



Validation of short tandem repeat analysis for the investigation of cases of disputed paternity

J.A. Thomson^{a,*}, V. Pilotti^a, P. Stevens^b, K.L. Ayres^c,
P.G. Debenham^{a,b}

^a*Life Sciences Group, LGC, Queens Road, Teddington, Middlesex TW11 0LY, UK*

^b*UDL, Southbank Technopark, London, UK*

^c*Dept of Applied Statistics, University of Reading, Reading, UK*

Received 24 June 1998; received in revised form 30 October 1998; accepted 2 November 1998

Abstract

This study details validation of two separate multiplex STR systems for use in paternity investigations. These are the Second Generation Multiplex (SGM) developed by the UK Forensic Science Service and the PowerPlexTM1 multiplex commercially available from Promega Inc. (Madison, WI, USA). These multiplexes contain 12 different STR systems (two are duplicated in the two systems). Population databases from Caucasian, Asian and Afro-Caribbean populations have been compiled for all loci. In all but two of the 36 STR/ethnic group combinations, no evidence was obtained to indicate inconsistency with Hardy-Weinberg (HW) proportions. Empirical and theoretical approaches have been taken to validate these systems for paternity testing. Samples from 121 cases of disputed paternity were analysed using established Single Locus Probe (SLP) tests currently in use, and also using the two multiplex STR systems. Results of all three test systems were compared and no non-conformities in the conclusions were observed, although four examples of apparent germ line mutations in the STR systems were identified. The data was analysed to give information on expected paternity indices and exclusion rates for these STR systems. The 12 systems combined comprise a highly discriminating test suitable for paternity testing. 99.96% of non-fathers are excluded from paternity on two or more STR systems. Where no exclusion is found, Paternity Index (PI) values of >10 000 are expected in >96% of cases. © 1999 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Paternity; Parentage; STR; Short tandem repeat; Microsatellite

*Corresponding author. Tel.: +44-181-9437604; fax: +44-181-9437659.

E-mail address: jat@lgc.co.uk (J.A. Thomson)

1. Introduction

Short Tandem Repeat (STR) loci are polymorphic loci found throughout all eukaryotic genomes. They characteristically consist of tandem arrays of short repeated sequences of 2–6 base pairs in length. Polymorphism occurs when the number of copies of the repeat sequence present at a given STR locus varies between individual chromosomes.

DNA profiling by analysis of STR loci has been widely used in forensic casework since 1993. Although most STR systems are not as polymorphic as the previously used single-locus probe (SLP) analysis of variable number of tandem repeat (VNTR) loci, they have gained widespread acceptance due to some significant advantages. STR analysis is dependant on the Polymerase Chain Reaction (PCR) which confers much greater sensitivity on the test system. Typically, as little as 1 ng of genomic DNA will yield a full STR profile whereas SLP analysis requires at least 100 ng for reliable profiling. STR analysis is also a considerably less time consuming procedure than SLP profiling, which relied on a number of sequential Southern Blot hybridisations of restricted genomic DNA. The PCR based approach not only allows the simultaneous analysis of several STR loci but is also more amenable to automation.

These benefits are as pertinent in parentage testing as in forensic casework. Greater sensitivity will allow the use of more convenient sample types. Currently, blood samples of 1–5 ml are commonly taken for paternity testing, which can prove problematic when sampling very young children and babies. A testing method only requiring blood drops from finger or heel pricks taken onto paper stain cards, or buccal swabs would offer significant benefits in terms of ease of taking the sample, transportation to the laboratory and storage. Similarly, all laboratories can benefit from faster and, where appropriate, automated methodologies.

Despite these benefits, the adoption of STR profiling in cases of disputed parentage has lagged behind its use in forensic casework. This is primarily because the relative reduction in discriminating power with respect to SLP profiling has a more profound effect in parentage testing, where generally only one allele at each locus is informative, than in identity testing, where a match at both alleles is required. This distinction means that a six locus STR system such as the Second Generation Multiplex (SGM) has been designed which has an equivalent discriminating power in identity cases to four SLPs [1] but many more STRs are required to provide equivalent paternity indices to the six SLPs currently used by the UDL laboratory in cases of disputed paternity.

The continuing development and validation of STR systems for use in identity testing have now resulted in twenty or more suitable STR systems being available either commercially or via published primer sequences. This represents an abundance even to the paternity tester. Provided that these systems can be shown to satisfy the basic requirements of mendelian inheritance, independence and sufficiently low mutation rates, then a battery of such tests can be applied. To some, this progression from four or six SLP tests to perhaps 12 or more STR tests represents a step backwards and indeed recalls the arrays of serological assays and protein polymorphisms used in paternity investigations in the past. The key point, however, is that these 12 or more systems can be co-amplified in just one or two multiplex PCR reactions which are analysed sequentially on the same electrophoretic analysis platform.

This paper details validation of two separate multiplex STR systems for use in paternity investigations. Empirical and theoretical approaches have been taken to validate these systems. Samples from 121 cases of disputed paternity, routinely submitted to this laboratory, were analysed using the established SLP tests currently in use, and also using the two multiplex STR systems. Results from the three test systems were compared and any anomalies identified and investigated. The data was then analysed to give information on expected paternity indices and exclusion rates for these STR systems.

In addition, population frequency databases have been constructed for the three major ethnic groups in the UK; British Caucasians, Afro-Caribbeans and Asians. These databases were analysed for consistency with Hardy-Weinberg (HW) proportions and were subsequently used in the calculation of paternity indices and in other analysis carried out on the data.

The two multiplex STR systems used were the SGM system and the PowerPlexTM1 system. The SGM system, developed by the UK Forensic Science Service [2] consists of the loci HUMAMGX/Y, HUMTH01 (referred to in text as TH01), HUMVWFA31/A (vWA), D8S1179, D21S11, HUMFIBRA (FGA) and D18S51. The PowerPlexTM1 system, commercially available from Promega (Madison, WI, USA) [3], consists of the loci D5S818, D13S317, D7S820, D16S539, HUMVWFA31/A, HUMTH01, HUMTPOX (TPOX) and HUMCSF1PO (CSF1PO). This gives a total of 12 different STR loci plus the amelogenin sex test (HUMAMGX/Y). Two of the loci, HUMVWFA31/A and HUMTH01 are duplicated between the two systems.

Results are presented for the two multiplex systems separately and for the 12 loci combined.

2. Materials and methods

2.1. DNA samples

Samples used in the validation for parentage testing were obtained from 121 paternity cases routinely submitted to the UDL laboratory for SLP analysis. Each case consisted of whole blood samples (drawn in EDTA) from the mother, one child and one putative father. Samples for construction of population frequency databases were obtained from unrelated individuals submitted to the UDL laboratory. Additional Afro-Caribbean samples from unrelated individuals were kindly provided by Dr Mark Layton, King's College London.

2.2. DNA extraction and quantitation

DNA extraction was carried out using a standard proteinase K/Phenol-Chloroform organic extraction method on an Applied Biosystems 340A Nucleic Acid Extractor.

One microlitre aliquots of extracted DNA were diluted 100-fold with sterile distilled water and quantified using PicoGreen® (Molecular Probes) on a Labsystems Flouroskan Ascent fluorometer [4].

2.3. Single locus probe (SLP) analysis

SLP analysis was carried out using six probes, YNH24 [5], TBQ7 [6], EFD52 [7] (Promega Corporation, Madison, WI), LH1 (Life Technologies, Paisley, UK), MR24/1 [formerly Amersham UK, now UDL], 3' alpha HVR. [8].

Extracted DNA was digested with *Alu I* restriction endonuclease, run on 20 cm 0.8% agarose/1×TBE gels and Southern blotted onto Biodyne B membrane. Probes were labelled with ³²P dCTP by random priming and hybridised with the membranes at 65°C overnight.

The autoradiographs were analysed on the Biotrac system using BIOCRSE SLP data software (Foster and Freeman Ltd., Evesham, UK) and band matching performed to within a ±2.8% window. Statistical analysis for the six probes was performed with an in house computer program using SLP allele frequency data taken from the UDL databases.

2.4. DNA amplification

Primer sequences and fluorescent dye specifications for the seven SGM loci have been previously published [2]. Primer sequences for the PowerPlexTM1 are commercial information and are not available for publication.

SGM amplifications contained 2 ng of target DNA in a 50 µl reaction volume containing 1×PARR buffer (Cambio Ltd, Cambridge, UK), 1.25U AmpliTaq GoldTM DNA Polymerase (Perkin-Elmer, Norwalk, CT, USA) and 200 µM of each dNTP (Pharmacia, Uppsala, Sweden). Samples were amplified using the following conditions: 95°C for 18 min., then 30 cycles of 95°C for 30 sec., 58°C for 75 sec., 72°C for 15 sec., then 72°C for 25 min. Primer concentrations were 0.045 µM AMG 1/2, 0.0875 µM TH01 1/2, 0.125 µM vWA 1/2, 0.56 µM D8S1179 1/2, 0.21 µM D21S11 1/2, 0.1 µM FGA 1/2, 0.1 µM D18S51 1/2.

PowerPlexTM1 amplifications contained 2 ng of target DNA in a 25 µl reaction volume containing 1×PARR buffer (Cambio Ltd, Cambridge, UK), 2U AmpliTaq GoldTM DNA Polymerase (Perkin-Elmer, Norwalk, CT, USA) and 200 µM of each dNTP (Pharmacia, Uppsala, Sweden). Samples were amplified using the following conditions: 95°C for 11 min, followed by 96°C for 1 min.; then 10 cycles of 94°C for 30 sec, ramp 68 seconds to 60°C, hold for 30 sec, ramp 50 sec to 70°C, hold for 45 sec; then 20 cycles of 90°C for 30 seconds, ramp 60 sec to 60°C hold for 30 seconds, ramp 50 sec to 70°C, hold for 45 sec, then 60°C for 30 min. Primer concentrations were as specified by the manufacturers (Promega Inc., Madison, WI).

All PCR reactions were carried out in a Perkin-Elmer 9600 thermal cycler in 0.2 ml thin-walled tubes with no oil overlay.

2.5. Electrophoresis and analysis

For SGM analysis, 1 µL of each PCR product was mixed with 1 µL of formamide, 0.3 µL GS500 TAMRA dye-labelled internal size standard (PE-ABI, Warrington, UK) and 0.2 µL Dextran Blue solution (Promega Inc., Madison, WI). For PowerPlexTM1 analysis, 1 µL of each PCR product was mixed with 1 µL of Dextran Blue loading

Table 1
Allele frequency population databases for SGM loci and PowerPlex™ 1 loci

D16S51			
Allele	Caucasian n=364	Afro-Carib n=374	Asian n=372
9	0.000	0.000	0.000
16	0.505	0.003	0.005
10.2	0.500	0.000	0.000
11	0.013	0.005	0.024
12	0.133	0.033	0.067
13	0.102	0.037	0.129
13.2	0.000	0.005	0.000
14	0.167	0.043	0.255
14.2	0.000	0.003	0.000
15	0.141	0.147	0.177
16	0.146	0.147	0.132
17	0.167	0.130	0.091
17.2	0.000	0.003	0.000
18	0.078	0.112	0.035
19	0.021	0.120	0.040
19.2	0.000	0.000	0.000
20	0.018	0.053	0.016
21	0.010	0.035	0.008
22	0.005	0.008	0.008
23	0.000	0.005	0.008
24	0.000	0.000	0.003

FGA			
Allele	Caucasian n=364	Afro-Carib n=378	Asian n=368
16	0.003	0.000	0.003
17	0.000	0.003	0.003
18	0.005	0.008	0.011
18.2	0.000	0.016	0.000
19	0.068	0.063	0.065
19.2	0.000	0.011	0.000
20	0.120	0.048	0.087
20.2	0.000	0.005	0.000
21	0.206	0.114	0.152
21.2	0.003	0.000	0.011
22	0.195	0.206	0.179
22.2	0.005	0.000	0.016
23	0.148	0.183	0.160
23.2	0.008	0.000	0.003
24	0.122	0.156	0.133
24.2	0.003	0.000	0.000
25	0.073	0.095	0.084
25.2	0.000	0.000	0.003
26	0.034	0.037	0.027
27	0.005	0.032	0.003
28	0.003	0.016	0.000
28	0.000	0.003	0.003
30	0.000	0.003	0.000
46.2	0.000	0.003	0.000

CSF1PO			
Allele	Caucasian n=384	Afro-Carib n=314	Asian n=400
7	0.000	0.078	0.003
8	0.005	0.054	0.005
9	0.026	0.064	0.028
10	0.232	0.248	0.200
11	0.326	0.245	0.296
12	0.321	0.288	0.370
13	0.079	0.041	0.093
14	0.011	0.003	0.008

TPOX			
Allele	Caucasian n=434	Afro-Carib n=314	Asian n=400
6	0.002	0.064	0.003
7	0.000	0.033	0.000
8	0.518	0.277	0.433
9	0.113	0.287	0.143
10	0.036	0.102	0.075
11	0.274	0.204	0.315
12	0.056	0.025	0.028
13	0.002	0.003	0.005

D13S317			
Allele	Caucasian n=420	Afro-Carib n=314	Asian n=400
7	0.000	0.000	0.013
8	0.114	0.022	0.235
9	0.086	0.010	0.098
10	0.050	0.025	0.075
11	0.336	0.306	0.188
12	0.276	0.455	0.283
13	0.107	0.121	0.075
14	0.029	0.054	0.035
15	0.002	0.006	0.000

D5S618			
Allele	Caucasian n=432	Afro-Carib n=314	Asian n=400
8	0.006	0.054	0.000
9	0.039	0.013	0.050
10	0.044	0.073	0.123
11	0.354	0.285	0.390
12	0.400	0.325	0.270
13	0.150	0.239	0.163
14	0.007	0.025	0.005
15	0.000	0.003	0.000

D7S820			
Allele	Caucasian n=426	Afro-Carib n=314	Asian n=400
7	0.016	0.010	0.028
8	0.131	0.217	0.213
9	0.157	0.105	0.065
10	0.317	0.360	0.245
11	0.209	0.188	0.245
12	0.127	0.098	0.183
13	0.028	0.025	0.023
14	0.016	0.000	0.000

D21S11			
Allele	Caucasian n=384	Afro-Carib n=366	Asian n=366
54	0.000	0.003	0.000
56	0.000	0.000	0.003
57	0.003	0.000	0.000
59	0.052	0.041	0.016
61	0.133	0.257	0.146
62	0.000	0.000	0.000
63	0.193	0.238	0.213
64	0.000	0.000	0.003
65	0.271	0.190	0.180
66	0.031	0.006	0.030
67	0.073	0.050	0.044
68	0.107	0.057	0.104
69	0.016	0.019	0.014
70	0.078	0.077	0.183
71	0.000	0.000	0.000
72	0.038	0.038	0.049
73	0.000	0.011	0.000
74	0.003	0.003	0.011
75	0.003	0.036	0.000
76	0.003	0.003	0.003
77	0.000	0.000	0.000
78	0.000	0.003	0.000
79	0.000	0.003	0.000

D8S1179			
Allele	Caucasian n=384	Afro-Carib n=376	Asian n=368
8	0.016	0.000	0.022
9	0.016	0.000	0.005
10	0.096	0.019	0.179
11	0.076	0.040	0.090
12	0.146	0.098	0.098
13	0.307	0.197	0.177
14	0.206	0.340	0.165
15	0.112	0.245	0.168
16	0.026	0.056	0.071
17	0.000	0.005	0.005
18	0.000	0.000	0.000

D16S539			
Allele	Caucasian n=420	Afro-Carib n=314	Asian n=400
8	0.029	0.022	0.065
9	0.121	0.274	0.168
10	0.062	0.092	0.115
11	0.310	0.296	0.300
12	0.224	0.191	0.205
13	0.200	0.105	0.125
14	0.055	0.019	0.023

vWA			
Allele	Caucasian n=384	Afro-Carib n=372	Asian n=370
11	0.000	0.003	0.000
12	0.000	0.003	0.000
13	0.000	0.030	0.003
14	0.089	0.067	0.103
15	0.099	0.185	0.092
16	0.255	0.285	0.216
17	0.247	0.212	0.249
18	0.208	0.134	0.211
19	0.085	0.043	0.108
20	0.013	0.022	0.019
21	0.000	0.016	0.000

TH01			
Allele	Caucasian n=384	Afro-Carib n=384	Asian n=366
5	0.003	0.003	0.000
6	0.201	0.174	0.219
7	0.148	0.438	0.158
8	0.120	0.214	0.142
9	0.130	0.125	0.322
9.3	0.393	0.042	0.153
10	0.005	0.005	0.005
11	0.000	0.000	0.006

solution and 0.3 μ L CXR Fluorescein dye-labelled internal standard (Promega Inc., Madison, WI).

All samples were denatured at 94°C for 5 minutes and loaded onto a 36 cm well-to-read, 0.2mm thick, denaturing polyacrylamide gel (4% 19:1 acrylamide:*bis*-acrylamide, 8M urea). Gel buffer and running buffer was 1 \times TBE. Gels were run on a PE-ABI PRISM 377 analyser at constant 3000 V at 51°C for 2 h.

Data was analysed using GeneScan and GenoTyper analysis software (PE-ABI, Warrington, UK).

3. Results

3.1. Population databases

Databases of allele frequencies for the 12 STR loci were constructed (see Table 1) For SGM, profiles were obtained from between 183 and 192 individuals from each of the three major ethnic groups in the UK; white skinned caucasian, afro-caribbeans and sub-continental asians. For PowerPlexTM1, 157 to 217 individuals were analysed.

For each locus, a randomisation exact test for Hardy-Weinberg (HW) equilibrium was performed. *P* values are given in Table 2. The results show significant (at the 5% level) deviation from HW in two of the 36 locus/ethnic group combinations. This may be observed by chance because of the large number of tests performed or may represent genuine inconsistency with HW.

To further investigate this, posterior probability density curves for the inbreeding coefficient *f* (a measure of departure from HW) were obtained, using the method of Ayres and Balding [9]. Since *f*=0 is consistent with HW proportions, density curves which give no obvious support to zero suggest that the locus may be inconsistent with HW (note however that it is possible for a population to undergo inbreeding and yet still

Table 2
P values from randomisation exact test of Hardy-Weinberg equilibrium, using 5000 random permutations

Multiplex	Locus	Caucasian	Afro-Car.	Asian
SGM	D21S11	0.228	0.538	0.175
	D18S51	0.510	0.190	0.489
	FGA	0.702	0.796	0.779
	D8S1179	0.511	<i>0.029</i>	0.135
SGM+	THO1	0.910	0.258	0.473
PowerPlex	vWA	0.320	0.182	0.972
PowerPlex	CSF1PO	0.663	0.879	0.205
	D13S317	0.581	0.529	0.316
	D16S539	<i>0.035</i>	0.221	0.512
	D5S818	0.388	0.280	0.557
	D7S820	0.460	0.315	0.780
	TPOX	0.731	0.394	0.844

Significant deviations (*P*<5%) are shown in italics.

be consistent with HW, due to factors such as selection). Most of the posterior curves obtained for the three populations give strong support to $f=0$ (based on a prior belief that f is likely to be small, and less than 20%; results not shown). Of the small number that give very little support to $f=0$, the most notable is locus D16S539 for the Caucasian population, since this case also produced a P value significant at the 5% level.

Although our results suggest that $f=0$ cannot be unequivocally excluded at any locus in the three populations, it seems likely that there is some departure from HW. A possible explanation for this is population subdivision. This is a phenomenon long since recognised, and allowed for in forensic identification by incorporating into the analyses the population genetics parameter F_{ST} [10]. The arguments presented there are also valid in parentage testing. Hence, throughout this paper we incorporate the parameter F_{ST} into our analyses of the STR systems.

3.2. Comparison of SLP and STR results

The SLP results were taken as reference results for the validation exercise and these, together with the STR results are summarised in Table 3.

Of the 121 cases analysed, 88 putative fathers were not excluded from paternity using between four and six SLP systems. Paternity Index (PI) values for these cases ranged from 11 140 (four probes used) to 2×10^9 . Where six SLP probes were employed the lowest PI value obtained was 29 000.

The remaining 33 putative fathers were excluded from paternity. No specific analysis has been carried out on these excluded men as the exclusion rate analysis described below simulates a large number of families where the man is not the father.

Cases where only a single STR system did not give an analysable result were included in the full analysis. Cases where more than one STR locus did not give an analysable

Table 3
Summary of STR Results in 121 Paternity cases

Reference result	STR Result		SGM (no. of cases)	PowerPlex™1 (no. of cases)	SGM + P'Plex™1 (no. of cases)
SLP	Inclusion ¹	(0 mismatches)	81	78	71
Inclusion	Inclusion ²	(Incomplete profile)	5	8	13
88 cases	Inconclusive ³	(single mismatch)	2	2	4
	Exclusion	(>1 mismatch)	0	0	0
SLP	Exclusion	(0 exclusions)	0	0	0
Exclusion	Inconclusive	(single mismatch)	1	2	0
33 cases	Exclusion ¹	(>1 mismatch)	31	28	29
	Exclusion ²	(Incomplete profile)	1	3	4

Results shown are number of cases falling into each indicated category, referenced against the SLP results.

¹, Results for profiles showing 6/6 or 5/6 reportable results for individual systems (SGM or PowerPlex™1) and 12/12 or 11/12 reportable results for combined systems.

², Results for partial profiles showing <5 reportable results for individual systems or <11 reportable results for combined systems.

³, Subject to statistical analysis to incorporate mutation rates into paternity index calculation.

result were recorded as incomplete profiles but the acceptable results for these cases were recorded and compared in the same manner as the complete profiles.

These results show no non-conformities between the conclusions concerning paternity drawn from SLP and STR results. However, amongst the 88 cases where the putative father was not excluded using SLPs, four men showed a mismatch with the child at a single STR locus, which were scored as Inconclusive results. These are classified as likely mutations and were subjected to further investigation as discussed below. These samples are detailed in Table 4.

In addition to these anomalies, one instance was observed of an mismatch between mother and child at the D21S11 locus. As mother and child are both homozygous for this locus the most likely explanation is a null allele caused by a deletion or primer site mutation. This result is also detailed in Table 4.

All five of these families were reanalysed from the original samples to confirm these anomalous results. All four presumed mutations conform to the single step mutation pattern commonly associated with mutations in microsatellite loci [11], that is that the presumed *de novo* alleles in the child are a single STR repeat unit removed from one of the paternal alleles. Three of these presumptive single step mutations involve the loss of a repeat unit.

Mutation rates based on these observations can be estimated as follows; D16S539: 1.3×10^{-2} (2 in 156 meioses); D18S51: 6×10^{-3} (1 in 162 meioses); FGA: 6×10^{-3} (1 in 162 meioses); all other loci: $< 6 \times 10^{-3}$ (no mutations observed).

Table 4

Details of cases where a single excluding result was observed in either one of the SGM and PowerPlexTM1 multiplex systems

Case	Reference result (SLP)	Exclusion details				Comment
		System	Locus	Rel	Result	
A	Inclusion PI=380 000	PowerPlex TM 1	D16S539	PF	12,13	Putative mutation: D16S539 13→14 (or 12→14)
				C	13,14	
				M	13,13	
B	Inclusion PI=2.2 million	PowerPlex TM 1	D16S539	PF	12,12	Putative mutation: D16S539 12→11 (or 12→14)
				C	11,14	
				M	11,14	
C	Inclusion PI=13 million	SGM	D18S51	PF	16,19	Putative mutation: D18S51 19→18 (or 16→18)
				C	15,18	
				M	14,15	
D	Inclusion PI=38 000	SGM	FGA	PF	21,24	Putative mutation: FGA 24→23 (or 21→23)
				C	23,23	
				M	21,23	
E	Inclusion PI=29 000	SGM	D21S11	PF	59,61	Maternal Exclusion (possible null allele)
				C	61,61	
				M	65,65	

These cases represent presumed mutations in one of the 12 STR loci, or in case E as possible maternally inherited null allele.

3.3. Exclusion rate analysis

Two approaches were taken to investigate the power of the STR systems to exclude non-fathers from paternity.

An empirical approach analysed the data from 120 of the 121 cases (one case being omitted due to repetition of mother and child with different putative fathers). Each known mother-child pair was linked with each of the 119 unrelated men from every other case. This approach allows for 14 280 families to be constructed where the man is presumed not to be the biological father of the child (note, though, that not all combinations yield acceptable results due to the presence of partial profiles). STR results for each locus were examined to determine whether the man was excluded and the number of excluding STR systems against each non-father was summed.

A theoretical approach was also used to calculate frequencies for numbers of exclusions among the SGM and PowerPlexTM1 loci, and the two combined. Average exclusion rates for each locus were calculated from population allele frequencies via a weighted summation of the joint probability of all possible mother/child/excluded man configurations (father and excluded man were assumed to be of the same ethnicity). Specifically, for each choice of ethnicity for mother and father joint probabilities for each configuration were calculated via the recursive formula outlined in equation (12) of Balding and Nichols [10]. This formula expresses the probability that the $n+1$ th gene drawn from a population is of allele type A , as follows:

$$P_{n+1}(A) = \frac{n_A F + (1 - F)p_A}{1 + (n - 1)F} \quad (1)$$

where n_A is the number of alleles of type A observed thus far, p_A is the population (e.g. database) frequency of allele A , and F (often equivalent to F_{ST}) models uncertainty about the allele frequencies due to recent shared ancestry of the individuals in the population.

The weights were derived under the assumption that parents are equally likely to be of any race (other weightings were investigated, and found to have little effect on the overall inferences). We also make the assumption that if individuals belong to the same ethnic group, they share a common level of ancestry, modelled here by $F_{ST} = 1\%$ (this being a reasonable choice for each population, based on the HW analysis). Expected frequencies for numbers of excluding loci in the SGM and PowerPlexTM1 systems, and combined, were calculated under the assumption of independence between loci.

Overall exclusion rates for each locus and for all the loci combined (assuming independence between loci) are shown in Table 5. Observed exclusion rates are based on all cyclic permutations of presumed non-fathers with mother/child pairs as described above. All theoretical and observed exclusion rates are in close agreement and the overall exclusion rate based on combining individual locus rates suggests that less than 0.002% of non-father/child pairs (<1 in 50 000) would fail to show at least one mis-match. This corresponds to a power of exclusion of >0.99998 but should be considered in conjunction with later discussion regarding mutation rates.

Table 5
Individual and combined exclusion rates for all SGM and PowerPlex™1 loci

Multiplex	Locus	Obs. exc. rate	Theor. exc. rate
SGM	D21S11	0.711	0.685
	D18S51	0.737	0.720
	FGA	0.729	0.696
	D8S1179	0.642	0.631
SGM/ PowerPlex	THO1	0.565	0.525
	VWA	0.604	0.611
PowerPlex	CSF1PO	0.452	0.516
	D13S317	0.570	0.542
	D16S539	0.613	0.588
	D5S818	0.415	0.483
	D7S820	0.648	0.578
	TPOX	0.399	0.466
Combined SGM (6 loci)		0.998734	0.998177
Combined P'Plex™1 (8 loci)		0.998064	0.998034
SGM+P'Plex™1 (12 loci)		0.99998575	0.99998061

Observed exclusion rates are based on all cyclic permutations of presumed non-fathers with mother/child pairs (10 297–13 689 cases). Theoretical rates were calculated as described in the text. Combined rates assume independence between loci.

The results showing the numbers of excluding systems for each family are shown in Fig. 1.

These results indicate that the SGM and the PowerPlex™1 perform very similarly with regard to excluding non-fathers. For SGM, 0.14% of acceptable cases showed no exclusions in the empirical false family experiment (*0.18% by the theoretical method*), 1.91% (*2.08%*) showed just a single exclusion and 97.95% (*97.74%*) showed 2 or more exclusions. For PowerPlex™1, 0.23% (*0.19%*) of cases showed no exclusion, 2.34% (*1.87%*) showed a single exclusion and 97.43% (*97.94%*) showed 2 or more exclusions.

The combined SGM/PowerPlex™1 figures indicate that a single exclusion only is observed in 0.04% (*0.04%*) of cases where a non-father is tested and no examples of zero exclusions are observed in the 9402 acceptable cases constructed for this study (*0.002%*) Two or more exclusions were observed in 99.96% (*99.96%*) of cases.

3.4. Paternity index distribution

A number of experiments were carried out to investigate the distribution of paternity index (PI) values.

Balding and Nichols ([12], Tables 1 and 2) give the relevant PI formulae, incorporating F_{ST} , when the mother, alleged father and alternative father are assumed to share a common level of ancestry. When the mother is assumed not to share recent ancestry with the two men (e.g. she is of a different ethnic group) the necessary formulae can be calculated via repeated use of formula (1).

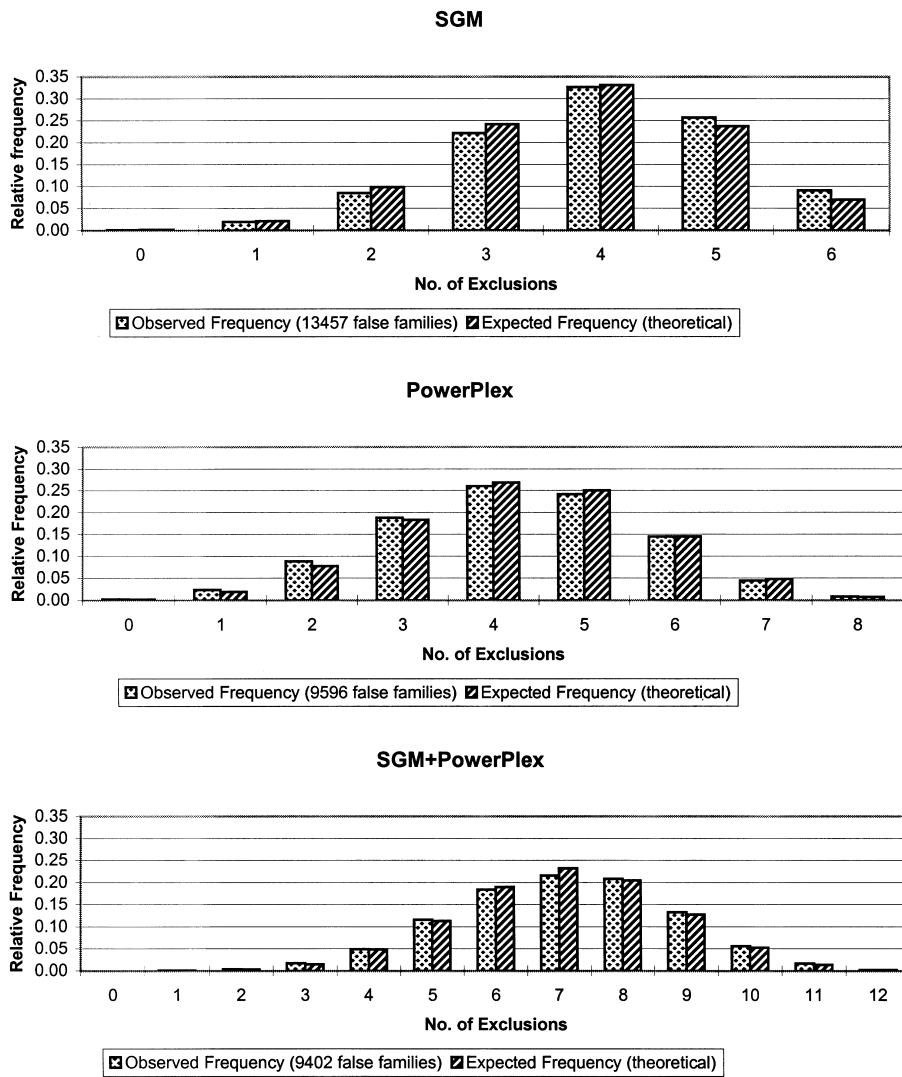


Fig. 1. Relative frequencies of number of excluding STR systems for SGM, PowerPlexTM1 and SGM+PowerPlexTM1. Expected frequencies are calculated from allele distributions in population frequency databases, incorporating an F_{ST} value of 1%. Observed frequencies were based on an empirical assessment of 120 mother-child pairs, compared with presumed non-fathers. Each known non-father was cycled through all known mother-child pairs in the data set and numbers of excluding systems was summed for each false family.

PI values (incorporating F_{ST}) were calculated for the 81 SGM cases and 77 PowerPlexTM1 cases where a valid zero exclusion result was obtained, and where no exclusion was observed in the SLP results. We again assumed that individuals with the same ethnicity share a common level of ancestry, implementing the value of 1% for F_{ST} .

In addition, the expected distribution of PI values was approximated for SGM, PowerPlexTM1 and SGM+PowerPlexTM1. By using the population database allele frequencies, and assuming $F_{ST}=1\%$ and independence between loci, expected frequencies of single-locus and multi-locus genotypes (product of single-locus frequencies)

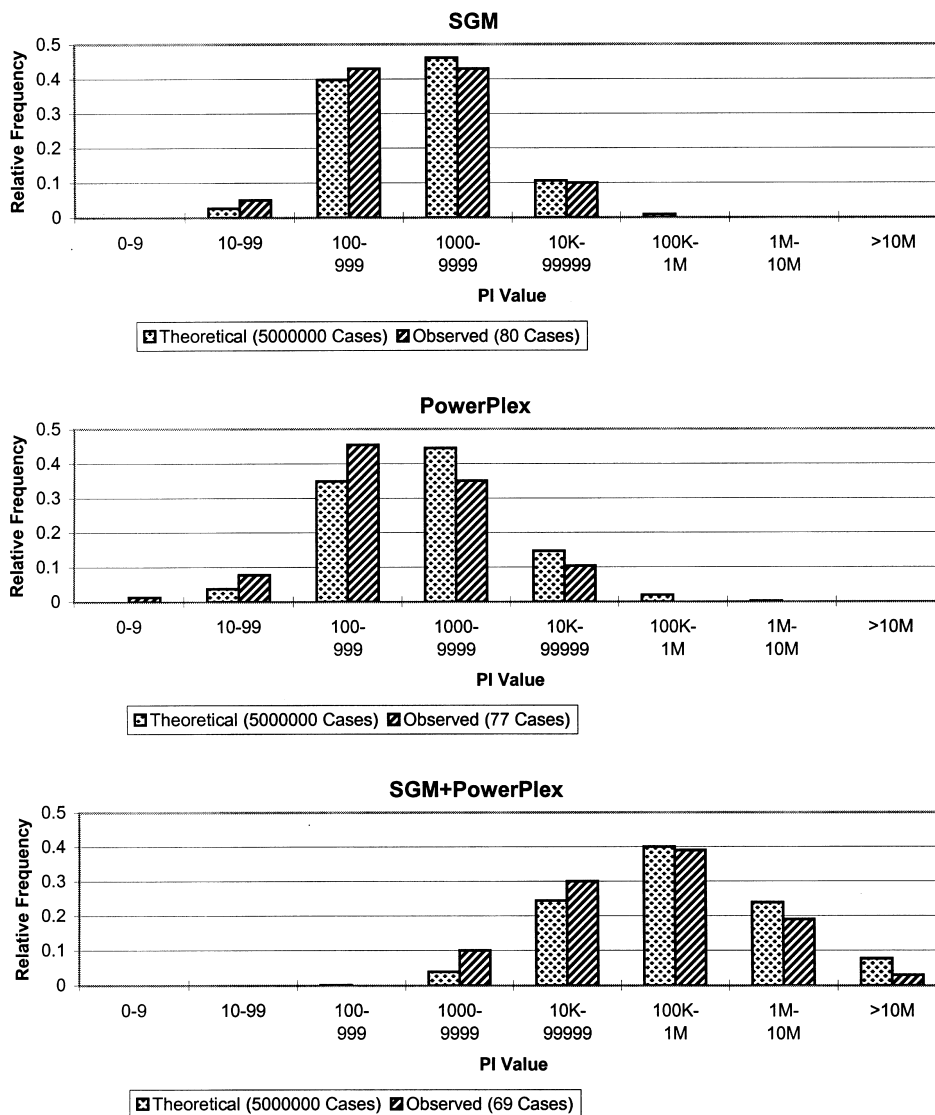


Fig. 2. Relative frequencies of Paternity Index (PI) ranges for SGM, PowerPlexTM1 and SGM+PowerPlexTM1. Theoretical values are obtained from 500 000 constructed families with genotype distributions derived from population databases. Observed values are from paternity casework samples submitted to this laboratory. An F_{ST} value of 1% was incorporated in all PI calculations.

can be calculated for each possible mother/father configuration. Hence, a theoretical “pool” of mothers and fathers can be created for each choice of ethnicity, and a number of theoretical parents randomly selected according to these genotype frequencies. In this experiment the sampling was biased such that 90% of mother/father pairs chosen were of the same race. This figure is somewhat arbitrary, but attempts to approximate the true position where mixed race pairings accounted for about 5% of the real cases in this study.

Having sampled a mother/father pair, a theoretical child can then be chosen, assuming Mendelian segregation. For each theoretical family created in this way, the paternity index can be calculated for each locus, and hence the relevant products taken to produce PI for the SGM and PowerPlex™1 groups of loci, as well as all loci combined.

In total, 500 000 random families were generated, and counts determined for various ranges of PI values. Both sets of data are shown in Fig. 2.

Fig. 3 shows a comparison of the combined STR PI values with the observed SLP PI values from the 85 non-exclusion cases analysed. For clarity, only the observed values for the STR cases are included. In addition to the STR values calculated with $F_{ST}=1\%$ (as used elsewhere in this paper), PI values calculated with $F_{ST}=0\%$ are also shown to allow a more direct comparison to be made with the SLP data. This indicates that the two approaches yield similar profiles of PI values. SLP values show 100% of cases with $PI > 10\,000$ while the observed STR values show 90% of cases with $PI > 10\,000$. Observed STR cases with PI values $< 10\,000$ were slightly elevated due to the inclusion of cases where only 11 of the 12 loci in the combined systems were successfully profiled. This highlights the need to continue to optimise PCR reactions to minimise this level of locus dropout.

A slightly higher proportion of SLP than STR families showed very high (> 10 million) PI values. 22% of SLP cases exceed this figure compared to 6% in the observed STR casework study (with $F_{ST}=0$).

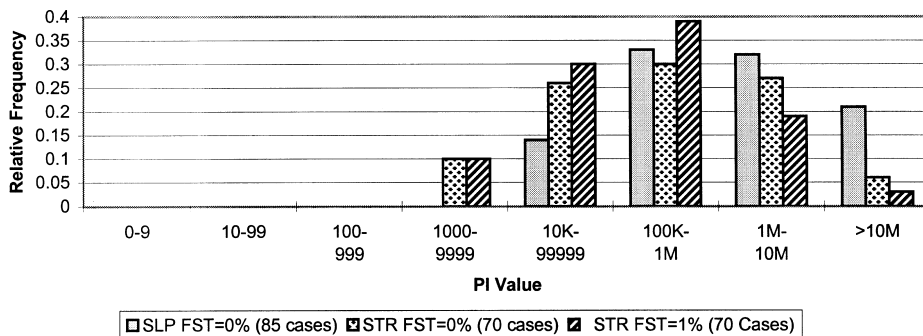


Fig. 3. Relative frequencies of Paternity Index (PI) ranges for combined SGM and PowerPlex™1 systems and for four to six Single Locus Probe (SLP) systems. Results are those observed in the reported cases. STR results are shown as calculated with F_{ST} values of 0% and 1%. SLP results are calculated with $F_{ST}=0\%$ as current practice in casework.

4. Discussion

The introduction of STR typing into routine paternity analysis is desirable as it will allow the use of smaller and more convenient sample types, faster turnaround times and is amenable to automation.

It has been suggested that STR loci do not offer sufficient levels of discrimination to replace the commonly used SLP analysis in paternity testing [13]. We have shown that a battery of 12 polymorphic STR loci offers a discriminating power only slightly less than that provided by the six SLP loci currently used in the UDL laboratory.

In our calculations for the STR systems, we have followed the recommendation of Balding and Nichols [11] and have modelled shared ancestry by incorporating the parameter F_{ST} into the PI and exclusion probability formulae. We have implemented the value of 1% which is reasonable given our preliminary analyses of the population databases. However, further analyses of subpopulation data are required for these STR loci in order to obtain more precise estimates of F_{ST} (see [15] for details of a method for estimating F_{ST}). Special circumstances, for example when individuals belong to an isolated community, may necessitate the use of a figure greater than 1%. It should be noted that unlike in criminal identity cases, where calculations and assumptions tend to be biased in favour of the defendant, in paternity cases every effort should be used to employ a realistic figure so neither party's interests are favoured. Consequently the "conservative" approach of incorporating unsupported high F_{ST} values of 3% or more is inappropriate here.

Overall exclusion power of the 12 loci combined is estimated at >99.998%, indicating that it is expected that less than 1 in 50 000 of non-fathers tested with a mother and child would fail to show at least one mismatch with the paternal allele in the child. In addition, empirical and theoretical studies investigating the number of excluding loci against non-fathers show that when 12 loci are analysed, less than 1 in 2000 of non-fathers would be excluded at only a single locus and all others (>99.96%) would be excluded at two or more loci.

This result is of particular significance when considering the conclusions to be drawn if only low numbers of mis-matches are observed between man and child. In these cases, the possibility that this is the result of mutation rather than evidence of non-paternity must be considered.

The observed mutation rate for D16S539 of 1.3×10^{-2} is surprisingly high and is in excess of previously reported mutation rates for STR loci. Brinkmann et al. reported rates of 0 to 7×10^{-3} for 10 844 meioses in 9 STR loci [14]. These included TH01 (no mutations in 10 844 meioses), vWA (2.0×10^{-3}), D21S11 (1.8×10^{-3}) and FGA (4×10^{-3}). However, Lins et al. [3] have reported no observed exclusions in 317 meioses for D16S539 so an overall estimate based on combining the published data with our own data results in 2 observed mutations in 473 meioses (4×10^{-3}). Lins et al. [3] also reported only a single mutation (in the D13S317 locus) among 317 meioses for all 8 PowerPlexTM1 loci and therefore estimated an average mutation rate of 2.6×10^{-4} for these loci. Further investigation and data are required before reliable estimates of mutation rates for all these STR loci can be made but it appears that the average mutation rate for these loci may be as high as 1×10^{-3} .

Using this estimated figure applied equally to all loci, it can be calculated by a binomial formula that in approximately 1 in 84 cases, a single mutation would be observed and in 1 in 15 000 cases, two mutations would occur. Three mutations would be expected to be seen in only 1 in 4.5 million cases.

Using these figures, likelihood ratios comparing the chances of observing these low numbers of mismatches given non-paternity (i.e. “real” exclusions) and given paternity (i.e. mutations) can be calculated. (see Table 6). From this it can be seen that based on prior expectations, observation of a single mismatch only would be approximately 30 times more likely if the man is the father than if he is not (and the true father is unrelated). Two mismatches would give support to the hypothesis of non-paternity (LR=46) but clearly this is insufficient to unequivocally exclude. At the three mismatch level, the likelihood ratio in favour of non-paternity (LR=77 000) provides powerful evidence to exclude paternity. In practice, the actual genotype allele frequencies and mismatching loci data will be incorporated into a calculation of PI. If the PI is weak (<10 000) then additional loci may be tested to provide further evidence.

In cases where the man is not excluded from paternity the UDL laboratory has, in the past, aimed to obtain paternity index (PI) values of at least 10 000. If STR Profiling is to be a suitable alternative to SLP profiling it is important that it can meet this self imposed requirement. Our results show that in the experimental data set which includes a number of cases with only 11 out of the 12 STR loci giving acceptable results, 90% of cases gave PI values >10 000 while 100% gave PI values >1000. In the theoretical data set (assuming 12 successful results in every case), approx. 96% of families gave PI values >10 000 and 99.9% of cases gave PI values >1000.

It is clear from this study that a large majority of cases will meet our required levels regarding numbers of exclusions against a non-father (i.e. >2) and PI value for a non-excluded man (i.e. >10 000). To summarise, over 99.6% of non-fathers would be excluded from paternity on three or more STR systems and no mismatches observed in only 0.002% of non-fathers. Where no mismatches are observed, PI values of >10 000 are expected in 96% of cases (where 12 loci are successfully typed).

However, in 4–10% of cases (depending on overall success rate of the STR profiling)

Table 6
Prior probabilities of observing low numbers of mis-matches between putative father and child

No. of PF-C mismatches (<i>n</i>)	$P(n non-paternity)^*$ i.e. exclusions	$P(n paternity)^\dagger$ i.e. mutations	LR
1	0.0004	0.012	0.033
2	0.003	6.53×10^{-5}	46
3	0.017	2.19×10^{-7}	77625
4	0.049	4.92×10^{-10}	100 million

Note: $P(n|non-paternity)$ is obtained from the false family exclusion study (Fig. 1). $p(n|paternity)$ is calculated using the binomial function: $P(n|paternity) = q^n(1-q)^{r-n} r!/((r-n)!n!)$, where q =mutation rate and r =number of tests performed. LR is likelihood ratio of the chance of observing n number of mismatches if the man is not the father and is unrelated to the true father against the chance if the man is the true biological father.

*, Assuming that the real father is unrelated to the named man.

†. Assuming mutation rate for all loci of 10^{-3} .

where the man is not excluded PI values $<10\,000$ are observed. Whilst it is likely that in these cases PI values will be >1000 and will therefore still offer strong evidence for paternity it may be that additional STR loci could be tested in such cases to increase the PI values to the levels expected in this laboratory. Further STR loci are readily available to allow such additional testing in the small number of cases requiring it and we anticipate that a simple three or four locus multiplex can easily be incorporated into the testing regime.

Acknowledgements

KLA was supported in part by a PhD studentship from the UK EPSRC. We thank David Balding for useful discussions and initial advice and laboratory staff at UDL for technical input.

References

- [1] P. Gill, C.P. Kimpton, A. Urquhart, et al., Automated short tandem repeat (STR) analysis in forensic casework – a strategy for the future, *Electrophoresis* 16 (1995) 1543–1552.
- [2] N.J. Oldroyd, A. Urquhart, C.P. Kimpton, et al., A highly discriminating octoplex short tandem repeat polymerase chain reaction system suitable for human individual identification, *Electrophoresis* 16 (1995) 334–337.
- [3] A.M. Lins, K.A. Micka, C.J. Sprecher, et al., *J. Forens. Sciences* 43 (1998) in press.
- [4] A. Hopwood, N. Oldroyd, S. Fellows, R. Ward, S.-A. Owen, K. Sullivan, Rapid Quantification of DNA samples extracted from Buccal scrapes prior to DNA profiling, *BioTechniques* 23 (1997) 18–20.
- [5] Y. Nakamura, S. Gillilan, P. O’Connell, et al., Isolation and mapping of a polymorphic DNA sequence pYNH24 on chromosome 2 (D2S44), *Nucl. Acid. Res.* 15 (1987) 23.
- [6] Y. Nakamura, E. Fujimoto, P. O’Connell, et al., Isolation and mapping of a polymorphic DNA cEFD52 on chromosome 17q (D17S26), *Nucl. Acid. Res.* 16 (1987) 786.
- [7] T. Bragg, Y. Nakamura, C. Jones, R. White, Isolation and mapping of a polymorphic DNA sequence (cTBQ7) on Chromosome 10 (D10S28), *Nucl. Acid. Res.* 16 (1987) 11395.
- [8] A. Jarman, R.D. Nichols, D.J. Wetherall, J.B. Clegg, D.R. Higgs, Molecular characterisation of a hypervariable region downstream of the human α -globin gene cluster, *EMBO J.* 5 (1986) 1857–1863.
- [9] K.L. Ayres, D.J. Balding, Measuring departures from Hardy-Weinberg: a Markov chain Monte Carlo method for estimating the inbreeding coefficient, *Heredity* 80 (1998) 769–777.
- [10] D.J. Balding, R.A. Nichols, DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands, *Forens. Sci. Int.* 64 (1994) 125–140.
- [11] J.L. Weber, C. Wong, Mutation of human short tandem repeats, *Hum. Mol. Genet.* 8 (1993) 1123–1128.
- [12] D.J. Balding, R.A. Nichols, A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity, *Genetica* 96 (1995) 3–12.
- [13] S.D.J. Pena, Pitfalls of paternity testing based solely on PCR typing of minisatellites and microsatellites, *Am. J. Hum. Genet.* 56 (1995) 1503–1506.
- [14] B. Brinkmann, M. Klintshar, F. Neuhuber, J. Huhne, B. Rolf, Mutation rate in human microsatellites: Influence of the structure and length of the tandem repeat, *Am. J. Hum. Gen.* 62 (1998) 1408–1415.
- [15] D.J. Balding, M. Greenhalgh, R.A. Nichols, Population genetics of STR loci in Caucasians, *Int. J. Legal Med.* 108 (1996) 300–305.