Legal Medicine 19 (2016) 122–126

Contents lists available at ScienceDirect

Legal Medicine

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

Short Communication

Sibling assessment based on likelihood ratio and total number of shared alleles using 21 short tandem repeat loci included in the GlobalFiler™ kit

Hiroyuki Inoue [a,b,*], Sho Manabe [c], Koji Fujii [d], Yasuki Iwashima [d], Saori Miyama [a], Ayako Tanaka [a], Hisako Saitoh [b], Hirotaro Iwase [b], Keiji Tamaki [c], Kazumasa Sekiguchi [d]

[a] Criminal Investigation Laboratory, Chiba Prefecture Police H.Q., Japan
[b] Department of Legal Medicine, Graduate School of Medicine, Chiba University, Japan
[c] Department of Forensic Medicine, Graduate School of Medicine, Kyoto University, Japan
[d] National Research Institute of Police Science, Japan

ARTICLE INFO

Article history:
Received 2 March 2015
Received in revised form 5 July 2015
Accepted 20 July 2015
Available online 21 July 2015

Keywords:
Forensic science
Sibship determination
Likelihood ratio (LR)
Total number of shared alleles (TNSA)
GlobalFiler kit
Identifiler kit

ABSTRACT

Sibling assessment using the 15 autosomal short tandem repeat (STR) loci included in the Identifiler® kit can be difficult when comparing an unidentified party to an alleged sibling. Therefore, we investigated the likelihood ratio (LR) and the total number of shared alleles (TNSA) for sibship determination using the 21 autosomal STR loci included in the GlobalFiler™ kit.

We computationally generated the genotypes of 10,000 sibling pairs and 10,000 unrelated pairs based on previously reported allele frequencies of the 15 Identifiler loci and the remaining 6 GlobalFiler loci. The LR and the TNSA were then calculated in each pair using the 15 and 21 loci. Next, these calculations were applied to 22 actual sibling pairs.

LR values $\geqslant$10,000 were observed in 48% of the sibling pairs using the 15 loci and in 80% of the sibling pairs using the 21 loci. The TNSA distribution between siblings and unrelated pairs was more divergent in GlobalFiler than in Identifiler. TNSA values $\geqslant$20 were found only in true siblings in Identifiler, while TNSA values $\geqslant$24 in GlobalFiler. In Identifiler, all pairs with TNSA $\geqslant$24 had LR values $\geqslant$10,000 and the same was true in GlobalFiler for TNSA $\geqslant$29. Therefore, increasing the number of loci is very efficient for sibship determination.

The LR is most reliable for determining sibship. However, TNSA values may be useful for the preliminary method of LR values because LR value demonstrated a significantly positive correlation with TNSA value in both Identifiler and GlobalFiler.

1. Introduction

Requests for personal identification have been increasing in Japanese forensic casework. The unidentified parties (UPs) can be identified by comparing them with reference materials such as personal belongings, facial and physical features, fingerprints, dental work, and DNA information. DNA information is especially effective for personal identification due to its reliability and power of discrimination.

The best way to identify unknown human remains by using DNA information, is to compare DNA profiles from the remains with profiles from direct reference samples obtained from personal items such as toothbrushes and razors. If the results match, the alleged individual can be simply identified from the remains. However, if such direct references are unavailable, indirect reference samples collected from relatives of the UP, i.e., parents, children, and/or siblings, have to be used. In parentage testing, identification of the parent–child relationship is straightforward because at least one allele is shared at each locus between the UP and the alleged parent or child, except for mutations. If DNA samples from parents or children cannot be obtained, a sibling test between the UP and alleged siblings (ASs) can be performed for personal identification. However, sibship determination may be difficult if the number of ASs is small, because the pair of siblings may not necessarily have common alleles [1].

Kinship analysis include the combined likelihood ratio (LR) calculated by the exact method [1], or the LR calculated by the Identity-By-State (IBS) method [2,3]. The LR compares the probabilities of the evidence under alternative hypotheses. In a sibling test, the two hypotheses of the LR are H_1: the UP is a sibling of

* Corresponding author at: Department of Legal Medicine, Graduate School of Medicine, Chiba University, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan.
E-mail address: inoue@restaff.chiba-u.jp (H. Inoue).

http://dx.doi.org/10.1016/j.legalmed.2015.07.008
1344-6223/

the ASs, and H_2: the UP is unrelated to the ASs. Let E denote the DNA evidence profiles from the UP and the ASs. The LR is then defined as follows:

$$\mathrm{LR} = \frac{\Pr(E|H_1)}{\Pr(E|H_2)}$$

If the LR is greater than one, the evidence supports H_1, but if it is less than one, the evidence supports H_2. However, it may be difficult to assess the strength of the evidence based on the LR because the interpretation of the LR is very complicated for non-specialists.

In Japanese forensic casework, DNA typing for 15 autosomal short tandem repeat (STR) loci using the AmpFℓSTR® Identifiler® PCR Amplification Kit (Life Technologies, Carlsbad, CA) is currently used for personal identification. When there is only one AS, we often experience the difficulties with sibship determination using the Identifiler kit (ID) because the LR value does not always demonstrate strong support for H_1. Recently, the GlobalFiler™ PCR Amplification Kit (Life Technologies, Carlsbad, CA) has been developed. The kit genotypes the 21 autosomal STRs (including the 15 ID loci), Amelogenin, DYS391 and Y-Indel loci. In general, increasing the number of tested loci improves the ability of the kit to determine whether or not a relationship exists [4,5].

In this study, we compared the efficacy of the GlobalFiler kit (GF) and ID for sibship determination. We synthesized sibling and unrelated pairs *in silico* using the 21 autosomal STR loci in GF, and estimated the LR in each pair using the 21 GF loci and the 15 ID loci. Next, we calculated the total number of shared alleles (TNSA) in each pair and investigated the correlation between the LR and the TNSA values. The TNSA approach has a more intuitive understanding than the LR approach because the TNSA values are easily calculated by summing up the number of alleles (0, 1, or 2) shared identical by state by pairs of individuals in each locus [6]. We further analyzed samples from 22 actual sibling pairs and discussed the effectiveness of 21 STR typing system.

2. Materials and methods

2.1. Simulation of the LR and the TNSA in full sibling assessments

We first generated 40,000 simulated individuals based on previously reported allele frequencies of 21 autosomal STR loci in GF. GF contains a total of 24 loci. However, the DYS399, Y Indel, and Amelogenin loci are primarily used to determine the sex and were therefore excluded from the study. We used Japanese allele frequencies [7] for 15 loci included in ID (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818, and FGA). For the remaining 6 loci (D12S391, SE33, D1S1656, D2S441, D10S1248, and D22S1045), we used NIST population data (i.e., African American ($n = 342$), Caucasian ($n = 361$), Hispanic ($n = 236$), and Asian ($n = 97$)) [8]. Next, we randomly selected 20,000 individuals to form pairs of parents. The parents were then simulated to give birth to two offspring whose alleles were randomly selected from the parents. We defined these 10,000 two-offspring pairs as sibling pairs. Using the remaining 20,000 individuals, we synthesized 10,000 unrelated pairs.

The LR for sibship determination was then calculated for each sibling and unrelated pair based on the kinship coefficient (k) equation proposed by Wenk [1]. There were several reports on the strength of the LR value supporting the evidence. In this study, we adopted the recommendation by Evett et al. who stated that a LR $\geqslant$10,000 indicates that the evidence is very strongly supportive [9]. Therefore, we determined sibling relationship between two simulated individuals as an LR $\geqslant$10,000. Additionally, we counted the TNSA for each sibling and unrelated pair, and investigated the correlation between the TNSA and the LR values by regression analysis. The LR and the TNSA were calculated using the 15 ID and 21 GF loci. All programs used for the simulations were written using the statistical software R (version 3.0.2) [10].

2.2. DNAprofiling using GF

We obtained 44 buccal samples (Bode Technology, Lorton, VA) from 22 pairs of siblings after informed consent, from which DNA was extracted using the EZ1 DNA Investigator Kit (Qiagen GmbH, Hilden, Germany). DNA profiling was carried out using the GlobalFiler™ PCR Amplification Kit (Life Technologies, Carlsbad, CA) according to the manufacturer's instructions. The PCR of GF was carried out using 1 ng of extracted DNA as the template in a reaction volume of 25 μL using GeneAmp® PCR System 9700 (Life Technologies, Carlsbad, CA). The PCR products were electrophoresed on an ABI 3500xL Genetic Analyzer (Life Technologies, Carlsbad, CA) and analyzed using GeneMapper IDX Software version 1.4 (Life Technologies, Carlsbad, CA) to obtain STR profiles. For each pair, we calculated the LR using the DNA-VIEW program developed by C.H. Brenner. The allele frequencies were used the same as the simulation. Then, we investigated the correlation between the LR and the TNSA values in each pair. This study was approved by the Ethics Committee of the National Research Institute of Police Science.

3. Results

3.1. Simulation of the LR in sibling assessments

In true sibling pairs, the LR distribution was shifted to the right, when the number of loci was increased from 15 of ID to 21 of GF (Fig. 1). The median LR values were 8.06×10^3 for ID and 1.46×10^6 for GF, respectively. LR values exceeding the temporal LR threshold of 10,000 were observed in 4823 pairs (48%) in ID and 7999 pairs (80%) in GF. Conversely, in unrelated pairs, the LR distribution was shifted to the left, when the number of loci was increased from 15 to 21. There were no cases where the LR value exceeded 10,000 in both of ID and GF.

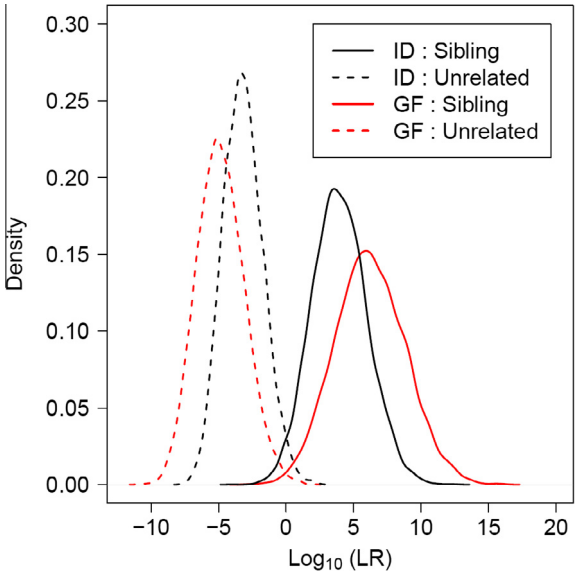

Fig. 1. LR distribution of sibling pairs (solid lines) and unrelated pairs (dashed lines) in each 10,000 simulation using ID 15 loci (black lines) and GF 21 loci (red lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)