Legal Medicine 17 (2015) 320–325

Contents lists available at ScienceDirect

Legal Medicine

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

Evaluation of advanced multiplex short tandem repeat systems in pairwise kinship analysis

Tomonori Tamura [a,b], Motoki Osawa [a,*], Eriko Ochiai [a], Takanori Suzuki [b], Takashi Nakamura [b]

[a] Department of Forensic Medicine, Tokai University School of Medicine, Shimokasuya 143, Isehara, Kanagawa 259-1193, Japan
[b] Scientific Crime Laboratory, Kanagawa Prefectural Police, Yamashitacho 155, Naka-ku, Yokohama 231-0023, Japan

ARTICLE INFO

Article history:
Received 24 September 2014
Received in revised form 23 March 2015
Accepted 23 March 2015
Available online 1 April 2015

Keywords:
Kinship analysis
Short tandem repeat
Likelihood ratio
PowerPlex Fusion
GlobalFiler
PowerPlex 21

ABSTRACT

The AmpFLSTR Identifiler Kit, comprising 15 autosomal short tandem repeat (STR) loci, is commonly employed in forensic practice for calculating match probabilities and parentage testing. The conventional system exhibits insufficient estimation for kinship analysis such as sibship testing because of shortness of examined loci. This study evaluated the power of the PowerPlex Fusion System, GlobalFiler Kit, and PowerPlex 21 System, which comprise more than 20 autosomal STR loci, to estimate pairwise blood relatedness (i.e., parent–child, full siblings, second-degree relatives, and first cousins). The genotypes of all 24 STR loci in 10,000 putative pedigrees were constructed by simulation. The likelihood ratio for each locus was calculated from joint probabilities for relatives and non-relatives. The combined likelihood ratio was calculated according to the product rule. The addition of STR loci improved separation between relatives and non-relatives. However, these systems were less effectively extended to the inference for first cousins. In conclusion, these advanced systems will be useful in forensic personal identification, especially in the evaluation of full siblings and second-degree relatives. Moreover, the additional loci may give rise to two major issues of more frequent mutational events and several pairs of linked loci on the same chromosome.

1. Introduction

Pedigree analysis in forensic science plays a major role in the identification of missing persons in mass disasters and accidents. In cases in which the personal remains of the deceased are unavailable, indirect identification is often performed by confirmation of any kinship with a potential family member. In addition to parentage testing, DNA analysis for other relatives such as full siblings, second-degree relatives, and first cousins is sometimes required in Japan. The clarification of the relationship between two individuals is required in the most common situation. For instance, forensic investigators are asked questions such as if the alleged father is the true father of the child or not and whether two males are brothers or not. The Paternity Testing Commission of the International Society of Forensic Genetics (ISFG) recommends the likelihood ratio (LR) approach if the weight of genetic evidence is calculated in parentage testing [1], although other parameters are also utilized for parentage testing (e.g., exclusion probability and probability based on Bayes' theorem).

The LR approach can be extended to other relationships such as full and half siblings [2,3]. When the genotypes used are mutually independent, the combined LR is calculated by multiplying LRs according to the product rule. If necessary, the LR and combined LR are adjusted for the effects of subpopulation and linkage [4]. Parentage is inferred according to typical cutoffs recommended by the ISFG: 100 or 1000 [5]. In sibship testing, a cutoff of $\leqslant$100 is usually acceptable according to the American Association of Blood Banks (AABB) [6]. However, there is no universal cutoff among laboratories. An LR of 500 [7], which is Hummel's paternity criterion during the blood group typing, is preferred in Japan regardless of the type of kinship.

Genetic polymorphisms of short tandem repeat (STR) play a critical role in personal identification. The AmpFLSTR® Identifiler® PCR Amplification Kit (Applied Biosystems, Foster City, CA, USA), which is composed of 15 autosomal STR loci (13 common loci of CODIS, D2S1338, and D19S433) and the Amelogenin locus, has been adopted as the nationwide standard in Japan. This system is used for the match probability calculation as well as LR calculation in duo and trio parentage testing.

Abbreviations: AABB, American Association of Blood Banks; ISFG, International Society of Forensic Genetics; STR, short tandem repeat; LR, likelihood ratio.

* Corresponding author. Tel.: +81 463 93 1121; fax: +81 463 92 0284.
E-mail address: osawa@is.icc.u-tokai.ac.jp (M. Osawa).

http://dx.doi.org/10.1016/j.legalmed.2015.03.005
1344-6223/

However, Tamaki et al. previously found that this kit occasionally has limitations when obtaining a significant LR in pairwise sibship testing [8,9]. This limitation can be overcome by analyzing additional autosomal STR loci [10]. There are three commercially available major multiplex STR systems that comprise more than 20 autosomal STR loci: the PowerPlex® Fusion System (Promega Corporation, Madison, WI, USA), which comprises 22 autosomal STR loci of 13 CODIS, D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, D22S1045, Penta D, and Penta E; the GlobalFiler® PCR Amplification Kit (Applied Biosystems), which comprises 21 autosomal STR loci of 13 CODIS, D1S1656, D2S441, D2S1338, SE33, D10S1248, D12S391, D19S433, and D22S1045; and the PowerPlex® 21 System (Promega Corporation), which comprises 20 autosomal STR loci of 13 CODIS, D1S1656, D2S1338, D6S1043, D12S391, D19S433, Penta D, and Penta E. These three systems include five to seven loci in addition to those in the Identifiler.

This study evaluated the power of the PowerPlex Fusion, GlobalFiler, and PowerPlex 21 to solve different cases of pairwise kinship analysis using simulated genotype data. In addition, the analyses were performed with the maximum set of 24 autosomal STR loci included in the three advanced systems. Finally, we discussed the problems of mutation and linkage associated with the addition of examined loci.

2. Materials and methods

2.1. Simulated STR type data for putative pedigrees

For the simulation study, 10,000 datasets of five different pairwise relations of parent–child, full siblings, second-degree relatives, first cousins and non-relatives were constructed under a putative pedigree chart, representatively showed in Fig. 1. Individuals #1 and #6 represent the sets as parent–child, #5 and #6 as full siblings, #5 and #7 as second-degree relatives, #7 and #8 as first cousins, and #3 and #4 as non-relatives. Second-degree relatives are two meioses away from a particular individual in a pedigree (i.e., grandparent–grandchild, uncle/aunt–nephew/niece, and half-siblings). First, a series of genotypes in individuals #1–#4 were constructed on the basis of the allele frequencies of the 24 STR loci in the Japanese population [11–13]; then, those in individuals #5–#8 were constructed by mating of those of #1–#4. Mutational events were not considered in this process. Uniform random numbers were produced using R version 3.1.0 [14], and the construction of STR type data in the putative pedigrees were conducted in Microsoft Excel (2010).

2.2. Combined LR

Four different case scenarios—parent–child (H_1) vs. non-relatives (H_2), full siblings (H_1) vs. non-relatives (H_2), second-degree relatives (H_1) vs. non-relatives (H_2), and first cousins (H_1) vs. non-relatives (H_2)—were investigated by simulations. The simulation was performed with 10,000 sets of STR types where hypothesis H_1 was assumed to be true, and those where hypothesis H_2 was assumed to be true. The LR for each locus was calculated by dividing the joint probability assuming the two individuals were relatives by that assuming the two individuals were non-relatives on the basis of the allele frequencies in the Japanese population. For the five sets of the 24 STR loci, PowerPlex Fusion, GlobalFiler, PowerPlex 21, and Identifiler, the combined LRs were calculated by multiplying LRs based on the product rule. The linkage and subpopulation were not considered in this calculation. The combined LR was calculated using Microsoft Excel (2010).

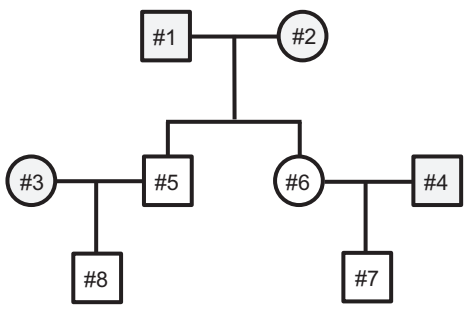

Fig. 1. Putative pedigree chart for the construction of pairwise kinship.

2.3. Distribution of combined LRs, sensitivity, and specificity

The distribution of the combined LR was depicted as a density curve using R version 3.1.0. In addition, sensitivity (the probability of judging relatives correctly as relatives), specificity (the probability of judging non-relatives correctly as non-relatives), positive predictive value (the proportion of subjects correctly judged as relatives), and negative predictive value (the proportion of subjects correctly judged as non-relatives) were calculated for different cutoffs of the combined LR. The power of the 24 STR loci, PowerPlex Fusion, GlobalFiler, and PowerPlex 21 to solve cases in the above mentioned scenarios was evaluated by comparing the sensitivity, specificity, positive predictive value, and negative predictive value of these advanced systems with those of Identifiler. All values were calculated using Microsoft Excel (2010).

This project was approved by the Ethics Committee of the Tokai University School of Medicine.

3. Results and discussion

3.1. Parent–child scenario

For the five sets of 24 STR loci, PowerPlex Fusion, GlobalFiler, PowerPlex 21, and Identifiler, the distribution of the common logarithm of the combined LR observed in the putative parent–child

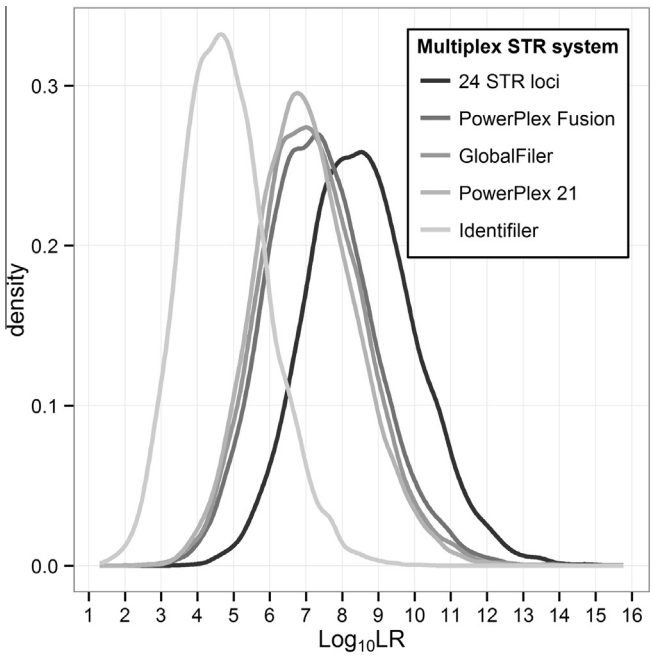

Fig. 2. Distribution of combined LR (common logarithm) in parent–child.