Forensic Science International: Genetics 4 (2010) 158–167

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Identification of racehorse and sample contamination by novel 24-plex STR system

Jin-Wen Chen [a], Cornelius E. Uboh [a,b,*], Lawrence R. Soma [a], Xiaoqing Li [a], Fuyu Guan [a], Youwen You [a], Ying Liu [a]

[a] University of Pennsylvania School of Veterinary Medicine, New Bolton Center Campus, 382 West Street Road, Kennett Square, PA 19348, USA
[b] Pennsylvania Equine Toxicology & Research Center, West Chester University, 220 E Rosedale Avenue, West Chester, PA 19382, USA

ARTICLE INFO

Article history:
Received 14 August 2008
Received in revised form 4 August 2009
Accepted 4 August 2009

Keywords:
Equine
Human
STR loci
Multiplex PCR
Probability of identity

ABSTRACT

Proper identification of racehorses competing in an official race and maintenance of defensible chain of custody are important in doping control regulations. The purpose of this study was to develop a reliable multiplex PCR method for providing genetic evidence for matching donors to test samples by using short tandem repeat (STR) loci. Amplification of 21 STR loci from blood, urine or hair root was achieved in a single tube and STR length polymorphism was analyzed using fluorescent labeled capillary electrophoresis. This novel approach showed an allele confidence interval of 0.19–0.43 bp and size estimation error of 0–0.48 bp. In 90 thoroughbred (TB) and 171 standardbred (STB) horses, the method was highly discriminating and reproducible with probability of false identification of 1 in 10^{11} (TB) and 1 in 10^{13} (STB). All loci were highly polymorphic with an average probability of identity of 0.18 (TB) and 0.13 (STB), heterozygosity of 0.65 (TB) and 0.68 (STB), and polymorphism information content (PIC) of 0.62 (TB) and 0.69 (STB). The highest allele frequency also reflected the degree of polymorphism due to high correlation with PIC. To obtain evidence of sample tampering with human material, three human specific STR markers were included in the panel. This method is the first in the horseracing industry, specifically designed for racehorse identification and detection of equine sample contamination by human DNA.

Published by Elsevier Ireland Ltd.

1. Introduction

Short tandem repeat (STR) loci technique has been used in equine paternity testing [1–3] and identification of individuals [4]. However, sample identification by STR is uncommon in routine testing of forensic samples in the horseracing industry. Positive drug test results have been legally challenged. Blood and urine are collected by Commission employees and witnessed by the horse trainer or representative. Samples are transported to the laboratory for drug screening. However, drug test results are sometimes challenged for possible breach in the chain of custody; therefore, sample identification to verify chain of custody under these circumstances is critical. Individual identification by STR typing with appropriate statistical analysis should resolve any dispute over identity of sample and donor. For this reason, genotyping a sample collected from a doped horse has become an integral component of equine forensic testing in Pennsylvania (PA). Fewer than ten STR markers might be sufficient for drawing a positive conclusion on a test sample. However, a method with more than ten markers increases the likelihood of obtaining a sufficiently large profile for individual identification even in cases where low-yield or degraded samples prevent some markers from being genotyped. Some markers cannot be accurately genotyped if samples are degraded, have low DNA integrity or have mutations in the primer binding site [1].

Probability of identity (PI) is the likelihood of observing two identical genotypes from two distinct individuals in a population and, is a measure of the power of genetic markers to distinguish individuals with a reasonable level of confidence [5,6]. This calculation for PI incorporates inbreeding coefficient, F_{IS} and F_{ST}, rendering a more accurate estimation of PI for related individuals [7,8]. In this study, PI and other polymorphism measures were employed to evaluate the variability of STR loci and to estimate the probability for sample matching or exclusion in racehorses.

Abbreviations: STR, short tandem repeat; bp, base pair; TB, thoroughbred; SB, standardbred; CE, capillary electrophoresis; PIC, polymorphism information content; IACUC, Institutional Animal Care and Use Committee; PA, Pennsylvania; PAR, peak area ratio; PCR, polymerase chain reaction; PI, probability of identity.

* Corresponding author at: University of Pennsylvania School of Veterinary Medicine, New Bolton Center Campus, 382 West Street Road, Kennett Square, PA 19348, USA. Tel.: +1 610 436 3501; fax: +1 610 436 3504.
E-mail address: ubohcorn@vet.upenn.edu (C.E. Uboh).

1872-4973/$ – see front matter. Published by Elsevier Ireland Ltd.
doi:10.1016/j.fsigen.2009.08.001

Equine samples could be contaminated with DNA from other species, particularly human, during pre-or post sample collection despite implementation of chain of custody procedures. Sample contamination may occur due to human error or willful supplementation of insufficient equine urine sample with that of human by the sample collector. Racehorse urine samples are collected by Commission employees other than veterinarians. In fact, the Commission employee who collected post-race urine sample from a racehorse later admitted to adding his urine to meet the minimum sample volume required after he could no longer wait for the horse to urinate (personal communication with the PA Racing Commission, 2007). This problem creates difficulties in defining the true nature of the sample and in identifying the true donor of the sample. This seemingly insurmountable problem in the horseracing industry could be overcome if genotyping of equine test samples incorporated human DNA testing as a first step approach. This paper presents, for the first time, a novel 24-plex STR typing method to identify equine samples and detect possible contamination by human DNA. In addition, the present study established allele frequency databases of 21 STR loci in both TB and STB horses competing in PA and evaluated the usefulness of loci in sample identification by comparing the degree of genetic polymorphism.

2. Materials and methods

2.1. Sample collection

Post competition blood and urine samples were collected for STR polymorphism analysis in 90 TB and 171 STB Pacer horses competing at five racetracks in PA. Samples were collected within a period of six days to avoid repeated use of samples from the same horse. Samples were routinely collected from winners in each race and a random selection of non-winners in the same race. Blood samples were collected in the presence of an anti-coagulant, potassium oxalate, and sodium fluoride to inhibit plasma esterase activity [9]. Samples were refrigerated at the tracks but transported at ambient temperature to the laboratory. Upon arrival and receipt of the samples, they were immediately stored at 4 °C. In a separate experiment, blood, urine and a single hair root samples were collected 3×, at an interval of two weeks, from the same nine TB horses housed at the University of Pennsylvania School of Veterinary Medicine. Samples were collected according to IACUC Guidelines and were analyzed for STR markers by multiplex PCR to compare their genotypes.

To test the multiplex, a total of 3508 DNA isolates were prepared from TB and STB horses racing in PA within a period of six months and from research horses at the Large Animal Hospital of the University of Pennsylvania School of Veterinary Medicine. Many of those isolates were either from samples obtained from the same horse that may have competed multiple times during a racing season or from different types of samples collected from the same animals. Although 3508 DNA isolates were tested by multiplex PCR, only 261 of these isolates were employed for STR polymorphism analysis and allele frequency database to avoid repetitive use of the same sample.

2.2. DNA sample preparation

DNA samples were prepared from whole blood using the Puregene Genomic DNA Purification Kit (Gentra Systems, USA) [10,11], and from urine and hair root by the Genorise DNA Purification Kit (Genorise Scientific, USA). DNA concentration was estimated by measuring OD_{260} using a BioPhotometer (Eppendorf, Germany) and verified by quantitation of DNA band on ethidium bromide stained agarose gel (0.8%) using Totallab TL100 software (Non-linear Dynamics, USA) [10].

2.3. Selection of STR markers

A total of 39 equine dinucleotide repeat markers including 12 loci in the International Society for Animal Genetics (ISAG) panel [4] were initially tested for PCR amplification and 21 loci including 8 in ISAG panel were chosen based on variability, allele size range, and ease of co-amplification of markers. Four other ISAG panel loci were not included in the final 21 panel due to difficulty in co-amplification. These 21 markers were distributed on 13 equine chromosomes [4] (http://locus.jouy.inra.fr, Table 1) and comprised: VHL20 [12], HTG4 [13], AHT4 [14], HMS1, HMS2, HMS3, HMS6, HMS7 [15], ASB2, ASB9 [16], COR045 [17], LEX073, LEX074 [18], COR008, COR018 [19], UM015 [20], UMNe116 [21], UMNe156, UMNe191, UMNe222 [22], and UMNe479 [23]. To detect human DNA, three human STR markers comprising TH01[24], TPOX [25], and D18S51 [26] were incorporated into the method.

2.4. PCR optimization and co-amplification of equine and human STR markers

The primer pairs to amplify the core sequence of dinucleotide repeats including VHL20 [12], HTG4 [13], HMS6 and HMS7 [15] were previously described and the primer pair for AHT4 was described elsewhere [14] with minor modifications (Table 1) [10]. The forward primers for UMNe156 [22], ASB2 [16], HMS3, HMS2 [15] and UM015 [20] were also previously described. The paired PCR primers for the other markers were designed to amplify the entire STR sequence for equine and human loci (Table 1) and tested for primer–dimer and hairpin by Autodimer [27]. PCR was established using HotStar Taq Master Mix (Qiagen, USA). Each primer pair was tested to achieve a single PCR product with an expected size and all primers were mixed to amplify all loci in a single tube. Optimized PCR conditions were determined to demonstrate equal amplification of all loci according to product specificity, quantity, and minimization of stutter activity. PCR conditions were optimized for the most suitable combinations of DNA template, primer and magnesium concentrations, annealing temperature, and cycling number. DNA template was titrated to determine the minimum quantity of DNA needed for successful multiplex amplification. Primer concentration was minimized to achieve efficient amplification and fidelity of loci.

Optimized conditions for multiplex PCR amplifications were established to amplify 21 equine and three human STR loci in a 15 μl reaction volume as follows: 1× HotStarTaq DNA polymerase Master Mix, 3.0 mM $MgCl_2$, 20–50 ng DNA template, and appropriate primer concentrations (Table 1). PCR amplification was subjected to the following cycling conditions in a TC-512 Thermal Cycler (Techne, USA): 95 °C for 15 min, 30 cycles of 30 s at 94 °C, 30 s at 57 °C and 1 min at 72 °C, with final extension at 72 °C for 10 min, and tested in a total of 3508 DNA isolates from TB and STB horses.

2.5. Size analysis of multiplex PCR products on CEQ8800 genetic analysis system

One primer of each locus was labeled with one of the following fluorescent dyes: D2, D3, D4 and Cy5 (Table 1). Aliquots of 7 μl PCR products were initially analyzed by DNA electrophoresis on 1.2% agarose gel. PCR product was then analyzed for fragment size by capillary electrophoresis (CE) on CEQ8800 Genetic Analysis System (Beckman Coulter, USA). The parameters for CE were as follows: sample was denatured at 90 °C for 120 s, electrokinetically injected at 2.0 kV for 30 s, and separated at 4.8 kV for 53 min at 50 °C. The size of the amplified STR fragments was automatically determined against the DNA size standard and all markers were within the range of DNA ladder sizes. Alleles were identified as the highest

Table 1
Equine and human (*) STR loci, chromosomal distribution[a], GenBank accession, dye assignment, sequences and concentrations (nM) of PCR primers, and PCR product size (bp).

STR	Chromosome	GenBank	Dye	Primer 1	Primer 2	nM	Size
VHL20	30	X75970	D4	caagtcctcttacttgaagactag	aactcagggagaatcttcctcag	70	90–110
UMNe156	15	AF536268	D2	agactagcttcaaattgcccc	cctatgcttgaaggagtgtg	90	102–124
HTG4	9	AF169165	D3	ctatctcagtcttcgttgcaggac	ctccctccctccctctgttctc	40	128–140
AHT4	24q14	Y07733	D2	aaccgcctgagcaaggaagt	tcccagagagtttaccctgg	90	146–162
HMS6	4	X74635	D3	gaagctgccagtattcaaccattg	ctccatcttgtgaagtgtaactca	70	160–170
HMS7	1q25	X74636	D4	caggaaactcatgttgataccatc	tgttgttgaaacataccttgactgt	70	175–185
ASB9	10q21–q23	X93523	D2	tttcctcctccactacacac	tcccttatcacaatagagcag	100	201–215
ASB2	15q21.3–q23	X93516	D3	ccttccgtagtttaagcttctg	cttccccagaagttattttgc	90	190–226
COR045	10p13dist	AF108362	D4	taccgcaagtgaaaccagttc	ttgtgggactgagcccttaac	70	217–233
HMS3	9	X74632	D2	ccaactctttgtcacataacaaga	catcagtcagaagctgcgaacc	90	243–265
UMNe222	15	AF536300	D3	accaagctatgagtcaggag	agcatcttcatgtcctctgc	60	249–267
LEX074	24	AF213360	D4	ccccttaaattcagaaagagagcc	ggaatttggagattatctgtgggc	15	273–291
TH01*	11p15.5	D00269	D2	ctgttcctcccttatttccc	ggtacctggaaatgacactg	100	281–297
COR008	9	AF083451	D3	aggacactgaagggctgaaag	tagatagcgtctggagggttc	90	297–321
HMS2	10	X74631	D2	acggtggcaactgccaaggaag	gatctctagctcagtaaatcacagg	100	305–325
UM015	6	AF195133	Cy5	agtctggctgaggatactg	tttgctgtcacattagaggg	40	330–346
COR018	25	AF083461	D2	tgagtcttctgtactcctgg	ccacatctgggagtactaga	90	343–368
UMNe116	11	AY735236	D3	ctggctaaactcttattcc	acatgggagaaaatacacac	90	364–378
UMNe479	8	AY731401	D4	gagatggatggaatagcttg	tgcccagcctgaaagatttc	60	377–385
HMS1	15	X74630	D2	cacttatcagagaggccctcc	gtcatcccactctatcagggg	90	399–409
TPOX*	2p25.3	M68651	D2	cacagcttgatctcctcatg	tgaactcctcaggtccaatc	100	425–445
UMNe191	12	AF536279	D3	tgtcctcacttggcatgagtcc	ccagatggtgaaacaaggggc	100	414–442
LEX073	19	AF213359	D4	ttcagaacatcatccagcatcccc	cccaccactcaaatgtactaggc	40	462–492
D18S51*	18q21.33	X91254	D3	actgcacttcactctgagtg	cactttagccgacaaaaggc	100	470–494

[a] http://locus.jouy.inra.fr, http://www.cstl.nist.gov/biotech/strbase/.

peak and the second highest peak that was neither a stutter nor a plus A. Missed and miscalled alleles were manually corrected. Alleles were designated as the number of repeat units plus 0.1 for a partial repeat (one nucleotide) as determined by DNA sequencing [28].

2.6. Allelic band purification and DNA sequencing

PCR products were separated on 6% polyacrylamide non-denaturing gel as previously described [29] and visualized by silver staining. Briefly, the mixture of PCR products for each locus was electrophoresed on 6% polyacrylamide gels containing 7 M urea (gel size: 20 cm × 45 cm × 0.04 cm) in T-Rex™ Aluminum Backed Sequence system (Model FB-SEQ-2045, Fisher Scientific, USA). Electrophoresis was conducted at 800 V for 2 h to separate the allelic bands from those of stutters. Following electrophoresis, allelic and stutter fragments bands were visualized by silver staining (Silver Sequence™ DNA staining reagents, Promega, USA). Allelic band was excised and purified by MiniElute Gel Extraction Kit (Qiagen, USA). Single PCR product was re-amplified prior to purification by Qiaquick PCR Purification Kit. All alleles were sequenced by GenomeLab™ Dye Terminator Cycle Sequencing (DTCS) Quick Start Kit (Beckman Coulter, USA). Briefly, DNA sequencing reaction was assembled in 20 μl including 80 fmol purified PCR product, 3.2 pmol primer 1 or 2 (Table 1) and 8 μl DTCS Quick Start Master Mix followed by 30 cycles of 20 s at 96 °C, 20 s at 55 °C and 4 min at 60 °C. Sequencing product was then purified by ethanol prior to sequencing in the CEQ8800 Genetic Analysis System.

2.7. Determination of peak area ratios

Peak area ratios (PAR) between two heterozygous alleles and those between stutter peaks and the corresponding allele were calculated as previously described [30].

2.8. Admixture of equine and human urine or DNA samples

In order to detect human DNA contamination in equine sample, human and equine urine samples were mixed at proportions of 10:1, 6:1, 5:1, 3:1, 1:1, 1:3, 1:5, 1:6, and 1:10. In addition, DNA samples from each species were mixed at proportions of 20:1, 15:1, 10:1, 5:1, 3:1, 1:1, 1:3, 1:5, 1:10, 1:15, and 1:20 with 50 ng/μl. Eight urine or DNA samples from each species were employed for the admixture study.

2.9. Statistical analysis

Allele number and frequencies, and genotype heterozygosity were acquired by Microsatellite Analyzer [31]. Polymorphism information content (PIC) was computed based on the allele frequencies [32]. Hardy–Weinberg proportions were tested to examine each locus studied. Genetic diversity was expressed by total inbreeding coefficient within breed (F_{IS}) and by co-ancestry coefficient between two breeds (F_{ST}). The parameters, F_{IS} and F_{ST}, were estimated according to Weir and Cockerham [33] using Genepop 4.05 software [34]. The average PI for a single locus was estimated by API-CALC 1.0 computer program incorporating F_{IS} and F_{ST} [5]. The values of F_{IS} and F_{ST} used in PI calculations are shown in Table 4. The cumulative PI was a product of the average PI from independent loci.

3. Results

3.1. Effect of DNA template on 21-plex STR co-amplification

Success of 24-plex STR typing depends on quantity of DNA template. With sufficient DNA template, the 24-plex PCR for amplification of 24 STR loci showed reliable results. When 1 ng or less genomic DNA template was used, only few STR markers were detectable with low intensity of signals. However, when the quantity of template was increased to 10 ng, all 21 STR loci were detectable but significant imbalance of allele peak areas was observed. The use of 20 ng DNA improved balance of allele peak area among 21 markers with slightly improved signal intensity with 50 ng.

This result indicated that the multiplex was optimized to work with 20–50 ng of DNA. DNA quantity in equine urine varied from 1 ng to 4601 ng per ml of fresh urine with an average of 2200 ng in 58 samples. Although 10 μl of urine would, theoretically, yield

Table 2
Correlation coefficient (r^2) between apparent and nominal sizes, allele confidence interval (ACI, bp) and standard deviation (SD, bp) for average of allele sizes, peak area ratios (PAR) for heterozygous alleles (PAR_H, mean ± SD) and the PAR for stutters vs the corresponding alleles (PAR_S, mean ± SD) in 261 horses by 21-plex STR typing system.

STR	r^2	ACI	SD	PAR_H	PAR_S
VHL20	0.999846	0.19	0.04–0.17	0.65 ± 0.06	0.27 ± 0.11
UMNe156	0.999801	0.21	0.03–0.35	0.81 ± 0.13	0.32 ± 0.04
HTG4	0.999777	0.23	0–0.48	0.80 ± 0.11	0.32 ± 0.11
AHT4	0.999917	0.23	0.08–0.33	0.77 ± 0.06	0.37 ± 0.18
HMS6	0.999509	0.27	0.08–0.24	0.75 ± 0.11	0.31 ± 0.08
HMS7	0.999971	0.32	0.09–0.18	0.79 ± 0.16	0.41 ± 0.14
ASB9	0.999943	0.29	0.03–0.41	0.84 ± 0.07	0.39 ± 0.11
ASB2	0.999871	0.30	0.06–0.19	0.72 ± 0.07	0.45 ± 0.20
COR045	0.999877	0.25	0.06–0.19	0.73 ± 0.06	0.37 ± 0.17
HMS3	0.999840	0.24	0.08–0.46	0.61 ± 0.21	0.27 ± 0.17
UMNe222	0.999991	0.28	0.10–0.38	0.80 ± 0.10	0.31 ± 0.10
LEX074	0.999808	0.28	0.06–0.24	0.76 ± 0.13	0.40 ± 0.10
COR008	0.999696	0.23	0.07–0.16	0.70 ± 0.05	0.46 ± 0.17
HMS2	0.999967	0.21	0.06–0.23	0.77 ± 0.06	0.40 ± 0.05
UM015	0.999975	0.28	0.07–0.19	0.83 ± 0.05	0.31 ± 0.06
COR018	0.999877	0.21	0.06–0.18	0.75 ± 0.03	0.32 ± 0.05
UMNe116	0.999951	0.20	0.01–0.36	0.76 ± 0.19	0.41 ± 0.16
UMNe479	0.999935	0.22	0.04–0.37	0.81 ± 0.15	0.29 ± 0.11
HMS1	0.999731	0.24	0.07–0.48	0.75 ± 0.12	0.39 ± 0.09
UMNe191	0.999767	0.43	0.19–0.33	0.53 ± 0.36	0.28 ± 0.14
LEX073	0.999937	0.41	0.09–0.31	0.82 ± 0.14	0.63 ± 0.16

sufficient DNA (20–50 ng) for one multiplex PCR reaction from most urine samples, 0.1–1 ml is an initial volume to ensure sufficient quantity for replicate and multiple analyses. Urine volume was increased for some samples with less abundant DNA. In contrast, a complete STR profile was obtained from a single hair root.

3.2. 21-Plex equine STR typing system showed high accuracy and precision with high power of discrimination

Nominal allele sizes were determined by DNA sequencing and were consistent with fragment size analysis. Allele sizes were well estimated with high correlation coefficients ($r^2 = 0.999$) between apparent and nominal allele sizes across all 21 loci (Table 2). Apparent sizes were very close to the nominal sizes with standard deviation <0.50 bp for over 21 markers. The 21-plex STR typing system generated an allele confidence interval of ≤0.43 bp in 261 horses. When the method was tested in 261 racehorse samples, none of the samples showed an identical genotype, indicating that the samples were obtained from different horses. Electropherograms of two samples (A and B) showed two distinct patterns in STR lengths (Fig. 1). Sixteen markers showed distinctly different genotypes while the remaining five markers showed identical profiles between two samples (Table 3). Among those five identical loci, two (ASB9 and HMS2) were homozygotes.

3.3. Heterozygous balance and stutter activity

Most markers had PAR of approximately 0.75 of heterozygous alleles although three markers (VHL20, HMS3 and UMNe191) had lower PAR of 0.65, 0.61, and 0.53, respectively (Table 2). This result indicated relatively balanced heterozygous alleles in most loci. Stutter peaks were observed in all STR loci and PAR of stutters vs. the corresponding alleles was 0.27–0.46 except for LEX073 with 0.63. In total, stutter activities were ~35% of the corresponding alleles.

3.4. Identical STR profile from blood, urine and hair roots of the same horses by 21-plex STR typing system

A total of 36 horses were evaluated to determine whether the same genotype profile could be acquired from different sample types obtained from the same horse. The results indicated that STR profiles from blood, hair root and urine from the same horse were identical. Identical STR profile was also obtained when the samples were collected from the same horse at different time periods indicating that STR profiling is independent of sample type and collection time from the same donor. The result demonstrated reliability and reproducibility of the method.

3.5. Detection of contamination of equine sample by human DNA

All drug positive test results are subject to legal challenge. Presence of human fluid or that of multiple human drugs in equine samples may complicate the interpretation of the positive report. To address this potential problem of contamination by human DNA, 24-plex STR typing system that incorporated three human specific STR markers (TH01, TPOX and D18S51) was developed [10]. The results indicated that all three human DNA markers were detected in DNA templates derived from mixed equine and human urine or DNA samples (Fig. 2). Limit of detection for human markers was 20% human urine and 10% human DNA. In addition, all 21 equine markers were detected in the mixed samples. Limit of detection for equine markers was 15% equine urine and 20% equine DNA. The results demonstrated that this novel 24-plex STR typing method not only identified individual horses, it detected contamination of equine sample by human DNA.

3.6. Specificity of 21-plex and 3-plex STR genotyping system

To determine whether 21-plex equine STR typing system was equine specific and 3-plex human STR system was human specific, human genomic DNA templates from 10 individuals were tested. No signal of equine marker was detected in genomic DNA templates derived from human. Moreover, none of the three human markers was detected in 520 equine samples. The result demonstrated that this method was a highly specific equine genotyping system with specificity for the detection of human DNA.

3.7. Probability of identity and number of markers

Probability of identity reflects the likelihood of observing two distinct individuals with identical genetic marker profiles and is the probability of making a false identification. Thus, a lower value of PI reflects higher power of identification. Expected and observed PI for all 21 markers in the racehorses did not show significant

Table 3
STR profiles of two STB racehorses (A and B).

Horse	VHL20	UMNe156	HTG4	AHT4	HMS6	HMS7	ASB9	ASB2	COR045	HMS3	UMNe222
A	14/19	22/22	16/18	29/33	17/20	20/21	15/15	25/26	19/19	18/24	10/15
B	17/18	21/21	16/18	28/33	15/17	18/20	15/15	12/29	21/23	18/25	10/10

Horse	LEX074	COR008	HMS2	UM015	COR018	UMNe116	UMNe479	HMS1	UMNe191	LEX073
A	17/21	24/24	18/18	14/17	17/18	20/20	24/24	19/19	20/21	38/38
B	14/21	24/30	18/18	14/17	17/17	18/20	24/27	19/22	20/21	38/47

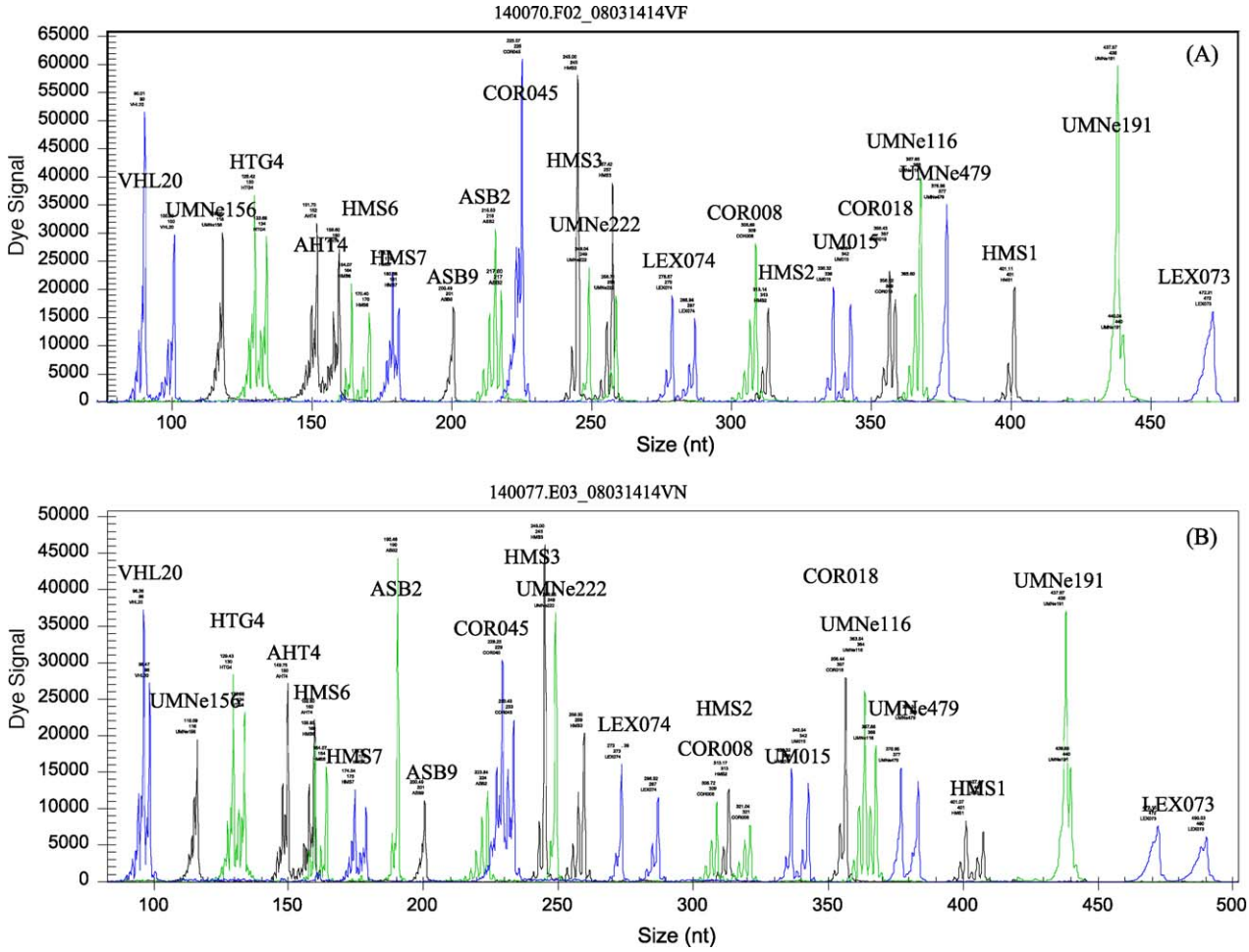

Fig. 1. Individual identification. Two individual horses were discriminated using 21-plex STR system. Electropherograms of two STR profiles were acquired from blood collected from two racehorses. Twelve markers in horse A and 16 markers in horse B showed heterozygosity. These two individual horses shared 22 of 42 alleles.

difference. However, it was evident that power of identification varied among markers; theoretically, the greatest power was reached by LEX073 with a PI of 0.044 STB while the lowest power was by COR018 with a PI of 0.396 in TB (Table 4). When the marker with the highest PI was used, a total of seven markers were needed for a PI of 2×10^{-6} (STB) while only four markers with the lowest PI were needed for the same probability. Seven of the most polymorphic markers attained a PI of 6.8×10^{-8} (TB) and 4.4×10^{-9} (STB) with an assumption of locus independency. Higher power of discrimination was reached when more markers

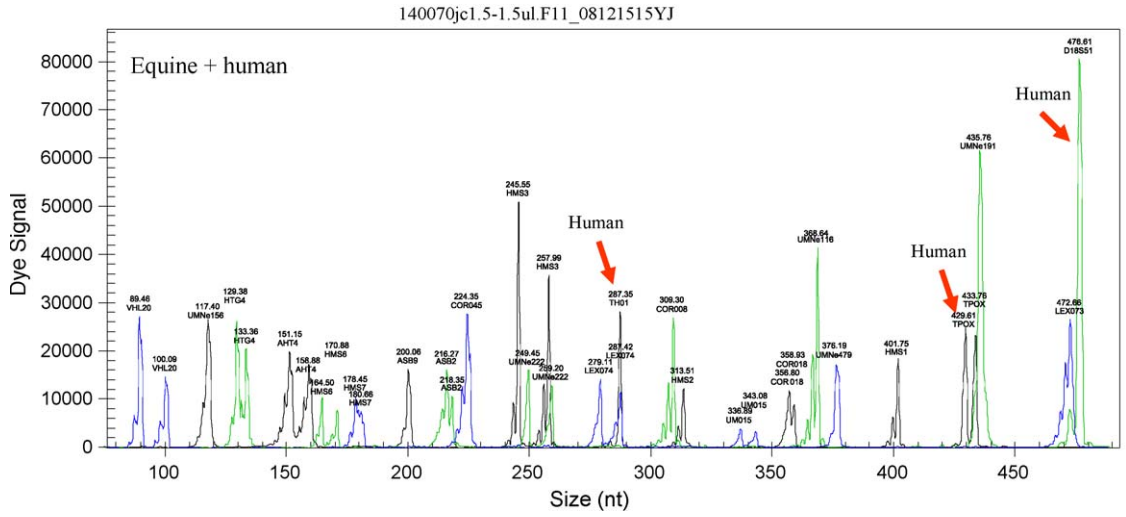

Fig. 2. Identification of human DNA using 24-plex STR typing system. Electropherograms of 24-plex STR profiles indicated specificity of 24-plex STR typing system for simultaneous identification of equine and human DNA. All three human specific STR markers (indicated by arrows) were clearly detected along with all 21 equine markers in a mixture of human and equine DNA samples.

Table 4
Polymorphism information for 21 equine STR loci with repeat motif in 90 TB and 171 STB racehorses: number of alleles, F_{IS}, F_{ST}, observed heterozygosity (H_O), expected heterozygosity (H_E), PIC, observed PI (PI_O), expected PI (PI_E), number of alleles of which frequency ≥0.05 (AF05), and the highest allele frequency (HAF).

STR	Motif	Allele[a]		F_{IS}		F_{ST}	H_O		H_E		PIC		PI_O		PI_E		AF05		HAF	
		TB	STB	TB	STB		TB	STB	TB	STB	TB	STB	TB	STB	TB	STB	TB	STB	TB	STB
VHL20#	TG	5	10	0.063	0.030	0.023	0.700	0.743	0.747	0.765	0.719	0.747	0.106	0.077	0.104	0.081	4	4	0.306	0.377
UMNe156	CA	5	9	−0.102	−0.025	0.083	0.633	0.632	0.575	0.616	0.531	0.587	0.265	0.199	0.243	0.186	3	3	0.539	0.544
HTG4	TG	4	7	0.120	−0.034	0.058	0.456	0.760	0.517	0.735	0.464	0.716	0.294	0.102	0.309	0.095	2	5	0.578	0.436
AHT4#	CA	5	7	−0.070	0.032	0.044	0.778	0.684	0.727	0.707	0.699	0.679	0.121	0.117	0.114	0.112	4	5	0.394	0.474
HMS6#	GT	4	5	−0.116	−0.015	0.175	0.689	0.690	0.617	0.680	0.586	0.649	0.198	0.152	0.189	0.144	3	4	0.544	0.447
HMS7#	CA	5	6	−0.099	0.183	0.112	0.856	0.520	0.779	0.587	0.757	0.545	0.086	0.233	0.080	0.230	5	3	0.300	0.518
ASB9	CA	6	6	0.174	0.100	0.111	0.483	0.655	0.524	0.727	0.489	0.703	0.267	0.111	0.254	0.109	3	5	0.656	0.389
ASB2#	GT	8	9	0.033	−0.006	0.031	0.867	0.819	0.839	0.814	0.823	0.798	0.045	0.061	0.045	0.057	6	5	0.217	0.257
COR045	GT	8	7	0.022	0.022	0.035	0.722	0.760	0.707	0.777	0.684	0.761	0.107	0.073	0.108	0.071	5	6	0.489	0.386
HMS3#	CA	6	13	0.044	0.182	0.192	0.633	0.726	0.663	0.764	0.631	0.747	0.138	0.086	0.146	0.079	4	5	0.517	0.383
UMNe222	CA	6	7	0.009	0.024	0.251	0.767	0.468	0.760	0.479	0.735	0.457	0.091	0.297	0.093	0.286	5	3	0.311	0.693
LEX074	GT	7	8	0.038	0.07	0.088	0.756	0.708	0.785	0.760	0.764	0.737	0.071	0.091	0.074	0.090	5	4	0.317	0.301
COR008	CA	5	9	0.012	0.010	0.201	0.500	0.702	0.506	0.709	0.482	0.691	0.256	0.113	0.264	0.108	4	5	0.672	0.477
HMS2#	CA	5	8	0.087	−0.006	0.129	0.367	0.789	0.401	0.785	0.382	0.771	0.376	0.066	0.362	0.065	2	6	0.756	0.386
UM015#	TG	5	11	−0.069	0.118	0.128	0.678	0.698	0.634	0.769	0.597	0.752	0.185	0.082	0.175	0.075	3	4	0.528	0.395
COR018#	CA	3	7	−0.057	0.066	0.092	0.422	0.556	0.400	0.595	0.370	0.573	0.396	0.195	0.382	0.187	2	4	0.744	0.594
UMNe116#	CA	4	5	0.222	0.034	0.092	0.578	0.737	0.613	0.762	0.572	0.742	0.197	0.085	0.210	0.086	3	5	0.494	0.360
UMNe479#	CA	3	4	−0.002	0.090	0.083	0.644	0.409	0.643	0.450	0.603	0.425	0.190	0.327	0.185	0.321	3	3	0.467	0.711
HMS1	TG	4	5	−0.071	0.146	0.007	0.689	0.629	0.643	0.678	0.602	0.641	0.201	0.151	0.188	0.159	3	3	0.411	0.348
UMNe191#	CA	12	17	0.030	0.106	0.108	0.822	0.754	0.798	0.843	0.775	0.832	0.078	0.046	0.066	0.039	5	5	0.278	0.243
LEX073#	CA	7	13	0.045	0.294	0.158	0.656	0.802	0.687	0.852	0.655	0.841	0.127	0.044	0.133	0.036	5	6	0.472	0.249
Mean	–	5.6	8.2	0.015	0.068	0.105	0.652	0.678	0.646	0.707	0.615	0.685	0.181	0.129	0.177	0.125	3.8	4.4	0.476	0.427
SD	–	2.0	3.2	0.090	0.083	0.064	0.141	0.110	0.125	0.108	0.129	0.113	0.099	0.079	0.097	0.078	1.2	1.0	0.154	0.130

[a] Total number of alleles. The cumulative values of PI over 13 independent loci (#) were 1.4×10^{-11} (TB, observed), 9.9×10^{-14} (STB, observed), 9.7×10^{-12} (TB, expected), and 4.8×10^{-14} (STB, expected).

were used where the PI of 9.7×10^{-12} (TB) and 4.8×10^{-14} (STB) could be expected with 13 independent loci.

3.8. Allele number and frequency

Twenty-one markers had an average allele number of 6 (TB) and 8 (STB) with the least number of three in TB by UMNe479 and the highest number of 17 in STB by UMNe191 (Table 4). All loci had 2–6 alleles with frequency of ≥0.05. Nine loci had alleles with frequency >0.5 while four loci (ASB2, LEX074, UMNe191 and LEX073) had alleles evenly distributed (Tables 5 and 6). Sequencing data indicated that the majority of STR loci (15 of 21) had the number of repeat motif from 14 to 30 although five loci (ASB2, UMNe222, UM015, COR018 and UMNe191) also showed shorter repeat sequences and one locus (LEX073) had longer repeat sequences. The partial repeat sequence was observed in six loci with allele frequencies of <0.05 except LEX073 that showed partial allele frequency of 0.19 for allele 33.1 and 0.13 for 34.1 in STB. All these partial sequences were found in STB and only two of them (UMNe191 and LEX073) were observed in TB horses. These partial allele sequences can be accessed through GenBank with the following accession numbers: FJ915129 (VHL20 allele 23.1), FJ915130 (HMS3 allele 23.1), FJ915131 (HMS3 25.1), FJ915132 (UM015 allele 14.1), FJ915133 (UMNe191 allele 8.1), FJ915134 (UMNe191 allele 16.1), FJ915135 (UMNe191 allele 17.1), FJ915136 (UMNe191 allele 19.1), FJ915137 (UMNe191 allele 20.1), FJ915138 (LEX073 allele 33.1), FJ915139 (LEX073 allele 34.1), FJ915140 (LEX073 allele 46.1), FJ915141 (UM015 allele 17.1), FJ915142 (COR018 allele 10.1), and FJ915143 (COR018 allele 22.1). All loci did not show significant deviation from Hardy–Weinberg proportion.

3.9. Heterozygosities and polymorphism information content

Many markers (11 of 21) showed heterozygosity of 0.5–0.8 while four markers (HMS7, ASB2 and UMNe191, LEX073) indicated a value >0.8 and six markers (HTG4, ASB9, UMNe222, HMS2, COR018 and UMNe479) were <0.5 (Table 4). Many loci (11 of 21) showed PIC of 0.5–0.8 whereas three markers (ASB2, UMNe191 and LEX073) were >0.8 while that of seven markers (HTG4, ASB9, UMNe222, COR008, HMS2, COR018, UMNe479) was <0.5. Heterozygosity and PIC showed consistent results regarding polymorphism for all 21 markers. In some loci (HTG4, HMS7, ASB9, UMNe222, COR008, HMS2, COR018) TB horses did not demonstrate the same heterozygosity and PIC patterns as STB despite other loci showed similar degree of polymorphism between two breeds. Observed heterozygosity did not differ significantly from expected value.

3.10. Correlation between polymorphism measurements

Correlation coefficients were acquired using the average values for the two breeds. High correlation was observed between PI, PIC, and heterozygosity ($|r^2| > 0.9$, Table 7). The highest alleles frequency was also highly associated with PIC, heterozygosity and PI ($|r^2| > 0.9$) while it had a moderate relationship with the number of alleles with frequency that was ≥0.05 and the total number of alleles ($|r^2| \approx 0.6$). In comparison with total number of alleles, the number of alleles with frequency ≥0.05 had higher correlations with PI, PIC and heterozygosity ($|r^2| \approx 0.8$) while the correlation ($r^2 = 0.66$) between those two parameters was moderate.

4. Discussion

This novel 21-plex STR typing panel included eight STR loci (VHL20, HTG4, AHT4, HMS6, HMS7, ASB2, HMS3 and HMS2) in the ISAG recommended panel for parentage testing [3,4] while four other ISAG recommended loci were not employed due to failure of co-amplification in the multiplex PCR. The system also included 13 other highly polymorphic loci that showed high power of individual identification in case studies (not shown). These 21 STR loci were distributed on 13 chromosomes while nine loci did not share a common chromosome with others. Among ISAG recommended panel, 12 loci were distributed on nine chromosomes [4]. Use of physically linked loci did not necessarily invalidate the typing method [1,3,4] but might over-estimate the

Table 5
Allelic frequencies of 21 equine STR loci in the Thoroughbred racehorse ($n = 90$).

Allele	VHL20	UMNe156	HTG4	AHT4	HMS6	HMS7	ASB9	ASB2	COR045	HMS3	UMNe222	LEX074	COR008	HMS2	UM015	COR018	UMNe116	UMNe479	HMS1	UMNe191	LEX073
10											0.289										
11																					
12								0.028												0.006	
13															0.011					0.006	
14	0.289										0.050	0.206		0.022	0.222						
15					0.183		0.656		0.016		0.311				0.200					0.012	
16		0.011	0.578		0.039				0.111		0.139			0.022	0.039					0.015	
17	0.189				0.233		0.094		0.006			0.028		0.167	0.528	0.744				0.056	
17.1																				0.044	
18	0.306		0.389			0.083				0.517	0.206	0.011	0.050	0.756		0.217	0.139		0.189	0.278	
19	0.211	0.083	0.017				0.217		0.161		0.006	0.089	0.017	0.033					0.394	0.061	
19.1																				0.022	
20	0.006	0.006			0.544	0.172	0.011		0.072			0.317					0.494			0.25	
21		0.361	0.017			0.228	0.011		0.028			0.233								0.239	
22		0.539				0.300	0.011	0.111	0.489	0.100						0.039	0.356		0.411	0.011	
23						0.217			0.117	0.011		0.117	0.672				0.011		0.006		
24								0.144		0.139			0.194					0.467			
25								0.167		0.211								0.250			
26				0.222				0.111													
27								0.011		0.022								0.283			
28				0.172				0.217													
29				0.206				0.211													
30													0.067								
33				0.394																	0.278
33.1																					0.050
34				0.006																	0.072
34.1																					0.011
38																					0.472
41																					0.023
47																					0.094

Table 6
Allelic frequencies of 21 equine STR loci in the Standardbred Pacer racehorse ($n = 171$).

Allele	VHL20	UMNe156	HTG4	AHT4	HMS6	HMS7	ASB9	ASB2	COR045	HMS3	UMNe222	LEX074	COR008	HMS2	UM015	COR018	UMNe116	UMNe479	HMS1	UMNe191	LEX073
8																				0.006	
8.1																				0.026	
9																				0.041	
10											0.175					0.015					
10.1																0.064					
11											0.003				0.035						
12															0.011					0.026	
13											0.006				0.123					0.237	
14	0.105	0.003										0.117		0.032	0.395					0.056	
14.1															0.012						
15	0.009		0.003		0.038		0.301				0.105			0.015	0.006					0.022	
16		0.108	0.436		0.056		0.006		0.117		0.693			0.137	0.047	0.187				0.009	
16.1																				0.009	
17	0.219	0.009	0.117		0.152		0.064		0.132	0.008	0.006	0.020		0.088	0.184	0.594				0.243	
17.1															0.026					0.029	
18	0.377		0.202			0.064	0.129			0.076	0.012	0.009		0.386	0.155	0.008	0.211		0.289	0.015	
19	0.178	0.047	0.108		0.447	0.006	0.389		0.082			0.023	0.091		0.006		0.184		0.345	0.003	
19.1																				0.006	
20	0.012	0.003	0.099		0.307	0.377		0.006	0.132	0.006		0.281					0.360			0.178	
20.1																				0.053	
21	0.018	0.544	0.035			0.020			0.146			0.009	0.129	0.123					0.009	0.041	
22	0.006	0.275				0.518	0.111	0.205	0.386	0.012		0.240	0.053	0.082		0.123	0.091		0.348		
22.1																0.009					
23						0.015			0.006	0.383		0.301	0.191				0.155		0.009		
23.1	0.044									0.023											
24	0.032	0.003						0.170		0.023			0.477	0.137				0.711			
25		0.006						0.026		0.272								0.196			
25.1										0.056											
26				0.178				0.257		0.073			0.003								
27				0.096				0.023		0.041								0.091			
27.1										0.015											
28				0.140				0.111		0.012			0.026					0.003			
29				0.006				0.196					0.012								
30				0.006				0.006					0.018								
33				0.474																	0.149
33.1																					0.190
34				0.099																	0.249
34.1																					0.126
35																					0.053
38																					0.009
39																					0.079
40																					0.014
43																					0.009
46																					0.029
46.1																					0.023
47																					0.058
48																					0.012

Table 7
Correlations between polymorphism measurements using average values for the two breeds (TB and STB).

r^2	PI_O	PI_E	PIC	H_O	H_E	HAF	Allele[a]	AF05
PI_O	1	0.99	−0.97	−0.95	−0.99	0.91	−0.68	−0.80
PI_E	0.99	1	−0.97	−0.95	−0.99	0.92	−0.69	−0.80
PIC	−0.97	−0.97	1	0.93	0.98	−0.92	0.67	0.75
H_O	−0.95	−0.95	0.93	1	0.96	−0.92	0.60	0.72
H_E	−0.99	−0.99	0.98	0.96	1	−0.95	0.69	0.77
HAF	0.91	0.92	−0.92	−0.92	−0.95	1	−0.59	−0.60
Allele	−0.68	−0.69	0.67	0.60	0.69	−0.59	1	0.66
AF05	−0.80	−0.80	0.75	0.72	0.77	−0.60	0.66	1

r^2: correlation coefficient; PI_O: observed PI; PI_E: expected PI; H_O: observed heterozygosity; H_E: expected heterozygosity; HAF: the highest allele frequency; AF05: number of alleles of which frequency was $\geq$0.05.
[a] Total number of alleles.

calculation of PI since the calculation assumes independence of loci. To obtain a more accurate estimation of PI, it would be best to use independent loci in the typing method. Thus, only 13 independent loci in the method were calculated to obtain average PI of 1.4×10^{-11} (TB) and 9.9×10^{-14} (STB), although 21 loci were employed in the method. The probability of observing matching profiles would be well presented by the PI in court [35,36].

Balance of peak intensity among STR markers was one of the greatest challenges in multiplex PCR amplification. Peak intensity of allele varied among STR markers and was associated with a number of factors that included the type of fluorescent dye used. Fluorescence intensity depends on the type of dye; there was a five-fold difference among dyes, D4 (blue) was 5-fold stronger than D3 (green) while D2 (black) was 5-fold weaker than D3. Cy5 appeared as an alternative fluorescent dye to D4 due to similar emission spectrum. In order to balance the peak height, primer concentrations should be increased or decreased according to the signal intensity of the dye label. Peak intensity did not change when the primer concentration reached its limit. For instance, a primer concentration higher than 90 nM did not increase the peak area of ASB2 marker. Limitation of amplification was also associated with complexity of multiplexing since primers compete for a limited quantity of DNA template, enzyme and other reagents, and complicated interactions between primers and components of the reaction [37].

Importance should be placed on heterozygous balance within the same locus. Imbalance of heterozygous alleles within the same locus was observed in the homogenous DNA isolate. For instance, the PARs of heterozygous alleles for UMNe191 were 0.45 and 0.23 in two DNA isolates (Fig. 1) although the average ratio in 261 samples was 0.53, the lowest in the panel (Table 2). Heterozygous imbalance occurs in DNA mixtures and PAR reflects the proportions of the components in the DNA mixtures [38]. However, due to implementation of strict chain of custody procedures in sampling, it was unlikely that mishandling of samples was the cause of the observed heterozygote imbalance. A mutation in primer binding site may influence amplification efficacy and thus, cause peak imbalance, particularly if such mutation were near the 3′ end of the primer binding site [39]. Insufficient DNA template or low DNA quality might also contribute to heterozygous imbalance. It seems likely that DNA quantity or quality contributed to peak imbalance in UMNe191 since perfect balance was observed in many samples and the current STR method was tested in several bloodstain cases and did not show any deficiency in amplification of this marker (data not shown). Thus, observations regarding heterozygous imbalance do not necessarily indicate invalidity of this marker in the panel and were probably rare occurrences due to low DNA quality or quantity.

Non-specific PCR products such as stutter peaks were observed across loci although amplification conditions were extensively optimized. The majority of stutter activity may result from replication slippage [2]. This undesired product was very common and appeared unavoidable because of the nature of STR locus and the characteristics of Taq DNA polymerase [40]. Stutter peak was the major disadvantage in using dinucleotide STR in sample identification due to difficulty of automatic analysis by computer software or risk of false allele designation [2]. However, lower stutter activity was evident with amplification of longer repeat motif such as tetranucleotide STR locus [40,41] and three human tetranucleotide loci showed extremely low stutter activity in this study (Fig. 2). Currently, only very few of such STR loci have been described in horses and, thus, introduction of more polymorphic tetranucleotide STR markers would add significance to equine genotyping.

Allele designation should be guided so that data from different systems and laboratories can be compared. A numeric allele designation, widely employed in human DNA testing, indicates the number of repeat units was a clear expression of the genotype [28]. However, this numeric allele designation method had not previously been introduced in horse genotyping system until the present study was undertaken. Alphabetic allele codes were referred to as an unwieldy nomenclature [28] but was extensively used in the description of equine genotype data without definition of the codes [1–3]. With assistance from DNA sequencing, partial repeat sequence was identified in six loci although the result was not corroborated by other investigators for the same markers, including VHL20 and HMS3 [1]. However, partial VLH20 and HMS3 sequences were not observed in TB horses. This result indicated that the current method detected a single nucleotide difference with high accuracy.

The theoretical probability of making false identification or reaching false conclusion using the present novel genotyping system was nearly zero (9.7×10^{-12} for TB and 4.8×10^{-14} for STB). The PI value was a valuable parameter in determining the number of markers needed to reach a confident level of identification. A minimum of four markers with the least PI values was sufficient for reaching a matching power of 0.99997 in TB and 0.99999 in STB.

5. Conclusion

A novel 24-plex STR typing system to identify a racehorse and detect sample contamination by human biological material was developed to meet the increasing demand for sample identification and to verify chain of custody procedures in the horseracing industry. The method has high power of discrimination and reproducibility with high accuracy and precision. Polymorphism index and usefulness of the markers were evaluated and allele frequency databases were established (Tables 5 and 6). Probability of identity was considered an important measurement to estimate power of discrimination from sample matching or exclusion.

Acknowledgements

Financial support for this study was provided by the PA Racing Commissions. Financial contribution was received from The PA

Harness Horsemen Association at Pocono Downs. We thank Anne Hess for the excellent technical support with sampling, and Johanna Mylet of Beckman Coulter, USA for assistance with automatic allelic calling, and Drs. ADJ Overall (University of Reading, UK) and F. Rousset (Université Montpellier 2, France) for their support in the use of API-CALC and Genepop 4.05 software, respectively.

References

[1] A.T. Bowling, M.L. Eggleston-Stott, G. Byrns, R.S. Clark, S. Dileanis, E. Wictum, Validation of microsatellite markers for routine horse parentage testing, Anim. Genet. 28 (1997) 247–252.
[2] T. Tozaki, H. Kakoi, S. Mashima, K. Hirota, T. Hasegawa, N. Ishida, N. Miura, N.H. Choi-Miura, M. Tomita, Population study and validation of paternity testing for thoroughbred horses by 15 microsatellite loci, J. Vet. Med. Sci. 63 (2001) 1191–1197.
[3] S.Y. Lee, G.J. Cho, Parentage testing of thoroughbred horse in korea using microsatellite DNA typing, J. Vet. Sci. 7 (2006) 63–67.
[4] P. Dimsoski, Development of a 17-plex microsatellite polymerase chain reaction kit for genotyping horses, Croat. Med. J. 44 (2003) 332–335.
[5] K.L. Ayres, A.D.J. Overall, Api-calc 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives, Mol. Ecol. Notes 4 (2004) 315–318.
[6] K.J. Craft, J.D. Owens, M.V. Ashley, Application of plant DNA markers in forensic botany: genetic comparison of quercus evidence leaves to crime scene trees using microsatellites, Forensic Sci. Int. 165 (2007) 64–70.
[7] K.L. Ayres, A.D.J. Overall, Allowing for within-subpopulation inbreeding in forensic match probabilities, Forensic Sci. Int. 103 (1999) 207–216.
[8] N. Dawnay, R. Ogden, R.S. Thorpe, L.C. Pope, D.A. Dawson, R. McEwing, A forensic STR profiling system for the eurasian badger: a framework for developing profiling systems for wildlife species, Forensic Sci. Int. Genet. 2 (2008) 47–53.
[9] A.J. Stevenson, M.P. Weber, F. Todi, M. Mendonca, J.D. Fenwick, L. Young, E. Kwong, F. Chen, P. Beaumier, S. Timmings, et al., Determination of procaine in equine plasma and urine by high-performance liquid chromatography, J. Anal. Toxicol. 16 (1992) 93–96.
[10] J.-W. Chen, C.E. Uboh, L.R. Soma, X. Li, F. Guan, Y. You, Microsatellite loci in urine supernatant and stored samples from racehorses, Am. J. Vet. Res. 70 (2009) 648–657.
[11] D.R. Linfert, A.H. Wu, G.J. Tsongalis, The effect of pathologic substances and adulterants on the DNA typing of urine, J. Forensic Sci. 43 (1998) 1041–1045.
[12] H. van Haeringen, A.T. Bowling, M.L. Stott, J.A. Lenstra, K.A. Zwaagstra, A highly polymorphic horse microsatellite locus: VHL20, Anim. Genet. 25 (1994) 207.
[13] H. Ellegren, M. Johansson, K. Sandberg, L. Andersson, Cloning of highly polymorphic microsatellites in the horse, Anim. Genet. 23 (1992) 133–142.
[14] M.M. Binns, N.G. Holmes, A. Holliman, A.M. Scott, The identification of polymorphic microsatellite loci in the horse and their use in thoroughbred parentage testing, Br. Vet. J. 151 (1995) 9–15.
[15] G. Guerin, M. Bertaud, Y. Amigues, Characterization of seven new horse microsatellites: HMS1, HMS2, HMS3, HMS5, HMS6, HMS7 and HMS8, Anim. Genet. 25 (1994) 62.
[16] M. Breen, G. Lindgren, M.M. Binns, J. Norman, Z. Irvin, K. Bell, K. Sandberg, H. Ellegren, Genetical physical assignments of equine microsatellites—first integration of anchored markers in horse genome mapping, Mamm. Genome 8 (1997) 267–273.
[17] L.S. Ruth, T.J. Hopman, M.D. Schug, C.F. Aquadro, A.T. Bowling, J.D. Murray, A.R. Caetano, D.F. Antczak, Equine dinucleotide repeat loci COR041-COR060, Anim. Genet. 30 (1999) 320–321.
[18] E. Bailey, L. Skow, D. Bernoco, A. DelValle, M.D. Scavone, A.T. Bowling, J.D. Murray, Equine dinucleotide repeat loci LEX071 through LEX078, Anim. Genet. 31 (2000) 286–287.
[19] T.J. Hopman, E.B. Han, M.R. Story, M.D. Schug, C.F. Aquadro, A.T. Bowling, J.D. Murray, A.R. Caetano, D.F. Antczak, Equine dinucleotide repeat loci COR001-COR020, Anim. Genet. 30 (1999) 225–226.
[20] A.H. Meyer, S.J. Valberg, K.R. Hillers, J.K. Schweitzer, J.R. Mickelson, Sixteen new polymorphic equine microsatellites, Anim. Genet. 28 (1997) 69–70.
[21] M.L. Wagner, G. Goh, J.T. Wu, L.Y. Morrison, L.J. Alexander, T. Raudsepp, L.C. Skow, B.P. Chowdhary, J.R. Mickelson, Sixty-seven new equine microsatellite loci assigned to the equine radiation hybrid map, Anim. Genet. 35 (2004) 484–486.
[22] J.R. Mickelson, J.T. Wu, L.Y. Morrison, J.E. Swinburne, M.M. Binns, K.M. Reed, L.J. Alexander, Eighty-three previously unreported equine microsatellite loci, Anim. Genet. 34 (2003) 71–74.
[23] J.R. Mickelson, M.L. Wagner, G. Goh, J.T. Wu, L.Y. Morrison, L.J. Alexander, T. Raudsepp, L.C. Skow, B.P. Chowdhary, J.E. Swinburne, M.M. Binns, Thirty-five new equine microsatellite loci assigned to genetic linkage and radiation hybrid maps, Anim. Genet. 35 (2004) 481–484.
[24] A. Edwards, A. Civitello, H.A. Hammond, C.T. Caskey, DNA typing and genetic mapping with trimeric and tetrameric tandem repeats, Am. J. Hum. Genet. 49 (1991) 746–756.
[25] R. Anker, T. Steinbrueck, H. Donis-Keller, Tetranucleotide repeat polymorphism at the human thyroid peroxidase (HTPO) locus, Hum. Mol. Genet. 1 (1992) 137.
[26] R.E. Straub, M.C. Speer, Y. Luo, K. Rojas, J. Overhauser, J. Ott, T.C. Gilliam, A microsatellite genetic linkage map of human chromosome 18, Genomics 15 (1993) 48–56.
[27] P.M. Vallone, J.M. Butler, Autodimer:, A screening tool for primer–dimer and hairpin structures, Biotechniques 37 (2004) 226–231.
[28] B. Budowle, P. Garofano, A. Hellman, M. Ketchum, S. Kanthaswamy, W. Parson, W. van Haeringen, S. Fain, T. Broad, Recommendations for animal DNA forensic and identity testing, Int. J. Legal Med. 119 (2005) 295–302.
[29] A. Moller, E. Meyer, B. Brinkmann, Different types of structural variation in STRs: HUMFES/FPS, HUMVWA and HUMD21S11, Int. J. Legal Med. 106 (1994) 319–323.
[30] P. Gill, R. Sparkes, R. Pinchin, T. Clayton, J. Whitaker, J. Buckleton, Interpreting simple str mixtures using allele peak areas, Forensic Sci. Int. 91 (1998) 41–53.
[31] D. Dieringer, C. Schlotterer, Microsatellite analyser (MSA): A platform independent analysis tool for large microsatellite data sets, Mol. Ecol. Notes 3 (2003) 167–169.
[32] D. Botstein, R.L. White, M. Skolnick, R.W. Davis, Construction of a genetic linkage map in man using restriction fragment length polymorphisms, Am. J. Hum. Genet. 32 (1980) 314–331.
[33] B.S. Weir, C.C. Cockerham, Estimating *F*-statistics for the analysis of population structure, Evolution 38 (1984) 1358–1370.
[34] F. Rousset, Genepop'007: a complete re-implementation of the genepop software for windows and linux, Mol. Ecol. Resourc. 8 (2008) 103–106.
[35] W.K. Fung, Y.Q. Hu, Y.K. Chung, On statistical analysis of forensic DNA: theory, methods and computer programs, Forensic Sci. Int. 162 (2006) 17–23.
[36] P. Taberlet, G. Luikart, Non-invasive genetic sampling and individual identification, Biol. J. Linnean Soc. 68 (1999) 41–55.
[37] C.P. Kimpton, N.J. Oldroyd, S.K. Watson, R.R. Frazier, P.E. Johnson, E.S. Millican, A. Urquhart, B.L. Sparkes, P. Gill, Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification, Electrophoresis 17 (1996) 1283–1293.
[38] T.M. Clayton, J.P. Whitaker, R. Sparkes, P. Gill, Analysis and interpretation of mixed forensic stains using DNA str profiling, Forensic Sci Int. 91 (1998) 55–70.
[39] D.F. Callen, A.D. Thompson, Y. Shen, H.A. Phillips, R.I. Richards, J.C. Mulley, G.R. Sutherland, Incidence and origin of "Null" Alleles in the (AC)n microsatellite markers, Am. J. Hum. Genet. 52 (1993) 922–927.
[40] P.S. Walsh, N.J. Fildes, R. Reynolds, Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus VWA, Nucleic Acids Res. 24 (1996) 2807–2812.
[41] T.R. Moretti, A.L. Baumstark, D.A. Defenbaugh, K.M. Keys, J.B. Smerick, B. Budowle, Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples, J. Forensic Sci. 46 (2001) 647–660.