Int J Legal Med (2012) 126:345–356
DOI 10.1007/s00414-011-0582-5

ORIGINAL ARTICLE

Forensic validation of the PowerPlex® ESI 16 STR Multiplex and comparison of performance with Amp*Fl*STR® SGM Plus®

Valerie C. Tucker · Amanda J. Kirkham · Andrew J. Hopwood

Received: 8 April 2011 / Accepted: 13 May 2011 / Published online: 28 May 2011

Abstract We describe the forensic validation of Promega's PowerPlex® European Standard Investigator 16 (ESI 16) multiplex kit and compare results generated with the AmpF*l*STR® SGM Plus® (SGM+) multiplex. ESI 16 combines the loci contained within the SGM+ multiplex with five additional loci: D2S441, D10S1248, D22S1045, D1S1656, and D12S391. A relative reduction in amplicon size of the SGM+ loci facilitates an increased robustness and amplification success of these amplicons with degraded DNA samples. Tests performed herein supplement ESI 16 data published previously with sensitivity, profile quality, mock casework, inhibitor and mixture study data collected in our laboratories in alignment with our internal technical and quality guidelines and those issued by the Scientific Working Group on DNA Analysis Methods (SWGDAM), the DNA Advisory Board (DAB) and the DNA working group (DNAWG) of the European Network of Forensic Science Institutes (ENFSI). Full profiles were routinely generated from a fully heterozygous single source DNA template using 62.5 pg for ESI 16 and 500 pg for SGM+. This increase in sensitivity has a consequent effect on mixture analyses and the detection of minor mixture components. The improved PCR chemistry confers enhanced tolerance to high levels of laboratory prepared inhibitors compared with SGM+ results. In summary, our results demonstrate that the ESI 16 multiplex kit is more robust and sensitive compared with SGM+ and will be a suitable replacement system for the analysis of forensic DNA samples providing compliance with the European standard set of STR loci.

Keywords Forensic · DNA · Validation · Degraded · Inhibitors · Mini-STRs

Electronic supplementary material The online version of this article (doi:10.1007/s00414-011-0582-5) contains supplementary material, which is available to authorized users.

V. C. Tucker (✉) · A. J. Kirkham · A. J. Hopwood
Research and Development, the Forensic Science Service,
2960 Trident Court, Birmingham Business Park,
Solihull Parkway,
Birmingham B37 7YN, UK
e-mail: Valerie.Tucker@fss.pnn.police.uk

Introduction

In the decade since the implementation of SGM+ [1], the work horse of forensic DNA typing in the UK, novel PCR enzymes have been developed, and a greater understanding of buffer technologies and dye chemistries has led to the development of greatly improved amplification chemistries for short tandem repeats (STRs). Advances in terms of sensitivity, robustness to inhibitors and perhaps of secondary importance, speed, are of particular interest to the forensic practitioner: the quality and quantity of samples submitted for analysis can be highly variable and may be contaminated with substances that can inhibit the DNA process. This improved performance, together with the increase in number of co-amplified loci, in line with the ENFSI and European DNA Profiling group (EDNAP) recommendation [2, 3], will provide European forensic practitioners with both improved performance and a reduced incidence of adventitious matches, particularly with partial profiles when sharing data cross-border, compared with the performance of SGM+.

To date, three manufacturers have undertaken to develop STR multiplexes to fulfil the ENFSI and EDNAP objectives. Of the commercially available kits, the Promega ESI multiplexes have been specifically designed to reduce the amplicon size of the UK National Database Loci (UKNDL)

found in the SGM+ PCR Amplification Kit, conferring an increased potential for success with degraded samples [4] for those loci.

Here, we present forensic validation work performed with the ESI 16 multiplex. Tests were designed to supplement the developmental validation data [5] and comply with internal technical and quality requirements, guidelines issued by SWGDAM [6], DAB [7], and by ENFSI DNAWG [8].

Materials and methods

DNA extraction

Samples were extracted using Qiagen chemistry, either the DNA Investigator Kit for the BioRobot® EZ1® Workstation, the QIAamp 96 DNA Swab BioRobot® Kit for the BioRobot® 8000 Workstation, or the manual QIAamp DNA Blood Mini Kit (Qiagen, Crawley, UK), according to manufacturer's instructions.

DNA quantification

Unless otherwise stated, DNA samples were quantified with PicoGreen® (Invitrogen, Paisley, UK) as described previously [5].

Amplification

Where DNA samples were diluted prior to amplification, a 1× Tris $EDTA^{-4}$ buffer ("DNA Suspension Buffer", Teknova, via Bioquote, York, UK) was used, unless otherwise stated. SGM+ (Applied Biosystems, Warrington, UK) amplifications were carried out using 1.5 ng DNA template measured by PicoGreen® or 2 ng if measured by Quantifiler® Duo in a total reaction volume of 25 μl, unless otherwise stated. Thermal cycling was performed in accordance with the manufacturer's instructions [9] for 28 cycles, except the final hold was performed at 4°C. ESI 16 (Promega, Southampton, UK) amplifications were performed according to the manufacturer's instructions [10]. All amplifications were performed on a MJ Research Tetrad thermocycler (Bio-Rad, Hemel Hempstead, UK).

Capillary electrophoresis and data analysis

For SGM+, 1.5 μl of PCR product was combined with 13.5 μl Hi-Di™ Formamide (Applied Biosystems, Warrington, UK) mixed in a 36:1 ratio with GeneScan™ HD400 ROX™ Size Standard (Applied Biosystems, Warrington, UK). Samples were denatured using an MJ Research Tetrad thermocycler at 90°C for 2 min prior to quick chilling on a crushed wet ice bath for at least 3 min prior to separation. Separation and detection were performed using an ABI PRISM® 3130*xl* Genetic Analyser (Applied Biosystems, Warrington, UK). SGM+ samples were injected at 1.5 kV for 10 s. ESI 16 PCR products were processed and separated according to the manufacturers' instructions [10], apart from injection parameters, which were set at 3 kV for 10 s following an initial sensitivity study. A detection threshold of 50 relative fluorescence units (rfu) was imposed during analysis using the GeneMapper® ID v3.2 software (Applied Biosystems, Warrington, UK). The terminology described by Gill et al. [11] was used to distinguish between adenylated amplification products ($n+1$) and non-adenylated products, (n).

Sensitivity and peak balance

In order to establish the sensitivity and working range of ESI 16 in our laboratory, a fully heterozygous donor blood sample was extracted, diluted in sterile distilled water (SDW), and amplified ten times at each of 14 template levels, 6,000, 3,000, 1,500, 1,000, 500, 250, 125, 62.5, 31.3, 15.6, 7.8, 3.9, 1.95 and 0.98 pg. Products were injected at the manufacturer's recommended 3 kV for 5 s and at 3 kV for 10 s.

Percentage profile and mean peak height was determined for each template level at both injection parameters. Profile balance was assessed by plotting mean peak height (rfu) against locus for the 500 pg recommended template level for the 3 kV for 10 s data set only. Peak height ratios were calculated (lowest peak height/largest peak height) for the 3 kV for 10 s injection data set only.

Limit of detection (LoD) was measured using 20 negative control samples from different amplifications, separated on a single 3130*xl* sequencer. The baseline fluorescence was measured using a 1 rfu analysis threshold for each dye used to separate sample amplicons. The standard deviation of these measurements was determined and multiplied by three to obtain the LoD [12].

Specimens, accuracy and precision

National Institute of Standards and Technology (NIST) Standard Reference Materials 2391b (Gaithersburg, MD, U.S.A.) were genotyped and the allele size was compared to the allelic ladder in order to assess accuracy.

Precision was measured using 22 allelic ladder samples run on each of three different 3130*xl* Genetic Analysers. Data was collected over a 1–4 month time frame. The base pair size range, mean, and standard deviation for each ladder allele were calculated for each Genetic Analyser, and then an average of the three datasets was calculated.

Repeatability and reproducibility

The ESI 16 within batch variation (repeatability) was measured using the data generated from the ten amplifications of 500 pg performed as part of the sensitivity experiments injected for 3 kV 10 s and compared with nine SGM+ amplifications of 1,500 pg injected at 1.5 kV for 10 s. The between batch variation (reproducibility) of ESI 16 was measured using 36 amplifications of the 9947a positive control DNA (1,000 pg) performed by five different operators on six different amplification instruments and three different 3130*xl* CE machines. Results were divided in to three 'test systems' by the location of the CE machine. The difference in the quality of results measured by total locus height was assessed by Mann–Whitney test with a 95% confidence interval.

ESI 16 profile characterisation

Eight fully heterozygous DNA samples were amplified five times at each of five template levels 15.6, 62.5, 125, 500 and 3,000 pg. A further four fully heterozygous DNA samples and three samples heterozygous at all but two loci were amplified five times at each of seven template levels 25, 50, 100, 200, 300, 400 and 500 pg. Profiles were annotated with size (bp), height (rfu) and area (rfu) data for all alleles and associated *n* peaks and stutter that exceeded a 50 rfu analysis threshold. Profile data were checked for accuracy and analysed using Minitab™ software (Minitab Ltd., Coventry, UK) and Matlab software (MathWorks, Cambridge, UK) in order to characterise the behaviour of the ESI 16 chemistry in our laboratory. Preferential amplification was assessed by dividing the lower molecular weight allele peak height by the larger molecular weight peak height. These values, transformed by $\log_{10}$ were plotted against the mean peak height of the allelic pair. Stutter proportion was calculated by dividing the stutter peak height by the peak height of the 'parent' allele.

Inhibitor spiking

Chemicals associated with common contaminants of casework samples were prepared in our laboratory and added to triplicate PCR reactions. Amplifications were performed in accordance with the manufacturers' recommendations [9, 10]. ESI 16 and SGM+ amplification reactions contained either 1,200, 900, 600, 300, 150 or 75 μM indigo (Fibrecrafts, Guildford, UK); 600, 400, 200 or 100 μM haematin (Sigma-Aldrich, Dorset, UK) or 240, 180, 120, 60, 30 ng/μl humic acid (Sigma-Aldrich, Dorset, UK).

Mixture studies

Mixture experiments were performed using DNA of known genotype and quantity. Five sample pairs (two male×female mixtures, two male×male mixtures and one female×female mixture) were used to ascertain the detection limit for the minor component when 500 pg total DNA was amplified in single reactions. Amplifications were performed using ESI 16 and SGM+ from mixture ratios 1:19, 1:14, 1:9, 1:4, 1:1, 1:0, 0:1, 4:1, 9:1, 14:1 and 19:1. Observed mixture ratios were calculated using peak area data for fully heterozygous loci of the mix pair and compared with the expected ratio. Percentage profiles were calculated and used to compare the relative success of ESI 16 with SGM+ results.

To ascertain detection limits for a range of total DNA templates, two male×female mix pairs were used at templates of 25, 50, 100, 250, 500 and 1,000 pg. Duplicate amplifications at mixture ratios of 1:9, 1:4, 1:2, 1:1, 2:1, 4:1 and 9:1 were performed with ESI 16 only.

Mock casework

A total of 51 mock casework samples were prepared, representing six cell types commonly encountered in forensic casework. Each sample type was prepared using three different consenting donors from an anonymised group comprising 11 male and 8 female donors.

Where samples were added to an item to simulate casework submissions, the item was first treated with 9,999 μJ/cm^2 UV using a CL-1000 UV crosslinker (UVP Ltd, Cambridge, UK) to reduce the incidence of background DNA contributing to the profiles.

Discrete 10 μl aliquots of liquid venous blood (with EDTA anti-coagulant) were spotted on to a square of 100% cotton material, a black leather shoe and a section of concrete paving, then allowed to air dry.

A 50 μl aliquot of saliva was spotted onto a square of material cut from a pair of 100% nylon tights and allowed to air dry.

Three volunteers provided soft drink cans from which they had drunk. The opening of the can was sampled with a cotton swab moistened with SDW.

Three volunteers provided an item of regularly worn jewellery. The surfaces of the items were sampled with a cotton swab moistened with SDW.

Three volunteers also provided a smoked cigarette butt, three plucked hairs and a sample of vaginal cells.

Blood on denim, semen on knickers, fingerprints on glass, telogen hairs, chewed gum and saliva on envelope samples were prepared as previously reported [5].

Where possible, 1.5 ng of extracted DNA was amplified per 25 μl reaction with SGM+ and 500 pg of DNA per 25 μl reaction amplified with ESI 16. In samples with insufficient DNA concentration including any zero quantification results, 10 μl of sample template was added to the amplification reaction. This is the maximum sample volume possible in a

25 μl SGM+ reaction. Following initial testing, seven low DNA template extracts were identified for further testing with ESI 16 alone to compare the results of amplifying 10 versus 17.5 μl extract, the maximum sample volume possible in a 25 μl ESI 16 reaction.

Degraded DNA

Samples were extracted, quantified using Quantifiler Duo® (Applied Biosystems, Warrington, UK) and amplified with ESI 16 and SGM+ as described previously [5], except that the injection conditions for ESI 16 were modified from the manufacturer's recommended protocol to 3 kV for 10 s.

Results

Sensitivity and peak balance

In our laboratory, between-instrument variability in terms of sensitivity has been observed with 3130*xl* Genetic Analysers. To ensure that results obtained mirror the kit performance reported by the manufacturer, a sensitivity test was performed.

Mean peak heights across all loci were calculated at each mass of DNA amplified (see Online Resource 1). Results demonstrated that in our laboratory, the 3 kV for 10 s injection of ESI 16 samples most closely resembled the peak heights achieved by the manufacturer [5], with saturation observed to occur at 3 ng templates and higher. Amplifications saturated by DNA are accompanied by a characteristic *n* peak present on amelogenin. The ESI 16 10 s injection data revealed a maximum fivefold increase in overall peak heights at templates around 250–500 pg compared to SGM+. The ESI 16 System generated 100% full profiles at template levels as low as 62.5 pg of DNA and between 81% and 97% profiles using 31.25 pg of DNA. By comparison, SGM+ consistently generated 100% full profiles at 500 pg DNA, between 95% and 100% profiles at 250 pg of DNA and between 86% and 100% profiles at 125 pg DNA. At the lower limit of the sensitivity scale, SGM+ amplified a mean 1% profile using 7.8 pg DNA, and no alleles at templates lower than this, whereas ESI 16 amplified a mean 3% profile at 0.98 pg (see Fig. 1).

Inter-locus balance using a fully heterozygous male DNA sample was examined across all dye channels for ESI 16 at 500 pg of DNA and SGM+ at 1,500 pg of DNA (see Online Resource 2). The difference in peak height between the most efficient locus and the least efficient locus is similar for the two multiplexes (2.5-fold and 2.7-fold difference for ESI 16 and SGM+, respectively). This illustrates that the 5-dye system used in ESI 16 performs as well as the four dyes employed in the SGM+ chemistry.

Intra-locus balance in a fully heterozygous male DNA sample was also examined. Mean peak height ratios (PHR) were calculated and their behaviour was observed with respect to template mass. Response of mean PHR to DNA template is shown in Online Resource 3. ESI 16 typically generated PHR values above 60%, even when using 62.5 pg of DNA, whereas SGM+ required 125 pg DNA to generate an average PHR of greater than 60%. An increase in the size of the error bars was observed with both multiplexes as template mass was reduced, but there was less variability in the PHR data at 125 and 62.5 pg with the ESI 16 System compared to SGM+.

Limit of detection calculations for each dye channel used to separate sample amplicons in ESI 16 are presented in Online Resource 4. Owing to restrictions in the GeneMapper® ID v3.2 software, a maximum of 50 data points can be written to the results table for each locus. The results suggest that an analysis threshold of 7 rfu could be applied in GeneMapper® in order to detect allelic peaks whilst filtering out 99% of the baseline. This would nevertheless, result in the automatic labelling of many non-allelic peaks, requiring deletion by the analyst. Binning the data illustrates that the frequency of baseline fluorescence falls to a minimal level around 14 to 16 rfu (see Online Resource 5). To inform the application of a minimum analysis threshold for interpretation of this system, 20 low template (ten 0.98 pg and ten 1.95 pg) sample profiles generated as part of the sensitivity study were analysed with a 10 rfu analysis threshold in GeneMapper®. Any peak smaller in size (bases) than the first bin of the first locus of each dye was disregarded. Of the remaining peaks, those attributed to baseline noise had a mean height of 10.4 rfu, with a maximum of 14 rfu (data not shown), which is of a similar level to that provided by the analysis of negative controls. Peaks with poor morphology of the expected profile or those not of the expected profile were analysed and demonstrated a range of 12 to 28 rfu with a mean of 17.5 rfu (data not shown). Alleles of good morphology from the expected profile were also analysed and demonstrated a range of 12 to 158 rfu with a mean of 47.7 rfu (data not shown).

Specimens, accuracy and precision

Accuracy was assessed by genotyping the ten NIST SRM 2391b DNA samples, all of which gave the expected profile. Allele peaks for each of the ten samples were also analysed for differences in base size compared to the corresponding allele in the allelic ladder. All alleles were found to fall within a ±0.5 base pair window of the corresponding allele in the ladder (see Online Resource 6), with vWA amplicons demonstrating the highest variance (−0.21 to +0.22 bases).

The precision study used allelic ladder data and revealed no single Genetic Analyser to have the most precise result

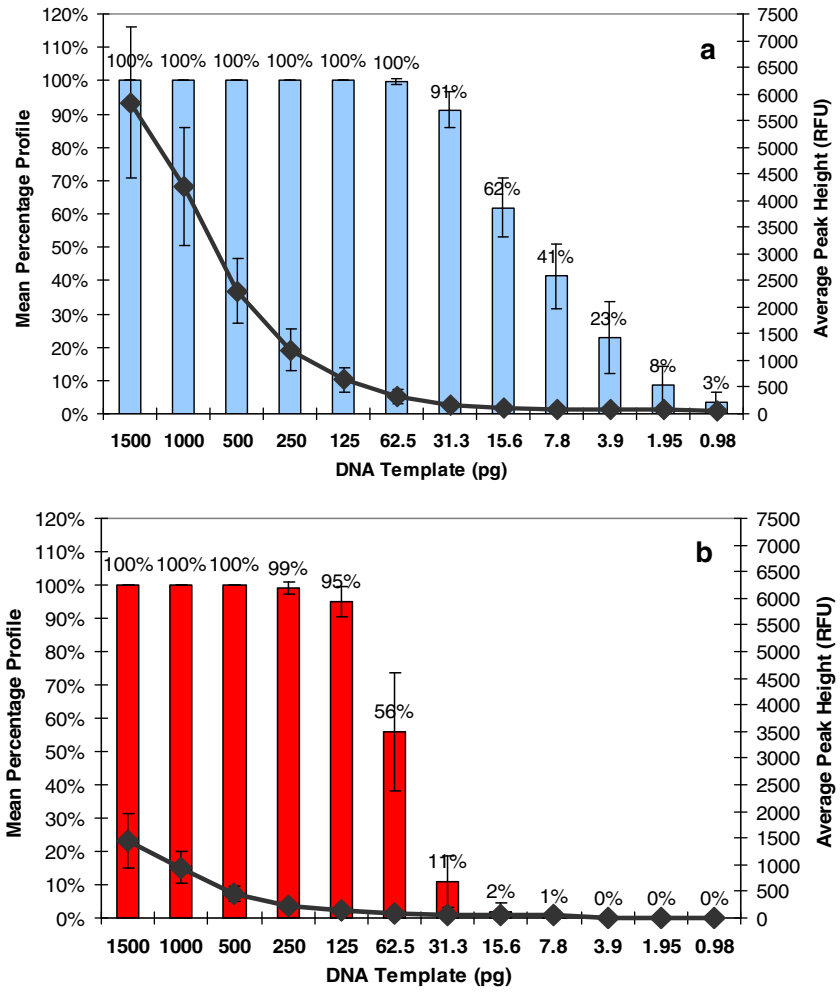

Fig. 1 Relative sensitivity of ESI 16 and SGM+mean peak height (right hand *Y*-axis) in rfu was determined across all loci for each STR multiplex and plotted against mass of DNA template (*line graph*). Percent full profile (left hand *Y*-axis) was also plotted against mass of DNA template (*bar graph*). Results are shown for ESI 16 samples injected at 3 kV for 10 s (**a**), and SGM+samples injected at 1.5 kV for 10 s (**b**)

for every locus of the ladder. On average, analyser #1 had the most precise result for 12 of the 16 loci. D16S539 was found to be the most consistent ladder locus with a mean 0.18 base pair variation (±0.05 base pairs), closely followed by vWA (0.2±0.04 base pairs). D2S441 was found to be the most variable ladder locus with a mean 0.45 base pair variation (±0.15 base pairs), closely followed by D2S1338 (0.41±0.15 base pairs). Although the two most consistent results were obtained for relatively low molecular weight loci, and the two most variable results were obtained for high molecular weight loci, precision does not appear to consistently correlate to allele size (see Table 1).

Repeatability and reproducibility

The samples assessed for repeatability were found to give the expected profile. Peak heights ranged from 834 rfu (D8S1179) to 4,115 rfu (D3S1358), with D22S1045 giving the greatest standard deviation of mean peak height (mean= 2,505±480 rfu). The amplicons of locus D22S1045 were also found to be the most variable in terms of base size in this data set, ≤0.21 base difference between the smallest and largest size recorded for a single allele (data not shown). Analysis of SGM+ sensitivity data from the 1,500 pg sample set revealed the amplicons of loci D18S51 and FGA to be the most variable in terms of base size, ≤0.24 base difference between the smallest and largest size recorded for a single allele (data not shown). SGM+ peak heights ranged from 517 rfu (FGA) to 3,145 rfu (D16S539), with D16S539 also giving the greatest standard deviation of mean peak height (mean= 1,997 rfu, SD=477 rfu). Comparing the results of D3S1358 amplicons (both loci are labelled with a form of FAM and occupy similar size ranges within both multiplexes), SGM+ mean peak heights (mean=1,607 rfu, SD=318 rfu) are roughly 60% of the height of ESI 16 mean peak heights (mean=2,716 rfu, SD=465 rfu).

The samples assessed for reproducibility were processed in three different laboratories, termed system 1, system 2 and system 3. All results were found to give the expected profile for each of the 36 replicates processed. Analysis of the data revealed that heterozygous loci, except TH01, demonstrated the low molecular weight (lmw) allele to have a higher average peak height compared with the average peak height of the high molecular weight (hmw) allele; however, this was determined

Table 1 Precision of ladder alleles

	LOCUS	Amelo	D3S1358	D19S433	D2S1338	D22S1045
	Size range (bp)	82.04–87.80	98.37–143.69	158.01–209.06	219.11–291.68	302.65–342.28
Mean size variance in base pairs	Analyser #1	0.19	0.21	0.22	0.30	0.29
	Analyser #2	0.22	0.27	0.28	0.58	0.42
	Analyser #3	0.27	0.29	0.21	0.35	0.33
	Mean	0.22	0.26	0.24	0.41	0.35
	SD	0.04	0.07	0.05	0.15	0.10
	LOCUS	D16S539	D18S51	D1S1656	D10S1248	D2S441
	Size range (bp)	78.10–127.66	131.0–212.13	220.82–268.79	280.61–323.71	340.77–377.26
Mean size variance in base pairs	Analyser #1	0.14	0.27	0.25	0.29	0.48
	Analyser #2	0.21	0.35	0.23	0.27	0.39
	Analyser #3	0.20	0.31	0.34	0.27	0.23
	Mean	0.18	0.30	0.26	0.29	0.45
	SD	0.05	0.08	0.08	0.08	0.15
	LOCUS	TH01	vWA	D21S11	D12S391	
	Size range (bp)	71.10–114.91	124.89–181.11	204.06–260.38	291.17–343.34	
Mean size variance in base pairs	Analyser #1	0.26	0.14	0.17	0.21	
	Analyser #2	0.29	0.25	0.27	0.24	
	Analyser #3	0.3	0.2	0.22	0.25	
	Mean	0.24	0.20	0.22	0.26	
	SD	0.08	0.04	0.06	0.06	
	LOCUS	D8S1179	FGA			
	Size range (bp)	77.06–126.64	145.31–293.5			
Mean size variance in base pairs	Analyser #1	0.25	0.24			
	Analyser #2	0.32	0.24			
	Analyser #3	0.35	0.38			
	Mean	0.29	0.35			
	SD	0.08	0.08			

Mean size variance (bases) of allelic ladder alleles at each locus calculated from 22 injections performed on each of three genetic analysers

not significant by Mann–Whitney test with a 95% confidence interval (data not shown). Analysis of the between-system reproducibility illustrated that there was a significant difference between system 1 and system 3, and between system 2 and system 3; where system 3 gave significantly reduced peak heights, see Online Resource 7. There was no significant difference between system 1 and system 2.

ESI 16 profile characterisation

Analysis of allele and locus drop-out at the lower DNA template levels (15.6–62.5 pg) within this data set demonstrates that there is no single locus most prone to drop-out. D8S1179 (15.6 pg), D22S1045 (25 and 50 pg) and D12S391 (62.5 pg) exhibit the highest rates of allele drop-out, whereas D8S1179 (15.6 pg), D22S1045 and D21S1045 (25 pg), D22S1045 (50 pg) and D10S1248 (62.5 pg) exhibited the highest rates of locus drop-out (data not shown).

Preferential amplification of the lmw allele of a heterozygous pair has been demonstrated where amplifications have generated peak heights in excess of 3,000 rfu. This generally arose in samples with excess DNA template (>500 pg). Below 3,000 rfu, the amplification of one allele in preference to the other is more variable (data not shown).

Forward and back stutters of up to two repeat units were observed within the data. Back stutter of a single repeat unit was the most commonly encountered form, occurring at DNA template levels of ≥50 pg (see Fig. 2). Whilst no locus was observed to stutter appreciably more than the others, TH01 appeared less prone to stutter compared with the other loci. D22S1045 presented the highest mean stutter proportional to the parent allele peak height (see Online Resource 8).

Back stutter of two repeat units occurred at DNA template levels ≥400 pg, most frequently from loci D3S1358, D1S1656 and D10S1248. The frequency of −2 repeat unit stutter was approximately 60-fold lower compared with −1 repeat unit stutter frequency (see Table 2).

Forward stutter was observed less frequently compared with back stutter. Single repeat unit forward stutter occurred

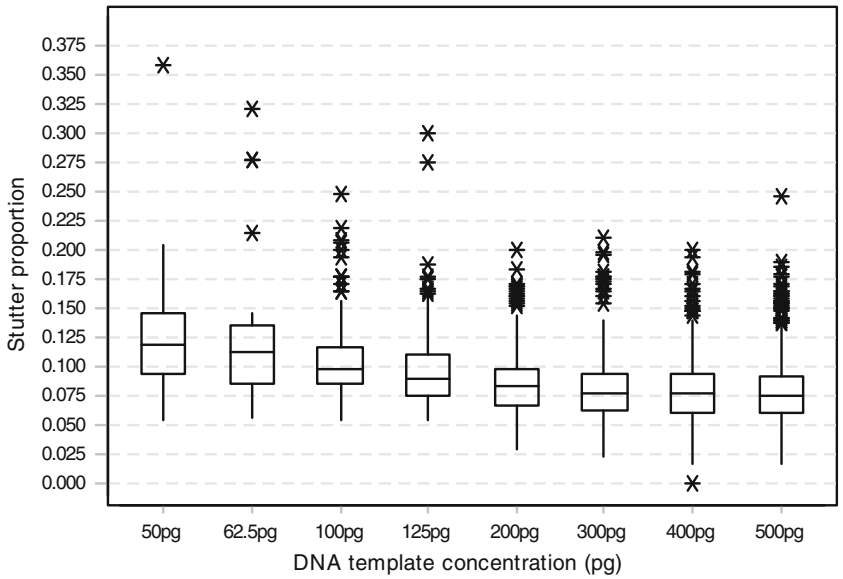

Fig. 2 −1 repeat unit stutter proportions plotted against DNA template concentration (pg) The 3,000 pg data set was excluded from this analysis

at DNA template levels ≥100 pg whereas +2 repeat unit stutter was only observed in the D3S1358 locus within the 3,000 pg DNA template data.

Three loci were observed to have *n* peaks associated with the higher template levels (100–3,000 pg): amelogenin, TH01 and D1S1656. Of these, the amelogenin X allele was most frequently observed to form an *n* peak.

In addition to the artefacts reported by the manufacturer [10], we have observed a number of other artefacts in our laboratory (see Online Resource 9).

Inhibitor spiking

Indigo

Compared with the control (no added inhibitors), the introduction of 75 μM indigo to SGM+ reactions affected the high molecular weight loci; D2S1338 and D18S51, in particular, exhibited reduced peak heights (data not shown).

Table 2 Stutter frequency calculated for 500 pg DNA template data

Locus	LMW frequency	HMW frequency	Combined frequency
−1 repeat	0.457	0.273	0.730
−2 repeat	0.008	0.005	0.012
+1 repeat	0.005	0.032	0.037
+2 repeat	0	0	0

Alleles (2,195) were analysed for stutter incidence in the ±1 and ±2 repeat unit positions relative to the parent allele. All autosomal loci were included in the analysis. Fewer high molecular weight allele stutters are recorded with back stutter owing to the opportunity for the low molecular allele to sit in the stutter position of the high molecular weight allele. Similarly, fewer low molecular weight allele stutters are recorded with forward stutter

Indigo (150 μM) almost entirely inhibited the SGM+ multiplex, with only amelogenin surviving. Higher concentrations of indigo completely inhibited SGM+.

ESI 16 results demonstrated no amplicon length-dependent effects for any of the indigo concentrations tested. Further, there was no overall decrease in mean peak height with increasing indigo concentration to 1,200 μM, compared with the control (see Fig. 3).

Haematin

SGM+ amplifications failed at all concentrations of haematin tested (see Fig. 3), mirroring recently reported data [5]. Previously unpublished work in our laboratory has shown SGM+ to give full profiles with 30 μM haematin and for complete inhibition to occur at 60 μM.

Full profiles were obtained from ESI 16 amplifications at all haematin concentrations tested. High concentrations of haematin caused the amelogenin alleles in ESI 16 to exhibit split peaks (*n* and *n*+1). *N* peak morphology changed from a ‘shoulder’ of the *n*+1 peak at 100 μM haematin, increasing in height with increasing haematin concentration until at 600 μM haematin the *n* and *n*+1 peaks were equal in height. A reduction in peak height from 1,600 rfu in the control to 400 rfu (600 μM haematin) was observed in D8S1179 and FGA loci with increasing haematin concentration. This contributed to a general decrease in the observed mean peak height as haematin concentration increased, compared with the control.

Sodium hydroxide is required to assist in the solubilisation of haematin. PCR reactions spiked with 600 μM haematin resulted in a sodium hydroxide concentration of 1.2 mM. This amount of sodium hydroxide alone is insufficient to cause any inhibition of SGM+ or ESI 16 (data not shown).

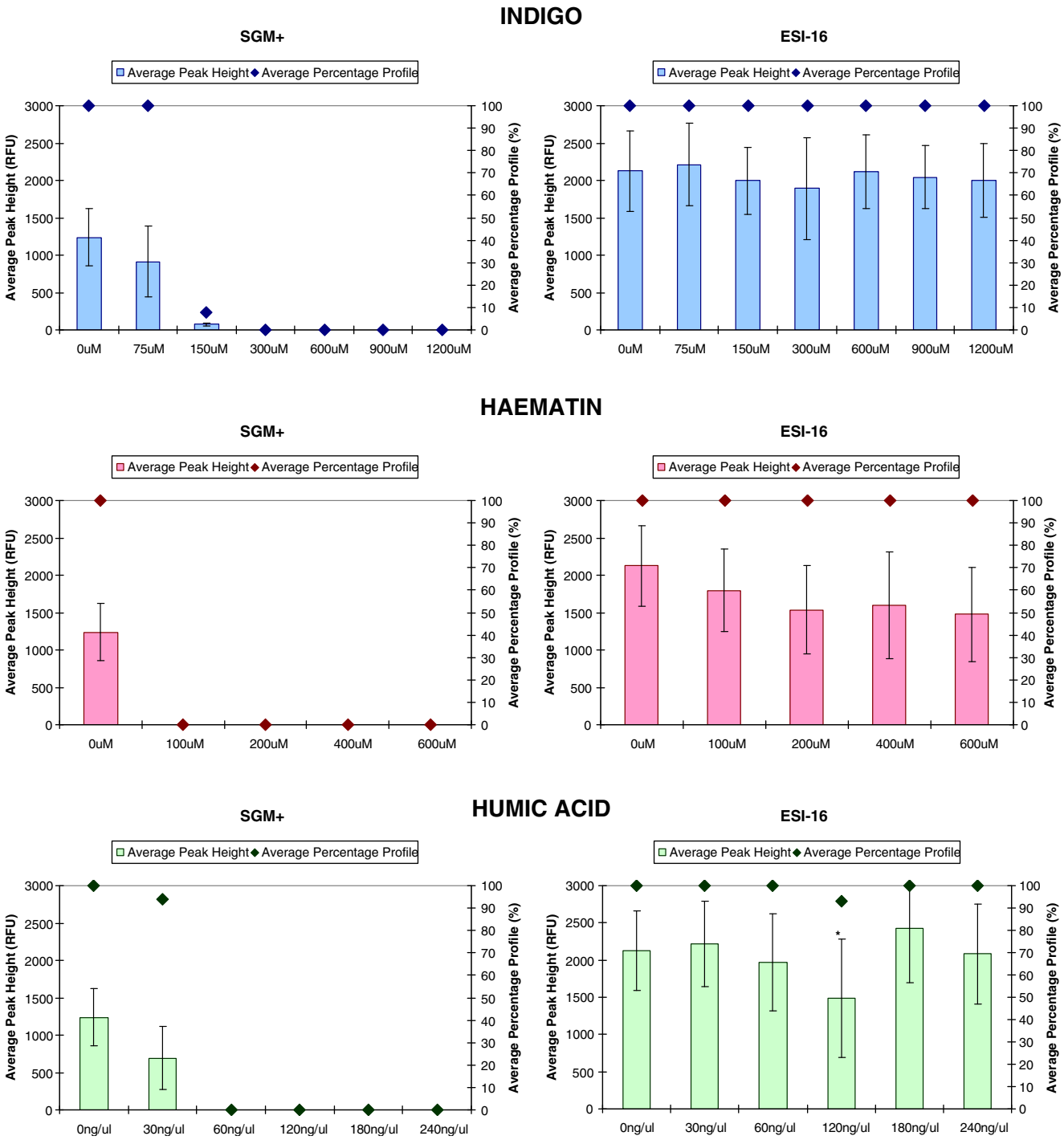

Fig. 3 Inhibition of AmpF*l*STR® SGM Plus® and PowerPlex® ESI 16 with indigo, haematin, and humic acid samples were amplified in triplicate with mean peak height and percentage profile scores calculated. *Error bars* represent standard error of the mean of the triplicate data. (*)Two of the three replicates gave 100% profiles (mean peak height of 1,909 rfu), the third replicate gave a 78% profile (mean peak height of 415 rfu)

Humic acid

SGM+ reactions spiked with 30 ng/μl humic acid produced a characteristic partial profile. D18S51 and FGA were prone to drop-out, with D3S1358 peak heights demonstrating an approximate 40% reduction in height compared with the neighbouring vWA and D16S539 alleles (data not shown). Humic acid concentrations in excess of 30 ng/μl gave no profile (see Fig. 3).

ESI 16 amplifications were capable of providing full profiles at all humic acid concentrations tested. Results indicate no overall decrease in mean peak height with increasing humic acid concentration compared with the control.

Mixture studies

ESI 16

Mixtures (1:4, 1:1 and 4:1) gave full profiles for major and minor components in all but one instance within the set of 500 pg total template mixtures tested. At the 1:9 ratio, the minor component (present at 50 pg) exhibited drop-out of up to one allele. The 1:14 or 14:1 and 1:19 or 19:1 ratios exhibited a drop-out of up to five and seven alleles, respectively, but both mix proportions gave at least one full profile for the minor component (present at 33 and 25 pg, respectively). The minor profile is visible but indistinguishable from stutter at the 1:19 and 19:1 ratios.

Mixture proportions (Mx) were calculated for loci exhibiting a full major and a full minor profile as described by Gill et al. [13]. The Mx observed for ESI 16 were more representative of the expected Mx compared with SGM+ (see Online Resource 10).

Where mixtures were tested at a range of total DNA template levels, the 1,000 pg data set demonstrated that all mixture ratios gave full profiles for the minor contributor, present at 100 pg in the 1:9 ratio. At the 500 pg template level (recommended by the manufacturer), the two person mixtures at the 1:9 ratio resulted in a partial profile (97%) for the minor contributor; this corresponds to a DNA template level of 50 pg. Mixture ratios of 1:1 failed to give full profiles for both contributors at 100 pg total DNA template and less. At the lowest total DNA template level tested (25 pg), the minor contributor in the 1:9 ratio, present at 2.5 pg, gave 15–20% profiles (data not shown).

SGM+

Only 1:1 mixtures (250 pg each donor DNA) gave full profiles for both donors in all five sample pairs within the set of 500 pg total template mixtures tested. At the 1:4 ratio, the minor component (present at 100 pg) typically exhibited drop-out of up to two alleles. At higher ratios, a full profile was not obtained for the minor component in any of the five sample pairs; this corresponds to a DNA input of 50 pg or less. Drop-out of up to 10, 12 and 14 alleles was observed for the 1:9 or 9:1, 1:14 or 14:1 and 1:19 or 19:1 ratios, respectively.

Interpretation of mixtures amplified with SGM+ was compromised by the high degree of drop-out observed, an expected result given the lower sensitivity of this multiplex.

Mx observed for SGM+ appeared to overestimate the Mx for minor components and underestimate the Mx for major components to a greater extent than ESI 16.

Mock casework

The majority of mock casework samples contained enough DNA to allow full profiles to be obtained from both STR systems tested. Where DNA content was limited, notably in the cellular samples obtained from jewellery and fingerprints deposited on glass and in the telogen hairs, the increased sensitivity of the ESI 16 system returned higher percentage profiles compared with the SGM+ result for the same sample (see Table 3). This increased sensitivity provided ESI 16 with enhanced first pass success results (see Online Resource 11).

Amplification of 17.5 μl extract from seven low template DNA samples revealed a general trend of increased percentage profiles compared with the amplification of 10 μl extract (see Online Resource 12). The increases do not represent a simple proportional increase in percentage profile in line with the quantity of extract added. While there was no adverse effect of adding more extract, a small increase in allelic drop-in was observed.

Degraded DNA

Compared with previously published data [5], the longer injection time used for ESI 16 in this study produced increased percentage profiles relative to results generated with the SGM+ kit in parallel. Previous data [5] demonstrated ESI 16 to provide a mean 38% improvement within the 11 UKNDL, equivalent to approximately eight alleles, compared with the SGM+ results. With the increased injection time, ESI 16 provided a mean 54% improvement within the 11 UKNDL, giving approximately 12 more alleles that could be searched against and loaded to the NDNAD compared with the SGM+ results.

Discussion

Any laboratory looking to validate a new multiplex with a view to offering it as a DNA profiling service would need to confirm if their existing interpretation guidelines fit the behaviour of the new multiplex. It is recommended that as part of a laboratory's forensic validation of any new multiplex, performing sensitivity and mixture studies, as well as processing a number of known DNA samples are required to provide statistical data necessary to inform the processing parameters (cycle number and CE injection settings) and interpretation processes suitable for individual laboratories to obtain comparable results. The study

Table 3 Results of mock casework samples exhibiting low DNA template

Sample	Donor	Quant (ng/μl)	SGM+		ESI 16		
			DNA template amplified	Percentage profile	DNA template amplified	Percentage profile	Percentage SGM+ profile
Cellular (watch)	♂ F	0.2534	1,500 pg	86	500 pg	100	100
Cellular (watch)	♀ F	0.0105	105 pg	64	105 pg	86	95
Cellular (ring)	♂ G	0.0025	25 pg	0	25 pg	63	65
Cellular (fingerprints)	♀ D	0	nc	0	nc	13	14
Cellular (fingerprints)	♂ F	0	nc	0	nc	76	76
Cellular (fingerprints)	♀ F	0	nc	100	nc	100	100
Telogen hair	♂ D	0	nc	0	nc	0	0
Telogen hair	♀ F	0	nc	11	nc	75	74
Telogen hair	♂ H	0	nc	43	nc	97	100

Samples were quantified with PicoGreen®. Where no quant value was obtained, no template quantity was calculated (nc) for amplification; in these cases, a volume of 10 μl of extract was added to the PCR. Alleles were called where they exceeded a 50 rfu analysis threshold

described here has demonstrated that ESI 16 is robust and fit for purpose. It is more robust to known inhibitors and more sensitive than SGM+, the system used routinely throughout the UK.

The greater sensitivity exhibited by the ESI 16 system compared to SGM+ could be attributable to three factors; the higher cycle number, the higher injection voltage, and the improved chemistry of the new kit. This is offset to some extent by the lower DNA template level recommended for ESI 16 amplifications compared with the quantity of DNA template used in SGM+ reactions but is not fully compensated for by the ESI 16 settings we have chosen to employ. There are several other factors which may also influence peak heights, such as PCR efficiency and the efficiency of the dyes used to label the fragments, but these are more difficult to estimate. The extent of the improved sensitivity ranges between twofold (illustrated by the difference in the DNA template level at which the last full profile is obtained), and eightfold (illustrated by the difference in the lowest mass of DNA that consistently yields a full profile) above the current SGM+ sensitivity. Unpublished data from a separate study performed in our laboratory demonstrated that the step change in sensitivity experienced with the ESI 16 System lay midway between that of standard SGM+ amplification and the increased cycle amplification of the Low Copy Number (LCN) method (see Online Resource 13). This indicates that there is scope to develop an enhanced sensitivity method for ESI 16.

As DNA template levels are reduced, a concomitant drop in peak height and the mean PHR with increasing standard deviation (SD) was observed and is a previously documented phenomenon [14–18]. PHR increased at the lower template levels with both the SGM+ and ESI 16 results and is attributed to stochastic effects of DNA sampling and amplification. It appears that PHR was not affected at DNA template levels above 62.5 pg with ESI 16, whereas the PHR became an issue for SGM+ at 125 pg, below which point the mean PHR declined rapidly. This difference in response could be attributed to improved efficiency of the ESI 16 chemistry and warrants further investigation.

The limit of detection may vary between laboratory, between different CE models and indeed between CE machines of the same model. The LOD data analysed for our study was generated by a single 3130*xl* genetic analyser, thus the recommendation of a minimum analysis threshold must include a measure of conservatism and be robust to the degree of instrument variation a laboratory observes between its CE machines.

The reproducibility test used three test systems, each with their own CE machines. Results demonstrated a significant difference between test systems 1 and 2 relative to 3: system 1 CE machine gave the lowest mean peak heights, system 2 CE machine gave 2.2-fold higher mean peak heights and the system 3 CE machine gave 1.8-fold higher compared to system 1. This observation does not support CE instrument variability as the main cause of the disparity. Some other element of the systems must be the source, such as PCR instrument or operator/robotic protocol; systems 1 and 2 are very similar manual systems, whereas system 3 utilised a robotic platform to prepare PCR reactions and CE plates.

The loci observed to have the most size variable amplicons in SGM+ (D18S51) and ESI 16 (D22S1045) both occupy positions at the high molecular weight end of the multiplex profile. The most size variable ladder alleles in ESI 16 were found in D2S441 and D2S1338, also relatively high molecular weight loci. Koumi et al. [19]

reported deterioration of precision at the high molecular weight end of profiles in a study using three different capillary array electrophoresis instruments. Their data also show that this is not a consistent correlation, an observation in parallel with the data presented herein.

Humic acid is reported to bind to DNA, affecting the template availability for amplification, whereas haematin is thought to affect *Taq* polymerase function [20]. The ESI 16 chemistry facilitates the amplification of STR fragments in the presence of both of these inhibitors at high concentrations, exhibiting a degree of robustness not seen with SGM+. It is questionable whether such concentrations of inhibitor would be found in forensic samples extracted by contemporary methods; however, a system exhibiting good robustness with known inhibitors could indicate a similar performance with samples carrying unidentified inhibitors.

Excessive amounts of DNA or the presence of a PCR inhibitor both provide an environment where ESI 16 will produce *n* peaks at the amelogenin alleles. This tendency was observed with both 3 and 6 ng of DNA added to an ESI 16 amplification (6 and 12 times the recommended template level, respectively) as well as in reactions spiked with haematin.

The mixture study demonstrated that the estimation of Mx from an ESI 16 profile followed expected behaviour. Further, ESI 16 demonstrates an improved adherence to the expected behaviour at the more extreme mixture ratios compared with SGM+. This may be linked to the ability of ESI 16 to provide more balanced PHRs at lower DNA template levels and could be a function of either or both the PCR efficiency or stability/availability of the template in the reaction mix. The product of these effects may be a reduction in the stochastic nature of the amplification reaction that would be most noticeable at reduced amounts of DNA template. Where mixtures were tested at a range of total DNA template levels, percentage profile results for each contributor mirrored the results of the sensitivity study. An amplification of the recommended 500 pg template, containing two contributing DNA sources can be resolved with full profiles at mix ratios up to 1:4. At other total DNA template levels, contribution of 50 pg or less resulted in the production of partial profiles.

Our mocked casework samples included a range of cell types and sample substrates commonly encountered in forensic casework. Among these, seven low template samples illustrated the effects of the increased sensitivity that the ESI 16 system confers. These samples also illustrated that quantity of extract (DNA template) and the resulting percentage profile or mean peak height does not follow a proportional linear relationship at low template levels. This supports our observations for variability and stochastic effects on PHR at low template levels. It is expected that the use of ESI 16 will bring the most benefit to the analysis of this low template type of sample, along with the degraded and inhibited samples that are encountered in forensic casework, when compared with standard SGM+ processing. It would likely reduce but not eliminate the use of low template DNA profiling approaches.

Conclusions

The developmental validation studies presented here illustrate the improved performance of the ESI 16 multiplex compared with standard SGM+ processing. However, it will not fully provide a replacement for LCN 34 cycle amplification methodology but will likely reduce the number of samples submitted to this process.

The ESI 16 System will provide reproducible results between laboratories where the methodology is comparable. It is recommended that laboratories wishing to implement this system should perform in-house studies to define appropriate interpretation guidelines in their hands. We found it to provide an acceptable level of accuracy and precision whilst exhibiting a comparable degree of repeatability to SGM+.

The more comprehensive allelic ladder provided with ESI 16 will assist forensic practitioners in the UK to confidently designate more of the alleles designated as “off ladder” using the National DNA Database® guidelines for SGM+ [21].

Deployment of this multiplex is anticipated to improve the success rates of standard casework samples that contain low DNA template levels as up to 17.5 μl of sample may be amplified in a 25 μl reaction. This has the potential to be maximised to 35 μl in a 50 μl reaction if desired and could prove to be a highly beneficial flexibility of the ESI 16 chemistry. Use of this multiplex is also expected to benefit ‘difficult’ samples that are affected by the co-extraction of PCR inhibitors or contain degraded DNA. The increased discrimination power compared with SGM+ will also assist in paternity testing and complex relationship elucidation. It may also prove to be a valuable new tool in the review of ‘cold cases’ where existing methods and technologies have been unable to provide fresh information to assist the resolution of unsolved cases.

Acknowledgements The authors would like to acknowledge the generosity of Promega Corporation, with special thanks to Dr. Doug Storts, Dr. Bob McLaren, Cindy Sprecher and Dawn Rabbach. Thanks also to FSS staff: Gareth Stead, Kate Lomas, Marcus Adams, Emily Rowlands, Laura Pelleymounter, Rebecca Lewis, Alison Peters, Nina Fairless and Tarek Leane for technical assistance with this work.

Ethical standards Experiments contained herein were performed in the UK and comply with current laws. All samples were collected with fully informed consent.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. The Crown Prosecution Service, B4. Adventitious (chance) DNA matches. Available at http://www.cps.gov.uk/legal/s_to_u/scientific_evidence/adventitious_dna_matches/. Accessed 15 December 2009
2. Gill P, Fereday L, Morling N, Schneider PM (2006) The evolution of DNA databases—recommendations for new European STR loci. Forensic Sci Int 156:242–244
3. Gill P, Fereday L, Morling N, Schneider PM (2006) New multiplexes for Europe—amendments and clarification of strategic development. Forensic Sci Int 163:155–157
4. Whitaker JP, Clayton TM, Urquhart AJ, Millican ES, Downes TJ, Kimpton CP, Gill P (1995) Short tandem repeat typing of bodies from a mass disaster: high success rate and characteristic amplification patterns in highly degraded samples. Biotechniques 18:670–677
5. Tucker VC, Hopwood AJ, Sprecher CJ, McLaren RS, Rabbach DR, Ensenberger MG, Thompson JM, Storts DR (2010) Developmental validation of the PowerPlex® ESI 16 and PowerPlex® ESI 17 Systems: STR multiplexes for the New European Standard. Forensic Sci Int Genet. doi:10.1016/j.fsigen.2010.09.004
6. Scientific Working Group on DNA Analysis Methods (SWGDAM) (2004) Revised validation guidelines. Forensic Sci Commun 6. Available at http://www2.fbi.gov/hq/lab/fsc/backissu/july2004/standards/2004_03_standards02.htm. Accessed 12 January 2009.
7. DNA Advisory Board (2000) Quality assurance standards for forensic DNA testing laboratories. Forensic Sci Commun 2. Available at http://www2.fbi.gov/hq/lab/fsc/backissu/july2000/codis2a.htm. Accessed 12 January 2009
8. ENFSI (2010) Recommended minimum criteria for the validation of various aspects of the DNA profiling process. Available at http://www.enfsi.eu/get_doc.php?uid=630. Accessed 03 March 2011
9. Applied Biosystems (2006) AmpF*l*STR® SGM Plus® PCR Amplification Kit User's Manual. Applied Biosystems Part Number 4309589D, Foster City, CA. Available at http://www3.appliedbiosystems.com/cms/groups/applied_markets_support/documents/generaldocuments/cms_041049.pdf
10. Promega Corporation (2009) Technical manual #TMD027-PowerPlex® ESI 16 System, Madison, WI. Available at http://www.promega.com/tbs/tmd027/tmd027.html
11. Gill P, Urquhart A, Millican E, Oldroyd N, Watson S, Sparkes R, Kimpton CP (1996) A new method of STR interpretation using inferential logic—development of a criminal intelligence database. Int J Leg Med 109:14–22
12. Eurachem (1998) The Fitness for purpose of analytical methods – a laboratory guide to method validation and related topics. English edition 1.0. Available at http://www.eurachem.org/guides/pdf/valid.pdf. Accessed 24 January 2007
13. Gill P, Sparkes R, Pinchin R, Clayton T, Whitaker J, Buckleton J (1998) Interpreting simple STR mixtures using allele peak areas. Forensic Sci Int 91:41–53
14. Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. Forensic Sci Int 112:17–40
15. Whitaker JP, Cotton EA, Gill P (2001) A comparison of the characteristics of profiles produced with the AmpF*l*STR® SGM Plus™ Multiplex System for both standard and low copy number (LCN) STR DNA analysis. Forensic Sci Int 123:215–223
16. Holt CL, Buoncristiani M, Wallin JM, Nguyen T, Lazaruk KD, Walsh PS (2002) TWGDAM validation of AmpF*l*STR™ PCR amplification kits for forensic DNA casework. J Forensic Sci 47:66–96
17. Opel KL, Chung DT, Drábek J, Butler LM, McCord BR (2007) Developmental validation of reduced-size STR miniplex primer sets. J Forensic Sci 52:1263–1271
18. Gilder JR, Inman K, Shields W, Krane DE (2011) Magnitude-dependent variation in peak height balance at heterozygous STR loci. Int J Leg Med 125:87–94
19. Koumi P, Green HE, Hartley S, Jordan D, Lahec S, Livett RJ, Tsang KW, Ward DM (2004) Evaluation and validation of the ABI 3700, ABI 3100, and the MegaBACE 1000 capillary array electrophoresis instruments for use with short tandem repeat microsatellite typing in a forensic environment. Electrophor 25:2227–2241
20. Opel KL, Chung D, McCord BR (2010) A study of PCR inhibition mechanisms using real time PCR. J Forensic Sci 55:25–33
21. NPIA, Rules for designation of alleles for inclusion on the National DNA Database. CUSTP-GS-004, issue 3 (20 August 2010)