and the number of loci detected. Comparing the rate of false positive and false negative of discriminant function between groups, a higher average DP value and larger number of loci detected were associated with a lower rate of miscarriage of justice and were more helpful for full-sib identification. *Conclusion:* STRs with higher DP values should be selected when additional autosomal markers are required for FS identification. Discriminant analysis with the IBS method is highly applicable for the FS-UI test.

1. Introduction

Full-sib identification plays an important role in individual identification in criminal and civil law cases and in searching for lost persons when the parents are absent [1]. Full-sib identification precisely can be difficult because the alleles at each short tandem repeat (STR) locus can be all the same, partially the same, or completely different between siblings. Autosomal STR analysis has been applied in full-sib identification using the identity by state (IBS) method [2–4]. According to Mendelian genetic laws, alleles of full-sibs are acquired from the parents. The numbers of 1/4, 1/2, or 1/4 are the probabilities that full-sibling pairs share 0, 1, or 2 alleles identical-by-descent, respectively. Unrelated individuals may have same alleles just because of random matching. In this theory, calculating the number of shared alleles can be adopted for full-sib identification.

The IBS method is simple and does not require knowledge of allele frequencies. It is under discussion to establish a Public Safety Industry Standard based on IBS method for the identification of full sibling and unrelated individuals in China. However, the effects of the number of genetic markers and the power of discrimination of tested STR loci on the IBS method are not well understood [5]. In the present study, we detected 51 autosomal STR loci for each individual, and made comparisons between the variant combinations, such as combinations of the same DP values but with different number of loci, and combinations of same number of loci but with different DP values. In addition, we investigated the effect of autosomal STRs on the IBS method in the context of full-sib identification.

2. Methods

2.1. Samples and STR analysis

Three hundred and forty-two Beijing Han full-sib pairs of all selected families agreed to sign informed consent to participate in the study. Selected full-sib pairs from triplet families with $\geqslant 2$ offspring were identified to be brothers and sisters of the same father and mother, with exclusion of identical twins and the presence of pedigree STR mutation after the "father-child-mother"

* Corresponding author at: Collaborative Innovation Center of Judicial Civilization, Beijing 100088, PR China.
E-mail address: yuanliwcy@126.com (L. Yuan).

http://dx.doi.org/10.1016/j.legalmed.2017.01.010
1344-6223/

paternity test. Additionally, 3900 unrelated pairs included randomly paired combinations of unrelated individuals. The sample collection, DNA analysis, and storage were conducted in accordance with humane and ethical research principles of Key Laboratory of Evidence Science, China University of Political Science and Law.

DNA samples were amplified using Goldeneye20A™ (Peoplespot Inc., Beijing, China), AGCU EX22, AGCU 21 + 1 (AGCU ScienTech Incorporation, Wuxi, China), and a 5-colored fluorescent multiplex amplification system established in our lab [6]. For each sample, 51 STR loci were detected (Table S1). PCR products were separated by electrophoresis on an ABI 3130 genetic analyzer, and ABI Genemapper ID 3.2 software was used for analysis (Applied Biosystems, Foster City, CA, USA). Internal control standards within the laboratory were used according to recommendation of the Paternity Testing Commission of the International Society for Forensic Genetics [7].

2.2. Genetic polymorphisms

In this study, 323 healthy unrelated individuals of the Beijing Han population were randomly selected and 51 STR loci were detected for genetic polymorphisms in this population. The genotype frequency distribution of STR loci was subject to Hardy-Weinberg equilibrium testing using GENEPOP (4.0) software [8], and linkage disequilibrium between loci was evaluated. PowerStatsV12 software from Promega Corporation (Madison, WI, USA) was used to calculate allele frequencies.

2.3. ibs and IBS

Based on the STR typing results of full-sibs (FSs) and unrelated individuals (UIs), in accordance with the value of each locus's discrimination power (DP) in descending order, the STR combinations were formed as follows: I, combination of the 19 STR loci with the lowest DP values; II, combination of 19 STR loci for routine work (including 13 CODIS loci); III, combination of 19 STR loci with the highest DP values; IV, combination of selected 29 STR loci (including combination II); and V, combination of selected 39 STR loci (including combination IV). Combinations IV and V had similar average DP values to combination II.

Then, ibs and IBS scores were calculated for each STR locus and for different STR combinations, respectively. In each STR locus, the number of shared alleles between individuals is known as the ibs score. The IBS score is the cumulative sum of ibs scores for a genotyping system.

2.4. Statistical analysis and discriminant function creation

SPSS 19.0 software (SPSS Inc., Chicago, IL, USA) was used for descriptive statistics for IBS, *t* test of mean IBS of FSs and UIs, Fisher discriminant analysis, and function creation. Matlab 2013a software (MathWorks, Natick, MA, USA) was used for IBS normal distribution curve fitting analysis of the five groups of FSs and UIs; State/MP 13.1 software was used to depict and compare the area and the difference under the receiver operator characteristic (ROC) curves [9].

3. Results and discussion

3.1. Hardy-Weinberg equilibrium test and linkage disequilibrium test

Statistical analysis was performed for the genotype distribution of 51 STR loci; there were no significant differences in Hardy-Weinberg equilibrium with Bonferroni correction (p-value, 0.05/51 = 0.00098) [10]. These STR loci showed p values of less than 0.05 in the linkage disequilibrium test (Table S2) for 93 of 1275 pairs of loci, and nine pairs showed significant differences after Bonferroni correction (p-value, 0.05/1275 = 0.000039), including FGA/D6S1043, FGA/D20S470, D7S820/D18S51, D7S3048/D20S470, D8S1179/CSF1PO, D9S1122/D17S1290, D12ATA63/D18S853, D12ATA63/D17S1290 and D13S317/Penta E. Given that these nine pairs of STR loci are located on different chromosomes, we did not consider the presence of linked inheritance for these loci. In practice, if two STR loci are in LD, the loci should not be used together simply in kinship analysis. Considering the relatively small population and the purpose of the present study, we did not exclude the use of these STR loci of potential LD, because many previous studies have showed no LD for these STR pairs [11,12]. Still, further study of a larger population is needed to clear the effects of LD and genetic linkage on IBS methods.

In the Beijing Han Chinese population, obvious differences were observed in the discrimination power for individual loci, such as DP values over 0.950 0 for Penta E, D7S3048, D6S1043, D18S51, D20S470, D2S1338, FGA, D15S659, D12S391, D11S2368 and D8S1179, whereas DP values of 0.8500 or lower for D1S1677, D1S1627, TH01, D1GATA113, and TPOX. Therefore, the impact of the DP values of the supplemental STRs on the final determination must be considered when performing FS-UI identification.

We combined the STRs according to the grouping method in accordance with the average DP values shown in Table 1.

3.2. ibs of single STR locus

Mean ibs and mean ibs differences at single STR locus between FSs and UIs are shown in Table S1. An important indicator used to evaluate whether an STR is beneficial for FS identification is the degree of difference in sharing alleles in FSs and UIs; greater differences make it easier to distinguish FS-UI and vice versa. Comparing the mean ibs difference of the 51 STR loci between FSs and UIs (shown in Table S1 and Fig. 1), as the DP value increases, the difference of mean ibs increased. These results indicated that high polymorphic genetic markers should be selected for FS identification.

In the I–V groups of STR combinations, normal distribution curve-fitting results for 342 FS pairs and 3900 UI pairs are shown in Fig. S1. In these five combinations, there were significant differences in IBS scores between FSs and UIs (Table S3). Table S4 shows the proportion of FSs and UIs in the overlapping area of same IBS on the two curves.

Table 1
STR loci grouping.

Combination	Loci	Average DP
Group I (19)	D1GATA113, D1S1627, D1S1677, D2S1776, TPOX, D3S1358, D3S4529, D4S2408, D5S2500, CSF1PO, D6S474, D6S1017, D9S1122, TH01, D12SATA63, D14S1434, D17S1301, D18S853 and D20S482;	0.8638
Group II (19)	D2S1338, TPOX, D3S1358, FGA, D5S818, CSF1PO, D6S1043, D7S820, D8S1179, TH01, D12S391, vWA, D13S317, Penta E, D16S539, D18S51, D19S433, D21S11and Penta D;	0.9220
Group III (19)	D2S1338, D3S1744, FGA, D6S1043, D7S3048, D8S1179, D11S2368, D12S391, D13S317, D14S608, D15S659, Penta E, D17S1290, D18S51, D18S535, D20S470, D21S11, Penta D and GATA198B05;	0.9515
Group IV (29)	Group II +D2S441, D4S2366, D7S3048, D10S1248, D11S4463, D15S659, D17S1290, D19S253, D20S470 and D22S1045;	0.9225
Group V (39)	Group IV + D2S1776, D3S1744, D6S474, D9S925, D10S1435, D11S2368, D14S608, D18S535, D18S853 and GATA198B05.	0.9224