Forensic Science International: Genetics 5 (2011) 316–328

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Forensic population genetics—original research

Low template STR typing: Effect of replicate number and consensus method on genotyping reliability and DNA database search results

Corina C.G. Benschop [a], Cornelis P. van der Beek [b], Hugo C. Meiland [c], Ankie G.M. van Gorp [a], Antoinette A. Westen [a], Titia Sijen [a,*]

[a] Netherlands Forensic Institute, Department of Human Biological Traces, P.O. Box 24044, 2490 AA The Hague, The Netherlands
[b] Netherlands Forensic Institute, Department Dutch DNA database, P.O. Box 24044, 2490 AA The Hague, The Netherlands
[c] Leiden Institute of Advanced Computer Science, Leiden University, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 12 March 2010
Received in revised form 25 May 2010
Accepted 24 June 2010

Keywords:
Forensic science
Low template DNA
Low copy number
Consensus
Database searches

ABSTRACT

To analyze DNA samples with very low DNA concentrations, various methods have been developed that sensitize short tandem repeat (STR) typing. Sensitized DNA typing is accompanied by stochastic amplification effects, such as allele drop-outs and drop-ins. Therefore low template (LT) DNA profiles are interpreted with care. One can either try to infer the genotype by a consensus method that uses alleles confirmed in replicate analyses, or one can use a statistical model to evaluate the strength of the evidence in a direct comparison with a known DNA profile. In this study we focused on the first strategy and we show that the procedure by which the consensus profile is assembled will affect genotyping reliability. In order to gain insight in the roles of replicate number and requested level of reproducibility, we generated six independent amplifications of samples of known donors. The LT methods included both increased cycling and enhanced capillary electrophoresis (CE) injection [1]. Consensus profiles were assembled from two to six of the replications using four methods: composite (include all alleles), $n-1$ (include alleles detected in all but one replicate), $n/2$ (include alleles detected in at least half of the replicates) and $2\times$ (include alleles detected twice). We compared the consensus DNA profiles with the DNA profile of the known donor, studied the stochastic amplification effects and examined the effect of the consensus procedure on DNA database search results. From all these analyses we conclude that the accuracy of LT DNA typing and the efficiency of database searching improve when the number of replicates is increased and the consensus method is $n/2$. The most functional number of replicates within this $n/2$ method is four (although a replicate number of three suffices for samples showing >25% of the alleles in standard STR typing). This approach was also the optimal strategy for the analysis of 2-person mixtures, although modified search strategies may be needed to retrieve the minor component in database searches. From the database searches follows the recommendation to specifically mark LT DNA profiles when entering them into the DNA database.

1. Introduction

Sensitized DNA typing (e.g. more amplification cycles or higher capillary electrophoresis injection settings [1]) is used to make partial DNA profiles more complete. Typically, these low template (LT) DNA profiles contain one or more so-called stochastic effects such as allele drop-out, locus drop-out, heterozygote peak imbalance, increased stutter, masking of alleles in mixtures, allele drop-in, near-threshold peaks and show differences between replicates from the same DNA extract [2–8]. To avoid false allele calling of increased stutter peaks, alternative analysis guidelines with raised stutter ratios are used for DNA profiles obtained after increased cycling [4,9–12]. Regarding the interpretation of LT DNA profiles, it is widely recognized that caution is required and that refined and validated protocols are needed to prevent interpretation bias [5,6,8,10,11,13–16].

Basically, two strategies for the interpretation of LT DNA profiles can be followed: the biological model [2–4,10–12,17,18] and the statistical model [2,3,7,19–21]. The biological model tries to obtain the genotype from the LT sample by deducing a consensus DNA profile from replicate analyses. This method is based on all or none allele detection (qualitative data) [22], but

* Corresponding author. Tel.: +31 708886888; fax: +31 708886555.
E-mail addresses: c.benschop@nfi.minjus.nl (Corina C.G. Benschop), k.v.d.beek@nfi.minjus.nl (C.P. van der Beek), hugo@chem.leidenuniv.nl (H.C. Meiland), a.van.gorp@nfi.minjus.nl (Ankie G.M. van Gorp), t.westen@nfi.minjus.nl (A.A. Westen), t.sijen@nfi.minjus.nl (T. Sijen).

1872-4973/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2010.06.006

can be extended to include peak heights [10]. The statistical model aims to assess the probability of the set of replicates from all possible genotypes and uses quantitative data [22]. A likelihood ratio framework is used to evaluate the strength of the scientific evidence in the context of the DNA profile of a known donor or suspect. Statistical models depend on comparison to a known DNA profile and are generally not suited for database searching. Parameters like drop-out probability and peak height or peak area are input for the statistical model. These are complex factors: drop-out probability shows variation for laboratories, machinery within the same laboratory and STR kits used for profiling [23]. Moreover, the inference of peak area is less reliable with lower peak heights due to a greater effect of the baseline noise [3,4]. Although software tools [3,7] have been developed, this probabilistic theory is hardly applied in forensic practice due to its time-consuming complexity. Meanwhile, forensic experts turn to the biological model and report only alleles that are reproduced in consecutive amplifications aiming to exclude spurious alleles. Often, this consensus method fails to fully infer the genotype of the donor, and the consensus profile may contain wrong locus designations like a false homozygote (when only one of the two alleles is reproduced). Incomplete designations may increase the number of adventitious matches when searching in a DNA database [24]. Altogether, the consensus approach for LT DNA profile interpretation is considered to be conservative.

The statistical model is regarded the way forward although it is recognized that big steps need to be taken with respect to both validation of the model and training of scientists, judges and lawyers. We feel that whilst this process is ongoing, it is important to use the most effective consensus procedure; after all, the number of replicates from which the consensus profile is assembled and the level of reproducibility that is called for will affect the accuracy by which a genotype is deduced.

In this study, we analyzed a large set of LT profiles in order to address specific issues raised during LT profile interpretation. These topics include the most frequent positions of allele drop-ins, the correlation between peak height and zygosity at a locus and the relationship between locus drop-outs and allelic drop-outs. In addition, we aimed to establish the optimal consensus procedure. Hence, a set of LT profiles was generated consisting of six independent amplifications of samples of known donors varying in the percentage of alleles detected by standard DNA analysis. DNA typing was sensitized by adding amplification cycles or using enhanced capillary electrophoresis (CE) injection settings [1]. From these amplifications we generated a series of consensus profiles for each sample based on a variable number of replicates (2–6) and using four consensus strategies: (1) composite (add all alleles), (2) $n-1$ (include alleles detected in all but one replicate), (3) $n/2$ (include alleles detected in half of the replicates) and (4) $2\times$ (include alleles detected twice). In our laboratory currently we perform three PCR amplifications and add alleles detected two or three times to the consensus profile. This method can be described as $n-1$, as $n/2$ but also as $2\times$. To gain insight in the effect of the consensus method when more amplifications are performed we tested each of these three consensus approaches. The composite approach was examined as well since it results in the maximum information theoretically. We determined the efficacy by which the consensus strategies detect the alleles of the donor. Peak heights or peak height ratios were not taken into account. Next, we determined the efficiency of retrieving the donor when using consensus profiles in searches against an anonymized copy of the Dutch DNA database to which the profiles of the donors were added. The findings for the single donor samples were substantiated using mixtures of two known donors.

2. Materials and methods

2.1. DNA samples

Twenty-three single donor low template DNA extracts were extracted from samples of different cell types (skin epithelium, vaginal epithelium, saliva and blood) originating from seven volunteers. For mixed low template samples, DNA was extracted from five skin epithelium samples which were obtained by mimicked strangulation (i.e. the hands of donor A rubbed on the arm of donor B). These mixtures contained DNA of two contributors (male–female) with a ratio of approximately 1:2. Six different donors contributed to these 2-person mixtures. Mock casework samples were chosen to include the chance of sporadic contamination as is common for forensic casework samples. Importantly, each mock casework DNA extract contained DNA from donor(s) from whom the STR profile was known.

2.2. STR-profiling

The 28 mock casework LT DNA extracts were amplified using the AMPF*l*STR® Identifiler™ kit (AB) in a volume of 25 μl for 28 cycles (recommended by the manufacturer). The DNA input varied per LT extract and was at most 0.28 ng. Six independent amplifications were generated for each sample (amplifications were done in independent wells of the same amplification plate). The 28 + 6 cycle amplifications were made by transferring 10 μl of each 28 cycling PCR product to a new tube, adding 0.5 μl fresh AmpliTaq Gold Polymerase (AB) and applying the Identifiler protocol for 6 cycles (based on internal validation and [12]). In addition, four different amounts of pristine DNA (DNA007) (10, 15, 20 and 40 pg) were amplified using the AMPF*l*STR® MiniFiler kit in a volume of 25 μl for 30 cycles (recommended by the manufacturer). Thirty independent amplifications were generated for each DNA input. All MiniFiler amplicons are small and range from 70 to 250 bp. We used this MiniFiler setup to gain deeper insight in the LT effects of drop-in and drop-out without the influence of DNA quality or amplicon length.

2.3. Capillary electrophoresis

All DNA fragments were separated and detected by capillary electrophoresis on the same 3130xl ABI Prism® Genetic Analyzer™ (AB). For the 28, 30 and 28 + 6 cycling PCR products, injection settings were 3 kV for 15 s and injection mixtures consisted of 1 μl PCR product, 0.3 μl LIZ™-500 size marker (AB) and 8.7 μl Hi-Di Formamide (AB). For the 28 cycling PCR products the sensitivity of STR typing was increased by performing a second capillary electrophoresis run with injection settings of 9 kV for 15 s preceded by removal of residual dye molecules and salts using Performa® DTR V3 96-Well Short Plates or 1.5 ml columns as described by the manufacturer (Edge BioSystems, Gaithersburg, USA) with minor adjustments as described in Westen et al. [1]. Mixtures for 9 kV injections were made using 2 μl of Performa DTR gel filtration cartridges purified PCR product, 0.015 μl LIZ™-500 and 8.98 μl Hi-Di Formamide. The amount of allelic ladder in runs with 9 kV injection settings was reduced 20-fold. Samples were denatured during 4 min at 95 °C and cooled on ice blocks for 5 min. CE resulted in 414 single donor Identifiler profiles (23 samples amplified in 6-fold and analyzed by 3 settings; 28 cycles + 3 kV, 28 cycles + 9 kV and 28 + 6 cycles + 3 kV), 90 mixed Identifiler profiles (5 samples amplified in 6-fold and analayzed by 3 settings) and 120 MiniFiler profiles (4 concentrations of pristine DNA amplified in 30-fold and analyzed by 1 setting).