Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Forensic Science International: Genetics 2 (2008) 292–300

FSI GENETICS

www.elsevier.com/locate/fsig

Performance of the SNP*for*ID 52 SNP-plex assay in paternity testing

Claus Børsting *, Juan J. Sanchez [1], Hanna E. Hansen, Anders J. Hansen, Hanne Q. Bruun, Niels Morling

Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health Sciences, University of Copenhagen, 11 Frederik V's Vej, DK-2100 Copenhagen, Denmark

Received 5 February 2008; received in revised form 25 March 2008; accepted 27 March 2008

Abstract

The performance of a multiplex assay with 52 autosomal single nucleotide polymorphisms (SNPs) developed for human identification was tested on 124 mother–child–father trios. The typical paternity indices (PIs) were 10^5–10^6 for the trios and 10^3–10^4 for the child–father duos. Using the SNP profiles from the randomly selected trios and 700 previously typed individuals, a total of 83,096 comparisons between mother, child and an unrelated man were performed. On average, 9–10 mismatches per comparison were detected. Four mismatches were genetic inconsistencies and 5–6 mismatches were opposite homozygosities. In only two of the 83,096 comparisons did an unrelated man match perfectly to a mother–child duo, and in both cases the PI of the true father was much higher than the PI of the unrelated man. The trios were also typed for 15 short tandem repeats (STRs) and seven variable number of tandem repeats (VNTRs). The typical PIs based on 15 STRs or seven VNTRs were 5–50 times higher than the typical PIs based on 52 SNPs. Six mutations in tandem repeats were detected among the randomly selected trios. In contrast, there was not found any mutations in the SNP loci. The results showed that the 52 SNP-plex assay is a very useful alternative to currently used methods in relationship testing. The usefulness of SNP markers with low mutation rates in paternity and immigration casework is discussed.
© 2008 Elsevier Ireland Ltd. All rights reserved.

Keywords: SNP; Forensic; Paternity testing; Mutation; PCR multiplex

1. Introduction

The use of single nucleotide polymorphisms (SNPs) in forensic genetic caseworks has been widely discussed in recent years, mainly because SNPs have three important advantages compared to short tandem repeats (STRs): (1) low mutation rates, (2) short amplicons sizes (<100 bp) are achievable, and (3) high throughput genotyping technologies are available. In addition, SNPs may become an important investigative tool for the police when SNP packages for physical characteristics or ethnic origin are fully developed and validated.

The low mutation rate of SNPs (approximately 10^{-8}) is a major advantage in relationship testing. A mutation will often lead to a genotype that is incompatible with Mendel's law of inheritance, and the results may indicate that more than one type of relationship between the tested individuals is possible. This is especially true for immigration cases, where the tested individuals are often related and a mutation may result in comparable likelihoods for different family scenarios. However, a locus with a low mutation rate is also a locus with few alleles, and SNPs are much less polymorphic than STRs. Krawczak [1] calculated the power of paternity exclusion for STRs and SNPs and estimated that approximately 4.2 SNPs with allele frequencies of 0.5 gave the same power of exclusion as one STR locus. Gill [2] showed that variations in allele frequencies between 0.2 and 0.8 had little effect on the probability of exclusion in paternity cases. For 50 SNP loci, the probability of exclusion was more than 0.999. Furthermore, Gill [2] calculated the match probability for SNPs and estimated that 50 SNPs with allele frequencies of 0.2–0.8 would give the same likelihood ratio as 12 STRs. Ayres [3] estimated that analysis of 30–35 SNPs with allele frequencies of 0.3–0.7 would result in one or more mismatches between the child and the tested man in 99.5% of all trios consisting of mother, child and an unrelated man. However, the number of SNP loci required to detect at least one mismatch in 99.5% of all motherless cases was estimated to 80. Ayres [3] also calculated the expected paternity index (PI) and suggested that

* Corresponding author. Tel.: +45 35 32 61 10; fax: +45 35 32 61 20.
E-mail addresses: claus.boersting@forensic.ku.dk (C. Børsting), jj.sanchez@mju.es (J.J. Sanchez).
[1] Current address: Instituto Nacional de Toxicología y Ciencias Forenses, Campus de Ciencias de la Salud, 38320 La Laguna, Tenerife, Spain.

1872-4973/$ – see front matter © 2008 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2008.03.007

50–60 SNPs with allele frequencies of 0.3–0.7 should be used for trio cases and 70–80 SNPs should be used for motherless cases.

Several autosomal SNP marker sets developed for the identification of humans were reported recently [4–10]. The largest of these marker sets and the only with more than 50 SNPs was the 52 SNP-plex assay developed by the SNP*for*ID consortium [8]. The main criteria for selection of the SNP loci in the SNP*for*ID multiplex were high levels of heterogeneity in the three major population groups and a minimum distance of 100 kb between the SNPs and neighbouring genes. All 52 SNP loci were amplified in one PCR reaction and the SNP alleles were detected in two single base extension (SBE) reactions by capillary electrophoresis. The two SBE reactions were injected into the same capillary with 5 min intervals and analysed in one data file. The SNPs were shown to be polymorphic in nine different populations from four different continents and there was not found any evidence of linkage disequilibrium for any pair of SNPs in the multiplex [8].

In this work, we typed 124 true mother–child–father trios from two genetically different populations (Danes and Somalis) for 15 STRs, seven variable number of tandem repeats (VNTRs), and 52 SNPs. The purpose of the study was two-fold: (1) to compare the performance of the SNP*for*ID 52 SNP-plex assay with well known STR and VNTR marker sets in typical paternity testing casework, and (2) to look for silent alleles in the selected SNP loci.

2. Material and methods

2.1. Samples

A total of 73 paternity cases from 2004 to 2005 and 54 immigration cases from 1997 to 2002 were selected (Table 1). The people involved in the paternity cases were of Danish origin and the people involved in the immigration cases were of Somali origin. Under the casework investigation, the reported PI in the paternity cases were calculated based on 15 STRs, whereas the paternity and maternity indices (MI) calculated in the immigration cases were based on seven VNTRs. For 50 of the paternity cases and all the immigration cases, PI >10,000 and MI >10,000, and there was not observed any mismatch between the investigated woman and the child, or between the investigated man and the child. In the remaining 23 paternity cases, one or two mismatches were observed between the investigated man and the child. The work was approved by the Danish ethical committee (KF-01-037/03).

2.2. STR analysis

A total of 15 autosomal STRs (CSF1PO, D13S317, D16S539, D18S51, D19S433, D21S11, D2S1338, D3S1358, D5S818, D7S820, D8S1179, FGA, TH01, TPOX and vWA) were typed using the AmpF*l*STR Identifiler PCR Amplification Kit (AB). A total of 0.5 μl chelex purified DNA was added to a mixture of 4 μl AmpF*l*STR PCR reaction mix (AB), 2 μl AmpF*l*STR Identifiler primer set (AB), 0.2 μl AmpliTaq Gold (AB) and 3.5 μl dH_2O. The following cycle programme was used: denaturation and enzyme activation at 95 °C for 11 min, 28 cycles of 94 °C for 1 min, 59 °C for 1 min and 72 °C for 1 min, followed by 60 min at 60 °C. PCR reactions were performed in the AB9700 thermal cycler (AB). One μl PCR product was mixed with 15 μl Hi-Di formamide (AB) and 0.1 μl GeneScan-400 HD LIZ size standard (AB), and analysed on an AB Prism 3100 Genetic Analyzer (AB) with a 36 cm capillary array, POP-4 polymer (AB) and 6 s injections at 3000 V. Data were analysed independently by two technicians using GeneScan analysis software v. 3.7 and GenoTyper analysis software v. 3.7 (AB), and the results were compared. The minimum peak height was set to 100 RFU for all dyes. All samples were analysed at least twice.

2.3. VNTR analysis

Seven VNTRs (D5S43, D7S21, D7S22, D12S11, D2S44, D16S309 and D5S110) were typed by RFLP techniques using a modified version of the methods described by Morling and Hansen [11]. Genomic DNA was isolated from 300 μl whole blood using the FlexiGene DNA kit (Qiagen) and digested with 40 U HinfI (Roche) overnight at 37 °C. Two micrograms digested DNA were fractionated by electrophoresis overnight through 20 cm 0.7% agarose gels using a GNA 200 electrophoresis instrument (Pharmacia Biotech). A mixture of 4 μl 117.5 μg/μl Low DNA mass ladder (Invitrogen) and 13 μl Gelstar nucleic acid gel stain (FMC BioProducts) was used as visual size marker. Electrophoresis was terminated when the 2000 bp fragment of the Low DNA mass ladder had migrated 15.5 cm through the agarose gel. After electrophoresis, the gels were washed in 0.25 M HCl for 30 min, in 1.5 M NaCl, 0.5 M NaOH for 30 min, and in 3 M NaCl, 0.5 M Tris–HCl, pH 7.5 for 30 min. DNA was blotted on to positively charged nylon membranes (Roche) using a Vacu-aid pressure blotter (Hybaid) and crosslinked to the membranes by UV irradiation (1200 μJ) using a UV Stratalinker 1800 (Stratagene). The membranes were pre-hybridised for 30 min at 50 °C

Table 1
Trio selection

	50 paternity cases	23 paternity cases	54 immigration cases
Ethnic origin	Danish	Danish	Somali
Casework investigation	15 STRs	15 STRs	Seven VNTRs
	No mismatches	1 or 2 mismatches	No mismatches
Additional analyses	52 SNPs	52 SNPs	52 SNPs
	Seven VNTRs	Seven VNTRs	15 STRs

in a mixture of 22.5 ml 10% SDS, 0.5 M Na_2HPO_4, pH 7.2 and 2.5 ml 100 g/L Casein Bovin Milk (Merck), $C_4H_4O_4$, pH 7.5. Hybridisation was performed at 50 °C for 30 min by addition of 10 μl alkaline phosphatase linked VNTR probe (MS8, MS31, MS43, MS205, MS621, g3 or pYNH24 (Cellmark Diagnostics)) and 10 μl alkaline phosphatase linked NICE MW100 probe (Orchid Cellmark Diagnostics). After hybridisation, the membranes were washed twice in 100 ml 10% SDS, 0.5 M Na_2HPO_4, pH 7.2 at 50 °C for 10 min, twice in 50 ml 100 g/L Casein Bovin Milk (Merck), $C_4H_4O_4$, pH 7.5 at room temperature for 10 min, and once in 25 ml 0.1 M $MgCl_2$, 1% $C_4H_{11}NO_2$, pH 9.5 at room temperature for 10 min. Finally, the membranes were washed in 30 ml 0.1 M $MgCl_2$, 1% $C_4H_{11}NO_2$, pH 9.5 containing 300 μl CDP-Star (Orchid Cellmark Diagnostics) at room temperature for 10 min. Prehybridisation, hybridisation and wash of the membranes were performed in a Midi oven (Hybaid). CDP-Star is a substrate for alkaline phosphatase and emits light at 477 nm. After exposure, X-ray films were analysed independently by two technicians using a KC3300 digitizer (Graphtec) and DNAVIEW 27.21 [11]. The NICE DNA analysis ladder (Orchid Cellmark diagnostics) was used as size marker. Probes were removed from the membranes by washing the membranes once in 10% SDS at 88 °C for 15 min.

2.4. SNP analysis

A total of 52 autosomal SNPs were amplified by PCR in 25 μl reactions containing 1–10 ng genomic DNA, 1× PCR buffer, 8 mM $MgCl_2$, 700 μM of each dNTP, 0.01–0.17 μM of each primer, and 2 U AmpliTaq Gold DNA polymerase (AB) [8]. The following cycle programme was used: denaturation at 94 °C for 5 min followed by 35 cycles of 95 °C for 30 s, 60 °C for 30 s and 65 °C for 30 s followed by 7 min at 65 °C. PCR reactions were performed in the GeneAmp 9600 (PerkinElmer) or Mastercycler gradient (Eppendorf) thermal cycler. Excess of PCR primers and dNTPs were removed by the addition of 0.75 μl (1 U/μl) shrimp alkaline phosphatase (Amersham Pharmacia Biotech) and 0.023 μl (10 U/μl) Exonuclease I (Amersham Pharmacia Biotech) to 2.5 μl PCR product and incubation at 37 °C for 1 h followed by incubation at 75 °C for 15 min.

SBE reactions were performed in 8 μl reaction volumes containing 1 μl purified PCR product, 4 μl SNaPshot reaction mix (AB), 1 μl SBE primer mix (0.01–0.27 μM of each primer) and 2 μl Milli-Q water. The SBE primer mix was diluted in 160 mM ammonium sulphate (Sigma–Aldrich) to minimize primer–dimer artefacts. The SBE reaction was performed in GeneAmp 9600 (PerkinElmer) or Mastercycler gradient (Eppendorf) thermal cyclers with 30 cycles of 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 30 s. Excess nucleotides were removed by addition of 1 μl (1 U/μl) shrimp alkaline phosphatase to the SBE mix and incubation at 37 °C for 45 min followed by 75 °C for 15 min. Two microlitres SBE products were mixed with 20 μl Hi-Di formamide (AB) and 0.1 μl GeneScan-120 Liz internal size standard (AB). The SBE products were analysed by capillary electrophoresis using AB Prism 3100 or AB Prism 3130XL Genetic Analyzers (AB) with 36 cm capillary arrays and POP-4 polymer (AB) as previously described [8,12]. Analysis was made using GeneScan Analysis 3.7 with peak thresholds set to a minimum of 120 relative fluorescence units (RFUs) (blue colour), 60 RFUs (green colour), and 30 RFUs (yellow, red, and orange colour). Allele calls were made independently by two technicians using Genotyper 3.7 macros (www.snpforid.org) and the results were compared. All peaks in the size standard had to be detected and the peak height of the largest peak in the electropherogram had to be a minimum of 1000 RFUs before analysis could proceed. Peaks detected in pre-defined allele windows with peak heights larger than 10% (blue colour), 7% (green colour) or 5% (yellow and red colour) of the maximum peak height in the respective colour were labelled with allele name, peak height and size [8,12]. A maximum peak height ratio of 3:1 after normalization was accepted as a heterozygote. A minimum peak height of 400 RFUs (blue), 200 RFUs (green) or 100 RFUs (yellow and red), and a peak height ratio of minimum 5:1 after normalization were accepted as a homozygote. A peak height ratio between 3:1 and 5:1 after normalization was considered inconclusive. All samples were analysed at least twice and the results from the two experiments were compared. If the results differed the sample was analysed a third time. Approximately 20% of the samples typed with the 52 SNP-plex assay in this work and in [8] were analysed a third time.

2.5. DNA sequencing

Amplification products of rs907100 and rs354439 were sequenced using the BigDye Terminator Kit (AB) as recommended by the manufacturer and analysed on an ABI Prism 377 Genetic Analyzer (AB).

2.6. Calculation of paternity indices

PI was calculated using DNAVIEW 27.21 and in-house frequency databases for Danes and Somalis. STR allele frequencies will be published elsewhere (Simonsen and Hallenberg, in preparation), and VNTR allele binns are shown in Tables S1 and S2. SNP allele frequencies were previously published [8]. Calculations were performed under the assumption that silent alleles exist. For loci where silent alleles had not previously been observed, the frequency of silent alleles was set to $1/(n + 1)$ (n is the total number of observed alleles).

The selected SNPs for the SNP*for*ID 52 SNP-plex assay were positioned on the distal parts of the chromosome arms except for the only SNP selected on chromosome 19. Two SNPs were selected on most chromosomes, but three or four SNPs were selected on chromosome 1, 9, 10, 13, 15, 21, and 22 [8]. On the smallest chromosomes, e.g. 21 and 22, the average number of recombination events per meiosis is one [13,14], and consequently, two of the three selected SNPs on these chromosomes will be on the same side of the recombination event, and the child will inherit two linked loci and one independent locus from the parent. However, PI may be

calculated as if the loci were independent, because the loci are in linkage equilibrium in the population [8] and because the haplotypes of the parents are unknown in a typical paternity case.

2.7. Determination of mismatches between a child and an unrelated man

Mismatches were divided into two groups: (1) genetic inconsistency and (2) opposite homozygosity. For a bi-allelic SNP locus, a genetic inconsistency is only observed when (1) the child is heterozygous, and (2) the mother and the tested man are homozygous for the same allele. Opposite homozygosity is observed when (1) the child and the tested man are homozygous, and (2) the child and the tested man have different alleles.

An Excel spreadsheet was constructed, where the SNP profiles of the mother and the child (or the child alone) could be imported and matched against a large number of possible fathers. For each tested man, the number of genetic inconsistencies and opposite homozygosities was calculated using the simple rules above. SNP profiles from 50 Danish mother–child duos and 54 Somali mother–child duos were matched against the SNP profiles from 800 database samples from nine different populations (this work and [8]). The true fathers of the children were also included in the database and were used as positive controls.

3. Results

3.1. PI calculations in true mother–child–father trios and child–father duos

Fifty Danish and 54 Somali mother–child–father trios were selected based on the results of forensic genetic investigations in paternity and immigration cases.

The 104 trios were typed for 15 STRs, seven VNTRs, and 52 SNPs, and the PIs were calculated for each set of loci (Fig. 1A and B) using population specific frequency databases. In addition, PIs were calculated for the 104 child–father duos (Fig. 1C and D) to simulate motherless casework investigations.

The typical PI for trios and duos was calculated as the geometric means of the obtained PIs (Tables 2A and 2B). Overall, the typical PIs based on 15 STRs and seven VNTRs were similar (10^7 for trios and 10^5 for duos), whereas the typical PIs based on 52 SNPs were 5–50 times lower. The typical PI obtained with the 52 SNP-plex assay in the Danish duos and trios were higher than the typical PI in the Somali duos and trios. This was in accordance with the previously performed estimations of PI based on the SNP allele frequencies found in Europeans, Somalis and Asians [8].

3.2. Trios with mutations

During the initial casework investigation, the 50 Danish trios were typed for 15 STRs and the 54 Somali trios were typed for seven VNTRs. The trios were selected because there was not observed any mismatch in any of the investigated loci during the casework investigation (Table 1). However, when the Danish trios were investigated for the seven VNTRs, one mismatch between the father and the child was observed in two trios (both in D7S22), and when the Somali trios were investigated for the 15 STRs, one mismatch between the father and the child was observed in two trios (in D8S1179 and in vWA31), and one mismatch between the mother and the child or between the father and the child was observed in two other trios (in D3S1358 and in D5S818). In contrast, there was not found any mismatch in the 52 SNP loci in any of the 104 trios.

Based on the mutation rates of the 15 STR loci and the seven VNTR loci, one mutation between the father and the child was expected per 38 and 52 trios, respectively. Similarly, one

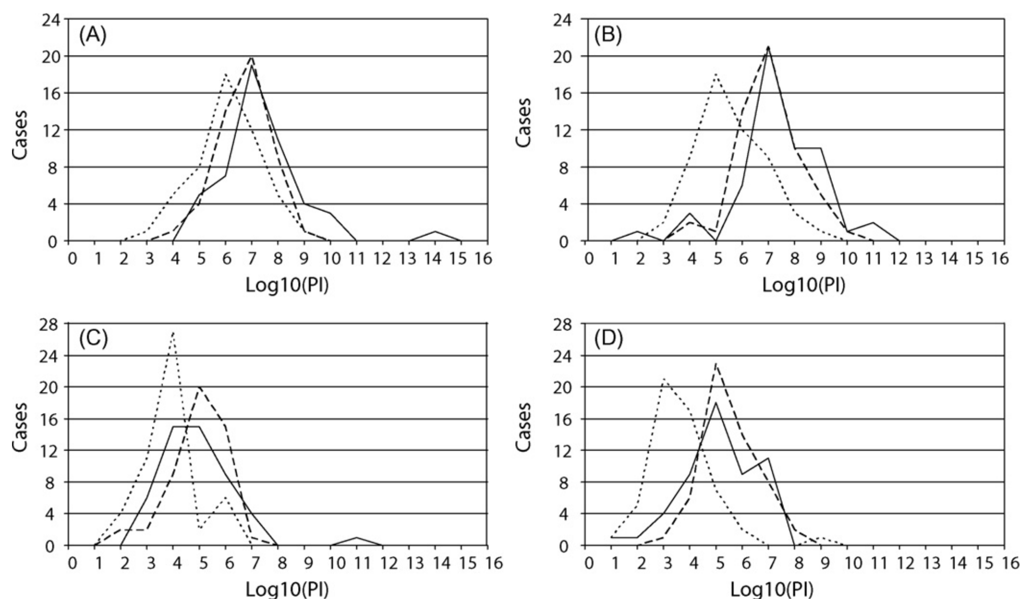

Fig. 1. Distribution of Log_{10}(PI). A total of 50 randomly selected Danish trios (A) and duos (C) and 54 randomly selected Somali trios (B) and duos (D) were typed for 15 STRs (full line), 52 SNPs (dotted line), and seven VNTRs (dashed line).

Table 2A
Typical paternity indices for Danish trios and duos after analysis of 15 STRs, 52 SNPs, and seven VNTRs

	Danish trio			Danish duo		
	STR	SNP	VNTR	STR	SNP	VNTR
Typical PI (geometric mean)	2.53E+07	1.32E+06	5.50E+06	9.75E+04	9.77E+03	9.65E+04
Minimum PI	9.54E+04	489	3.22E+03	708	90	66
Maximum PI	2.18E+14	6.31E+08	9.93E+08	2.50E+11	1.62E+06	1.55E+07

Table 2B
Typical paternity indices for Somali trios and duos after analysis of 15 STRs, 52 SNPs, and seven VNTRs

	Somali trio			Somali duo		
	STR	SNP	VNTR	STR	SNP	VNTR
Typical PI (geometric mean)	2.09E+07	4.04E+05	1.26E+07	1.04E+05	4.31E+03	3.41E+05
Minimum PI	202	1.68E+03	1.30E+04	1	12	467
Maximum PI	4.21E+10	1.30E+09	1.20E+10	2.73E+07	1.01E+09	9.23E+07

mutation between the mother and the child was expected in every 164 and 121 trios, respectively. The reason for the unexpected high number of mutations in the 104 trios is unknown. The VNTR results in the relevant trios were re-examined, and all the trios with mutations in one of the STR loci were typed a third time with the AmpF*l*STR Identifiler PCR Amplification Kit to confirm the results.

In addition to the 104 trios mentioned above, 23 Danish paternity cases from 2004 to 2005 were selected (Table 1). During the casework investigation, the trios were typed for 15 STRs. In 20 of these cases, one mismatch was observed between the tested man and the child, and in three cases, two mismatches were observed between the tested man and the child. The individuals in these cases were subsequently typed for seven VNTRs and 52 SNPs. In the 20 cases with one mismatch, no additional mismatch was found, indicating that the observed mismatch in each case was a mutation. The PI was calculated based on seven VNTRs and 52 SNPs, respectively, and the obtained PIs were similar to the PIs in Fig. 1 (data not shown). In the three cases with two mismatches, additional mismatches were found in 5, 5 and 7 of the VNTR loci and in 7, 7 and 2 of the SNP loci, respectively, demonstrating that the tested man in the three cases was not the father of the tested child.

3.3. Silent alleles and new alleles in the 52 SNP-plex assay

One of the purposes of typing trios with the 52 SNP-plex assay was to look for possible silent alleles, and early on in this work, a frequent silent allele was found in rs907100. In 13 of the 50 Danish trios, opposite homozygosity was observed between the mother and the child, or between the father and the child. Interestingly, the genotype distributions of rs907100 differed from expectations based on Hardy–Weinberg equilibrium in six of the nine previously tested populations because of an excess of individuals typed as homozygotes [8], which strongly indicated the presence of a silent allele. DNA from selected samples was amplified using another set of PCR primers for rs907100, and the PCR products were sequenced. An A/G SNP (rs11689319) in position −9 of the reverse PCR primer was detected (data not shown), and the A allele of this SNP was in all sequences linked to the G allele of rs907100. The PCR primer used in the 52 SNP-plex assay had a C nucleotide in this position, and when the rs11689319 A allele was present, only a very weak signal was detected from the rs907100 G allele.

The reverse PCR primer for rs907100 was redesigned with a degenerated nucleotide in position −9. With the new PCR primer, a strong signal for the rs907100 G allele was obtained also in the presence of the rs11689319 A allele. This primer was used in all subsequent work, and it was also included in the PCR primer mixes used for the EDNAP SNP exercise in 2006 [12].

There was not observed mutations in any of the 52 SNP loci in the 124 true trios analysed in this work. However, a new allele was detected for rs354439. This SNP was reported as an A/T SNP, but in one of the Danish trios, a G allele was detected in the mother and the child. The PCR products were sequenced and the G allele was confirmed (data not shown).

3.4. Number of mismatches for an unrelated man in trios

Mismatches were divided in two groups: (1) genetic inconsistency and (2) opposite homozygosity. For a bi-allelic SNP locus, a genetic inconsistency is only observed when the child is heterozygous, and the mother and the tested man are homozygous for the same allele. Opposite homozygosity is observed when the child and the tested man are homozygous for different alleles, irrespective of the genotype of the mother. In the case of opposite homozygosity, the child and the tested man may share a silent allele, and thus, opposite homozygosity is not *per se* a true mismatch. Nevertheless, opposite homozygosity speaks strongly against paternity ($PI \ll 1$).

SNP profiles from 50 Danish mother–child duos and 54 Somali mother–child duos were matched against the SNP profiles from 800 putative fathers from nine different

populations (83,200 comparisons). The real fathers of the 104 children were among the 800 putative fathers and, as expected, the real fathers matched perfectly to the mother–child duos. When the putative father was not the real father, the putative father matched perfectly to a mother–child duo in only two out of the 83,096 comparisons. A Turkish man matched perfectly to a Somali mother–child duo and a Greenlander matched perfectly to a Danish mother–child duo. The PIs were 660 and 1520 in the two cases, whereas the PIs of the true fathers were 370,000 and 95,900, respectively. The average number of mismatches for an unrelated man was approximately ten. On average, six of the 10 mismatches were opposite homozygosities and four of the 10 mismatches were genetic inconsistencies (Fig. 2). When the putative father had the same ethnic origin as the mother and the child, the average number of loci with opposite homozygosity was five and the average number of mismatches was nine (Table 3). In trios consisting of mother, child and an unrelated man, one and two mismatches were observed in 20 (0.02%) and 107 (0.13%) of the 83,096 comparisons, respectively.

3.5. Number of mismatches for an unrelated man in duos

Genetic inconsistencies in bi-allelic SNP loci cannot be determined with certainty in cases with only one putative parent (motherless or fatherless cases). Even though the child and the putative parent do not share any detectable allele (opposite homozygosity), there is always the possibility that they share a silent allele. The number of loci with opposite homozygosity calculated above (Fig. 2B and E, and Table 3) also applies to cases with only one putative parent, because the genotype of the other parent is irrelevant for the comparison. In duos consisting of a child and an unrelated individual, a perfect match was observed in 154 (0.19%) of the 83,096 comparisons, and there was observed two or fewer opposite homozygosities in 3673 (4.4%) of the 83,096 comparisons. Eighty (52%) of the 154 perfect matches between a child and an unrelated individual were a comparison between a child and an individual of the same ethnic origin. The PIs for the false associations were calculated (Fig. 3) using either Danish or Somali allele frequencies depending on the ethnic origin of the child. The

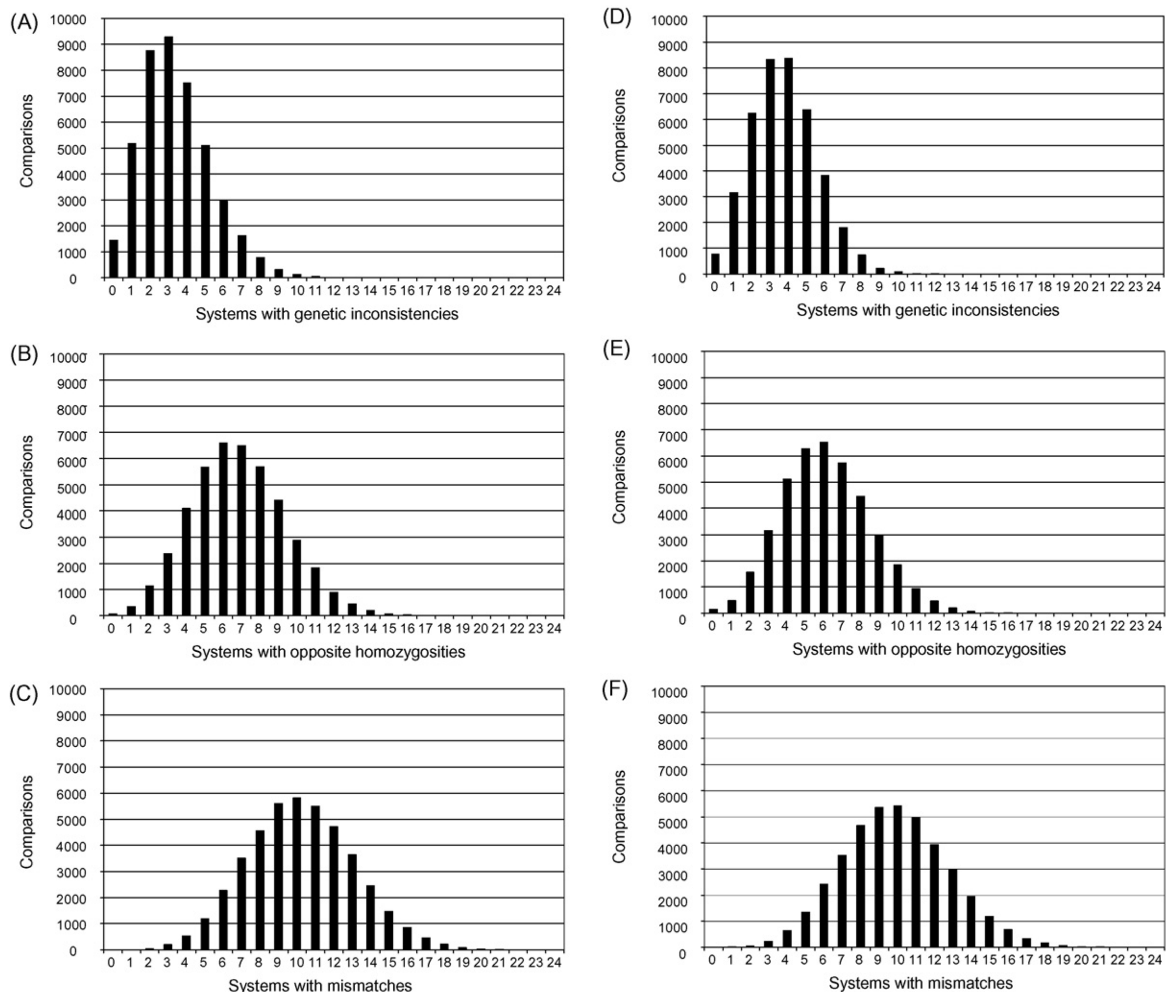

Fig. 2. The number of mismatches between child and tested man in trios consisting of mother, child and an unrelated man. A total of 50 randomly selected Danish mother–child duos (A–C) or 54 randomly selected Somali mother–child duos (D–F) were matched against 800 individuals. The number of genetic inconsistencies (A and D), opposite homozygosities (B and E), and the total number of mismatches (C and F) were determined. The mother–child duos were matched against the real fathers and no mismatches were found. However, only the results from matches between mother, child and an unrelated man are shown in the figure. The number of mismatches between child and tested man in duos consisting of child and an unrelated man are identical to the number of opposite homozygosities (B and E).

Table 3
Average number of mismatches in trios consisting of mother, child and an unrelated man

		Danes (256)	Somalis (104)	Greenlanders (149)	Turks (96)	Chinese (106)
Somali mother–child	Mismatches	10.3	8.7	10.7	9.9	10.8
	Genetic inconsistencies	3.1	3.6	3.6	3.2	3.6
	Opposite homozygosities	7.2	5.1	7.1	6.7	7.2
Danish mother–child	Mismatches	9.0	10.7	10.5	9.3	10.9
	Genetic inconsistencies	3.7	3.7	3.9	3.6	3.9
	Opposite homozygosities	5.3	7.0	6.6	5.7	7.0

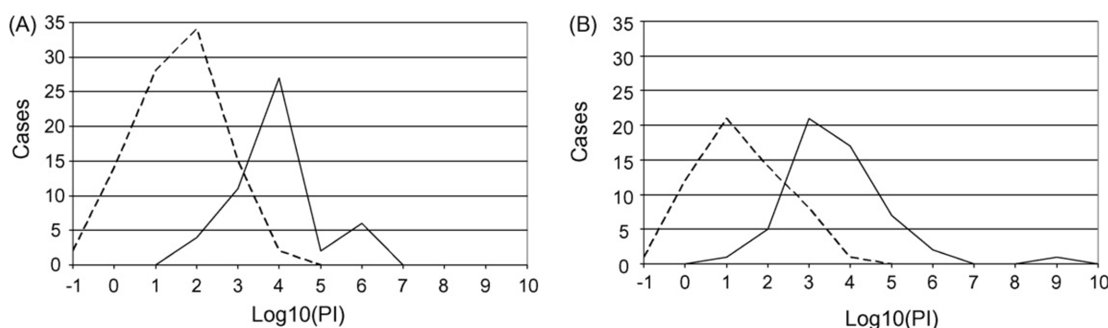

Fig. 3. Distribution of Log_{10}(PI) for false associations (dashed line) and true fathers (full line). Out of 83,096 comparisons between a child and an unrelated man, 95 false associations between a Danish child and an unrelated man (A) and 59 false associations between a Somali child and an unrelated man (B) were found.

typical PIs for duos consisting of a child and an unrelated man were 37 and 22 for a child of Danish and Somali origin, respectively, which was approximately 200 times lower than the typical PIs for the true fathers (Tables 2A and 2B).

4. Discussion

This study demonstrates that the SNP*for*ID 52 SNP-plex assay is a good alternative to standard STR or VNTR investigations in relationship testing. A total of 124 trios were investigated and the typical PIs based on the 52 SNP-plex assay were 10^5–10^6 for true fathers. Three or more mismatches were found in 99.85% of the 83,096 comparisons between mother, child and an unrelated man. In average, 9–10 mismatches per comparison were detected, and in only two of the 83,096 comparisons did an unrelated man match perfectly to a mother–child duo.

The typical PIs based on the 52 SNP-plex assay were 5–50 times lower than the typical PIs obtained when the same trios were typed for 15 STRs or seven VNTRs. However, the high heterogeneity of tandem repeats that result in high PIs comes with a price. Mutations in the most commonly used STR loci are frequently observed. From 1999 to 2003, our laboratory used the AmpFℓSTR® Profiler Plus® PCR Amplification Kit for paternity caseworks. Nine STRs were amplified with this kit and 56 mutations in 4191 true mother–child–father trios (1.3%) were observed. In 2004, the AmpFℓSTR® Profiler Plus® PCR Amplification Kit was replaced with the AmpFℓSTR® Identifiler® PCR Amplification Kit that amplified 15 STRs, and for a period of 2 years, 32 mutations in 1244 true trios (2.6%) were observed. In the current work, six mutations in tandem repeat loci (four in STRs and two in VNTRs) were detected among the 104 randomly selected trios. This is an unusual high number, but it emphasizes the weakness of tandem repeat loci in relationship testing. When one mismatch is observed, e.g. between a child and a tested man, it is usually a mutation. However, in 14 trios investigated in our laboratory (13 times with the Profiler Plus® PCR Amplification Kit and once with the Identifiler® PCR Amplification Kit), the tested man was later excluded as a father by supplementary investigation of seven VNTR loci. Some forensic laboratories choose to ignore single mismatches and calculate PI without the locus with the single mismatch. Other laboratories include the information from the locus with the single mismatch in the calculation by multiplying the PI with the mutation rate of that locus, but do not make supplementary investigations. There is a risk of reaching the wrong conclusion with either of these laboratory protocols. One mismatch is usually a mutation, but it may indicate that a close relative is the true parent, and supplementary investigations are recommendable in such cases. Many forensic laboratories choose to type additional STRs, or if applicable, Y-STRs, but by typing more tandem repeats, there is always a risk of finding more mutations. Our laboratory experienced three paternity cases in the past 2 years, where one mismatch between the child and the tested man was detected after analysis of 15 STRs, and a second mismatch was found after supplementary investigation of seven VNTR loci. The PI was >10,000 in all three cases, but the likelihood ratio of the tested man versus the tested mans brother was between 0.8 and 1.6. Thus, the tested man might as well have been the father as the uncle. Clearly, loci with low mutation rates, like SNPs, would be more appropriate for supplementary investigations and with the introduction of the SNP*for*ID 52 SNP-plex assay, the forensic community has the opportunity to work with a SNP

marker set that may even solve the question of parenthood in most paternity cases, without the need for typing tandem repeat markers.

In the three cases with two mismatched mentioned above, there was no mismatch between the tested man and the child when the individuals were typed with the SNP*for*ID 52 SNP-plex assay (data not shown), and the likelihood ratios of the tested man versus the tested man's brother were 19, 241, and 332 based on SNPs alone.

With an expected mutation rate of 10^{-8} for each of the 52 SNP loci in the SNP*for*ID assay, a mutation is expected once per 1,000,000 trios, and therefore, it was not surprising that there was not found any mutations in the 124 trios typed in this work. Nevertheless, a previously unknown allele was detected for the SNP locus rs354439, and a frequent silent allele was detected in rs907100. A new reverse PCR primer was designed for rs907100 to eliminate the silent allele and this primer was used in all subsequent work, including the EDNAP SNP exercise [12]. It is likely that more silent alleles will be uncovered as more trios are typed with the 52 SNP-plex assay, but as demonstrated here, at least some of them may be eliminated in a simple way.

It was important to evaluate how the SNP*for*ID 52 SNP-plex assay performed in single parent analyses (motherless or fatherless cases), because it may be difficult to exclude a putative parent when the only possible type of mismatch is opposite homozygosity where the putative parent and the child do not share any detectable allele but may share a silent allele. The typical PIs in true duos were 10^3–10^4 based on the 52 SNP-plex assay. Three or more opposite homozygosities were found in 95.6% of the 83,096 comparisons of a child and an unrelated individual, and on average, opposite homozygosity was found in 5–6 loci. A total of 154 perfect matches between a child and an unrelated individual were found, but the typical PI for the false associations was only 30.

In a locus with opposite homozygosity, PI $\sim f_o/f_p{*}f_q$, where f_p and f_q are the frequencies of the observed alleles, and f_o is the frequency of the silent allele. The allele frequencies of the selected SNPs in the SNP*for*ID 52 SNP-plex assay ranged from 0.2 to 0.8 [8]. Thus, PI in a locus with opposite homozygosity will range from $4f_o$ to $6.25f_o$. The frequency of silent alleles is not known for any of the 52 SNP loci, but a very conservative estimate would be $f_o = 0.001$. Consequently, the PI in one and two loci with opposite homozygosity will be maximally 6.3×10^{-3} and 3.9×10^{-5}, respectively, and with a typical PI of 30 for the false associations between a child and an unrelated man, the typical PI for all 52 loci is expected to be less than one in single parent analyses, where the child and the putative parent are opposite homozygotes in one or two loci.

In a recent study, 336 children and 348 men from northern Germany were typed for 15 STRs using the AmpFℓSTR® Identifiler® PCR Amplification Kit [15]. A total of 116,004 comparisons of a child and an unrelated man were performed, and three or more mismatches were detected in 98.5% of the comparisons [15]. In our work, the typical PI for the 124 motherless cases was approximately 10 times higher for the STR investigation than for the SNP investigation. Together, these numbers seem to imply that investigation of 15 STRs is more efficient than typing 52 SNPs in single parent analyses. However, one mutation in one of the STR loci will have a large impact on these cases, because fewer mismatches are expected for an unrelated man compared to the mismatches expected in cases involving mother and child. Therefore, the low mutation rate of SNPs fully compensates for the lower PI. It was suggested that more tandem repeats should be available for typing in cases with only one possible parent [15], but, as discussed above, this also increases the risk of finding more mutations. Ayres [3] suggested that 70–80 SNPs were used for motherless cases and a multiplex of this size would almost certainly be able to solve all cases of this kind. However, we would like to suggest that the strength of STRs and SNPs are combined, and since the SNP*for*ID multiplex was accredited under the ISO 17025 standard in our laboratory in the fall 2007 (Børsting, in preparation), we have analysed 15 STRs and 52 SNPs in all motherless or fatherless cases.

The cost per sample for the 52 SNP-plex assay was approximately 8.5 Euro in laboratory expenses (enzymes, buffers, oligonucleotides, disposables, kits and cost of electrophoresis), which was only 1.6 Euros more than the cost for the investigation of 15 STRs and 48.2 Euros cheaper than the VNTR investigation. The SBE reaction step adds a couple of hours to the time it takes to perform a PCR fragment analysis, but a SNP investigation from sample preparation to analysis of the electropherogram may be conducted in only two work days. In contrast, a VNTR investigation may take weeks to perform.

The SNP*for*ID 52 SNP-plex assay is not a commercial kit, and therefore the 156 oligonucleotides needed for the assay must be bought separately, and PCR and SBE primer mixtures must be made in house. All the equipment that is needed to perform the assay is a thermal cycler and a capillary electrophoresis instrument, which is standard equipment in most forensic genetic laboratories.

The SNP*for*ID multiplex was recently set up as an oligo ligation assay for the Genplex platform in collaboration between the SNP*for*ID consortium and Applied Biosystems [16]. This platform has some clear advantages compared to the SNaPshot assay in both data analysis and data management, which makes the Genplex assay suitable for high throughput analyses of SNPs. Our laboratory is currently developing a SBE multiplex for the MALDI-TOF mass spectrometer that target 50 of the SNPs in the SNP*for*ID 52-plex, and other detection methods have also been exploited [17–19]. It is difficult to predict exactly which method will become the preferred method for SNP typing in forensic laboratories, but it is clear from the work presented here that the SNP*for*ID 52 SNP marker set is a valuable alternative to and in some cases preferable to tandem repeat markers.

Acknowledgements

We thank Charles Brenner, Ph.D. for helpful discussions and Annemette H. Birk, Winnie Ruben, and Pia C. Roswell for technical assistance. The work was supported by grants to J.J. Sanchez from Ellen and Aage Andersen's Foundation, and by

the European Commission (GROWTH programme, SNP*for*ID project, contract G6RD-CT-2002-00844).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2008.03.007.

References

[1] M. Krawczak, Informativity assessment for biallelic single nucleotide polymorphisms, Electrophoresis 20 (1999) 1676–1681.

[2] P. Gill, An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes, Int. J. Legal Med. 114 (2001) 204–210.

[3] K.L. Ayres, The expected performance of single nucleotide polymorphism loci in paternity testing, Forensic Sci. Int. 154 (2005) 167–172.

[4] S. Inagaki, Y. Yamamoto, Y. Doi, T. Takata, T. Ishikawa, K. Imabayashi, K. Yoshitome, S. Miyaishi, H. Ishizu, A new 39-plex analysis method for SNPs including 15 blood group loci, Forensic Sci. Int. 144 (2004) 45–57.

[5] E. Petkovski, C. Keyser-Tracqui, R. Hienne, B. Ludes, SNPs and MALDI-TOF MS: tools for DNA typing in forensic paternity testing and anthropology, J. Forensic Sci. 50 (2005) 1–7.

[6] H.Y. Lee, M.J. Park, J. Yoo, U. Chung, G. Han, K. Shin, Selection of twenty-four highly informative SNP markers for human identification and paternity analysis in Koreans, Forensic Sci. Int. 148 (2005) 107–112.

[7] L.A. Dixon, C.M. Murray, E.J. Archer, A.E. Dobbins, P. Koumi, P. Gill, Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes, Forensic Sci. Int. 154 (2005) 62–77.

[8] J.J. Sanchez, C. Phillips, C. Børsting, K. Balogh, M. Bogus, M. Fondevila, C.D. Harrison, E. Musgrave-Brown, A. Salas, D. Syndercombe-Court, P.M. Schneider, A. Carracedo, N. Morling, A multiplex assay with 52 single nucleotide polymorphisms for human identification, Electrophoresis 27 (2006) 1713–1724.

[9] L. Li, C. Li, R. Li, Y. Liu, Y. Lin, T. Que, M. Sun, Y. Li, SNP genotyping by multiplex amplification and microarrays assay for forensic identification, Forensic Sci. Int. 162 (2006) 74–79.

[10] A.J. Pakstis, W.C. Speed, J.R. Kidd, K.K. Kidd, Candidate SNPs for a universal individual identification panel, Hum. Genet. 121 (2007) 305–317.

[11] N. Morling, H.E. Hansen, Paternity testing with VNTR DNA systems, Int. J. Legal Med. 105 (1993) 189–196.

[12] J.J. Sanchez, C. Børsting, K. Balogh, B. Berger, M. Bogus, J. Butler, A. Carracedo, D. Syndercombe Court, L.A. Dixon, B. Filipović, M. Fondevila, P. Gill, C.D. Harrison, C. Hohoff, R. Huel, B. Ludes, W. Parson, T.J. Parsons, C. Phillips, H. Schmitter, P.M. Schneider, P.M. Vallone, N. Morling, Forensic typing of autosomal SNPs with a 29 SNP-multiplex – results of a collaborative EDNAP exercise, Forensic Sci. Int. Genet. 2 (2008) 176–183.

[13] A. Kong, D.F. Gudbjartsson, J. Sainz, G.M. Jonsdottir, S.A. Gudjonsson, B. Richardsson, S. Sigurdardottir, J. Barnard, B. Hallbeck, G. Masson, A. Shlien, S.T. Palsson, M.L. Frigge, T.E. Thorgeirsson, J.R. Gulcher, K. Stefansson, A high-resolution recombination map of the human genome, Nat. Genet. 31 (2002) 241–247.

[14] F. Sun, M. Oliver-Bonet, T. Liehr, H. Starke, E. Ko, A. Rademaker, J. Navarro, J. Benet, R.H. Martin, Human male recombination maps for individual chromosomes, Am. J. Hum. Genet. 74 (2004) 521–531.

[15] M. Poetsch, C. Lüdcke, A. Repenning, L. Fischer, V. Mályusz, E. Simeoni, E. Lignitz, M. Oehmichen, N. von, Wurmb-Schwark, The problem of single parent/child paternity analysis – practical results involving 336 children and 348 unrelated men, Forensic Sci. Int. 159 (2006) 98–103.

[16] C. Phillips, R. Fang, D. Ballard, M. Fondevilla, C. Harrison, F. Hyland, E. Musgrave-Brown, C. Proff, E. Ramos-Luis, B. Sobrino, A. Carrecedo, M.R. Furtado, D. Syndercombe Court, P.M. Schneider, Evaluation of the Genplex SNP typing system and a 49plex forensic marker panel, Forensic Sci. Int. Genet. 1 (2007) 180–185.

[17] M.K. Balogh, K. Bender, P.M. Schneider, and the SNP*for*ID consortium, Application of Nanogen microarray technology for forensic SNP analysis, Int. Congress Series 1288 (2006) 43–45.

[18] K. Bender, C. Nehlich, C. Harrison, E. Musgrave-Brown, D. Syndercombe-Court, P.M. Schneider, and the SNP*for*ID consortium, A multiplex SNP typing approach for the DNA pyrosequencing technology, Int. Congress Series 1288 (2006) 73–75.

[19] M. Bogus, B. Sobrino, K. Bender, A. Carrecedo, P.M. Schneider, and the SNP*for*ID consortium, Rapid microarray-based typing of forensic SNPs, Int. Congress Series 1288 (2006) 37–39.