Legal Medicine 17 (2015) 205–209

Contents lists available at ScienceDirect

Legal Medicine

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

Development of a novel miniplex DNA identification system for the Japanese population

Munetaka Ozeki [a,*], Hiroki Kato [b], Keiji Tamaki [a]

[a] Department of Forensic Medicine, Graduate School of Medicine, Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
[b] Department of General Medicine/Family and Community Medicine, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8560, Japan

ARTICLE INFO

Article history:
Received 22 July 2014
Received in revised form 31 October 2014
Accepted 16 December 2014
Available online 24 December 2014

Keywords:
DNA typing
Japanese population
Kinship analysis
Short tandem repeat

ABSTRACT

To aid DNA identification using 36 short tandem repeat (STR) loci for kinship analysis, likelihood ratio (LR) distributions were estimated using the allele frequency data evaluated for the Japanese population in our previous study. The results revealed that the LR tended to be higher when kinship analysis was performed using the 36 STR loci than when the analysis was performed using Identifiler®, the most commonly used commercial DNA typing kit in Japan, even when a sibship case was analyzed. However, a typing kit targeting 36 STR loci is not suitable for the analysis of damaged DNA. In this study, we developed a novel miniplex DNA identification system targeting 7 STR loci (D3S1744, D5S818, D8S1179, D10S2325, Penta D, Penta E, and vWA) that was optimized for use in combination with MiniFiler™ for the Japanese population. The combined matching probability of the MiniFiler plus miniplex system was 4.8×10^{-19}, which is lower than that of Identifiler (4.3×10^{-17}). All expected alleles were detected successfully on DNA isolated from HepG2 human hepatocarcinoma cells by the miniplex system, but no significant signal was observed from a DNA sample isolated from COS-7 African green monkey kidney cells.

1. Introduction

DNA profiling, also known as DNA typing, is a powerful technique employed to assist the identification of individuals by their respective DNA profiles. The technique is routinely used in many fields, including forensic sciences and population genetics. It is also used to identify remains in historical samples [1]. Due to its high discriminatory power, data from DNA profiling analyses remain the most important pieces of evidence in many investigations, specifically those involving crime scenes, mass disasters, and missing persons, as well as in cases of paternity testing.

The Combined DNA Index System (CODIS) is one of the most popular DNA typing systems used worldwide. However, despite its high power of discrimination, CODIS needs to be improved through the addition or exchange of loci to the index system to enable efficient typing/profiling in the present-day scenario with respect to crime scene investigations or social requirements in the United States [2,3]. Notably, it is difficult to obtain allelic information from damaged DNA—such as those obtained in investigations involving long-term missing persons—because of DNA fragmentation and modifications; typing of such DNA is insufficient to identify the individual. For example, Fondevila et al. reported a DNA identification case involving a set of skeletal remains found after a forest fire, which was later determined as belonging to a person missing since 10 years [4]. However, in this case, they obtained insufficient DNA typing results even on using 3 short-amplicon short tandem repeat (STR) sets, and therefore had to identify the individual using single nucleotide polymorphism (SNP) analysis.

In this regard, many studies have focused on improving DNA characterization methodologies for degraded or damaged DNA. One potential approach to overcome this limitation is to target STRs on the Y chromosome or mitochondrial DNA, which provide additional information and act as lineage markers. However, these applications are limited to the identification of paternal or maternal relationships. SNPs have also been shown to be suitable in the analysis of damaged DNA because of their very short amplicon sizes [4]. However, they are not suitable for the analysis of mixed DNA samples.

In our previous study, we reported the allele frequencies for 37 STR loci using a combination of 4 commercially available DNA identification kits [5]. In this study, we performed kinship analysis using the extended locus set and developed a novel miniplex system that may be used in combination with the AmpFlSTR® MiniFiler™ polymerase chain reaction (PCR) Amplification Kit to enable the typing of damaged DNA. We propose that DNA typing using the extended locus set may be the most informative method available for the analysis of damaged DNA.

* Corresponding author. Tel.: +81 75 753 4474; fax: +81 75 761 9591.
E-mail address: ozeki@fp.med.kyoto-u.ac.jp (M. Ozeki).

http://dx.doi.org/10.1016/j.legalmed.2014.12.007
1344-6223/

2. Materials and methods

2.1. Kinship analysis simulation using 36 STR loci for the Japanese population

Likelihood ratio (LR) distributions for the kinship analyses performed using the Life Technologies's AmpFlSTR® Identifiler® (Identifiler), AmpFlSTR® MiniFiler™ (MiniFiler; Carlsbad, CA) PCR Amplification Kit locus sets, and the 36 loci characterized in our previous study were estimated using the calculation method described by Ge et al. [6]. To determine whether a missing person (mp) was related to a putative family reference person (f: parent/child, full sibling [full-sib], or half sibling [half-sib]), identification was assessed by comparing two alternative hypotheses: H_r, mp is related to the putative family member, and H_u, mp is unrelated to the putative family member. The LR was calculated based on a ratio of the probabilities of the DNA evidence under each hypothesis, represented by the general expression:

$$\text{Likelihood ratio} = \frac{Pr(G_{mp}, G_f|H_r)}{Pr(G_{mp}, G_f|H_u)}$$

where G_{mp} refers to the DNA profile of the missing person (obtained from remains) and G_f represents the DNA profile of the putative family member. The LR distribution for 10,000 cases was simulated using the Python 2.7 software and the allele frequency data for the Japanese population published previously [5].

2.2. DNA preparation

HepG2 human hepatocarcinoma cells were provided by the RIKEN BioResource Center (BRC) through the National BioResource project of the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan. COS-7 African green monkey kidney cells were obtained from the Japanese Collection of Research Bioresources (JCRB) Cell Bank (Osaka, Japan). The cells were maintained in minimum essential medium supplemented with 10% fetal bovine serum and antibiotics/antimycotics at 37 °C in a 5% CO_2 incubator. Genomic DNA from cultured cells was prepared using a QIAamp DNA mini kit (QIAGEN, Venlo, Netherlands) according to the manufacturer's instructions. DNA typing for the HepG2 cells was performed using the following DNA typing kits: Identifiler, Promega PowerPlex® ESX17, PowerPlex® CS7 custom (Madison WI), and QIAGEN HDplex, according to the manufacturers' instructions. All PCR products were applied onto an Applied Biosystems® 310 Genetic analyzer, and the results were analyzed using the GeneMapper® version 4.0 software.

2.3. Miniplex analysis

Fluorescence-labeled PCR primers used for the novel miniplex PCR designed in this study are listed in Table 1. All primers were designed to amplify all DNA fragments less than 250 base pairs (bp) in size. To avoid an extra adenine overhang at the 3′ terminal, KOD FX Neo DNA polymerase was used because of its 3′–5′ exonuclease activity [7]. Primer concentrations were adjusted empirically to balance peak areas within dye colors. Genomic DNA (2 ng) was added to 20 μL of a PCR master mix containing 1 × PCR buffer for KOD FX Neo DNA polymerase, deoxynucleotide triphosphates (0.4 mM), and 0.4 U of KOD FX Neo DNA polymerase (Takara, Shiga, Japan). The PCR was performed using an Astec thermal cycler (Astec Inc., model #482, Fukuoka, Japan) under the following conditions: 94 °C for 2 min followed by 28 cycles of 94 °C for 15 s, 59 °C for 30 s, and 68 °C for 30 s, and a final incubation at 25 °C. Loading samples were prepared by adding 2 μL of the PCR mixture to 15 μL of Hi-Di™ formamide containing 0.5 μL of GeneScan™ 500 LIZ® Size Standard (Life Technologies). Amplified DNA fragments were separated using a 310 Genetic Analyzer and analyzed with the GeneMapper v4.0 software. Reproducibility was evaluated by estimating the mean of peak areas in 3 independent analyses.

Table 1
Primers used for the miniplex analysis.

Locus	Direction	Sequence	Dye
Penta E	Forward	5′-AGATCACGCCATTGCACTC-3′	FAM
	Reverse	5′-TGGGTTATTAATTGAGAAAACTCCTT-3′	
Penta D	Forward	5′-CATCTCAAGAAAGAAAAAAAAGAAAG-3′	VIC
	Reverse	5′-TGCCTAACCTATGGTCATAACG-3′	
vWA	Forward	5′-TCCCACCTTCCAGAAGAAGA-3′	VIC
	Reverse	5′-GATGATAAATACATAGGATGGATGG-3′	
D10S2325	Forward	5′-CAGCATGAAGCTCACGAAAG-3′	NED
	Reverse	5′-TGGGTGACGAGCCCGT-3′	
D3S1744	Forward	5′-TCTGGCCCCATCTCTCTC-3′	NED
	Reverse	5′-CATTGGTGAATTGGGAAAGG-3′	
D8S1179	Forward	5′-TGTACATTCGTATCTATCTRTCT-3′	PET
	Reverse	5′-GTAGATTATTTTCACTGTGGGGAAT-3′	
D5S818	Forward	5′-TGTGACAAGGGTGATTTTCCT-3′	PET
	Reverse	5′-AGCGCTTTTTAGCCAAGTGA-3′	

3. Results

3.1. Kinship analysis simulation using 36 STR loci

In our previous study, we evaluated the allele frequencies for 37 STR loci in the Japanese population [5]. Therefore, all 37 loci were available to calculate the LR, which is often used as evidence to verify criminal investigations or pedigrees with extended locus sets. We hypothesized that applying all 37 STR loci for forensic DNA identification would improve kinship analysis. Therefore, we simulated kinship analysis using all available data, except the FESFPS locus, which was omitted because of a linkage-disequilibrium between the FESFPS and Penta E loci on chromosome 15. Penta E was chosen over FESFPS for its higher power of discrimination.

Fig. 1 shows the LR distribution for the kinship analysis, as evaluated using Identifiler, MiniFiler, and the 36 STR loci. DNA identification using all 36 STR loci resulted in a higher LR compared with the Identifiler or MiniFiler systems in all 3 cases using parent/child, full-sib, and half-sib as the reference family person. To avoid linkage-disequilibrium, 19 loci from independent chromosomes were chosen on the basis of their high powers of discrimination (Supplemental Table 1) and tested for kinship analysis. The LR distribution was similar to those obtained for the 36 loci shown in Supplemental Fig. 1, suggesting that the extended locus set improves kinship analysis.

However, the multiplex system for 36 STR loci has not been optimized for the analysis of damaged DNA, which is required in forensic cases. Therefore, we developed a new miniplex system that may be used to analyze damaged DNA with a high power of discrimination in combination with MiniFiler.

3.2. Development of a new miniplex system

On the basis of our previous analysis, we considered the following criteria to choose appropriate loci for the development of a novel miniplex system: (1) high power of discrimination (>0.9), (2) narrow range in length, (3) fewer inter-alleles, and (4) existence on independent chromosomes. For example, SE33 showed the highest power of discrimination, but was excluded because it has a wide range of allele variation and many inter-alleles. Finally, 7 STR loci, including D3S1744, D5S818, D8S1179, D10S2325, Penta D, Penta E, and vWA were chosen. The combined matching