Forensic Science International: Genetics 30 (2017) 81–92

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Population and performance analyses of four major populations with Illumina's FGx Forensic Genomics System

Jennifer D. Churchill[a,*], Nicole M.M. Novroski[a], Jonathan L. King[a], Lay Hong Seah[b], Bruce Budowle[a,c]

[a] Center for Human Identification, University of North Texas Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, TX 76107, USA
[b] Department of Chemistry Malaysia Kuching, Ministry of Science, Technology and Innovation (MOSTI) Malaysia
[c] Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 8 December 2016
Received in revised form 14 June 2017
Accepted 15 June 2017
Available online 17 June 2017

Keywords:
ForenSeq™
DNA Signature Prep kit
FGx Forensic Genomics System
Major populations
STR
SNP
Allele frequencies
Population genetics

ABSTRACT

The MiSeq FGx Forensic Genomics System (Illumina) enables amplification and massively parallel sequencing of 59 STRs, 94 identity informative SNPs, 54 ancestry informative SNPs, and 24 phenotypic informative SNPs. Allele frequency and population statistics data were generated for the 172 SNP loci included in this panel on four major population groups (Chinese, African Americans, US Caucasians, and Southwest Hispanics). Single-locus and combined random match probability values were generated for the identity informative SNPs. The average combined STR and identity informative SNP random match probabilities (assuming independence) across all four populations were 1.75E-67 and 2.30E-71 with length-based and sequence-based STR alleles, respectively. Ancestry and phenotype predictions were obtained using the ForenSeq™ Universal Analysis System (UAS; Illumina) based on the ancestry informative and phenotype informative SNP profiles generated for each sample. Additionally, performance metrics, including profile completeness, read depth, relative locus performance, and allele coverage ratios, were evaluated and detailed for the 725 samples included in this study. While some genetic markers included in this panel performed notably better than others, performance across populations was generally consistent. The performance and population data included in this study support that accurate and reliable profiles were generated and provide valuable background information for laboratories considering internal validation studies and implementation.

1. Introduction

Numerous studies have illustrated that massively parallel sequencing technologies offer a higher throughput of genetic markers than capillary electrophoresis-based technologies [1–18]. With these data, backward compatibility to current short tandem repeat (STR) marker databases can be maintained while increasing the power of discrimination with additional STRs and identity SNPs, as well as including new types of information with ancestry and phenotype SNPs [4,8,12,13,18,19,20–24]. The smaller amplicons afforded by analyzing SNP markers improve characterization of degraded or challenged samples often encountered in forensic evidence [2,12,19,25,26]. Furthermore, massively parallel sequencing's ability to identify repeat-based and sequence-based variation among STR alleles has the potential to increase the power of discrimination, especially for kinship analysis and mixture interpretation [1,4–6,9,16,27,28].

The MiSeq FGx Forensic Genomics System (Illumina, San Diego, CA, USA) is a commercially available massively parallel sequencing workflow that enables amplification and sequencing of 231 forensically-relevant genetic markers [29]. The ForenSeq DNA Signature Prep kit's (the FGx Forensic Genomics System's library prep kit; Illumina) multiplex primer panel includes primers for 59 STRs (including Amelogenin), 94 identity informative SNPs (iiSNPs), 54 ancestry informative SNPs (aiSNPs), and 24 phenotype informative SNPs (piSNPs) [29]. Previous studies have illustrated the MiSeq FGx Forensic Genomics System's sensitivity of detection [5,18], characterization of populations [6,9], and utility in forensically relevant samples (e.g., mixtures [5] and historical human remains [30,31]). However, prior to validation and implementation in forensic genetic laboratories, population studies are necessary for generating allele frequencies for calculating the strength of massively parallel sequencing-generated DNA profile results. Performance evaluations for each of the

* Corresponding author.
E-mail address: Jennifer.Churchill@unthsc.edu (J.D. Churchill).

http://dx.doi.org/10.1016/j.fsigen.2017.06.004
1872-4973/

genetic markers included in the MiSeq FGx Forensic Genomics System also provide valuable information and guidance for establishing parameters for interpretation of data generated by massively parallel sequencing.

Due to the substantial amount of genetic data produced in this study and the forensic DNA community's traditional interest in STRs, the allele frequencies, sequence variation, and population statistics for the 59 STRs included in the FGx Forensic Genomics System were described separately by Novroski et al. [6]. For the data described herein, allele frequencies and standard population statistics were generated for the 172 SNP loci included in the kit for 725 samples of Chinese, African American, Caucasian, and Southwest Hispanic ancestry. Additionally, a performance evaluation of all 231 loci included in the kit was completed with metrics such as profile completeness, read depth, relative locus performance, and allele coverage ratios reported for all markers in each population.

2. Materials and methods

2.1. Samples

DNA samples from 725 unrelated individuals (i.e., Chinese (ASN; n = 161), African Americans (AFR; n = 167), US Caucasians (CAU; n = 208), and Southwest Hispanics (HIS; n = 189); previously described by Novroski et al. [6]) were used for this study in accordance with the policies and procedures approved by the Institutional Review Board for the University of North Texas Health Science Center in Fort Worth, Texas. DNA was extracted with the QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA, USA) using the manufacturer's protocols [32]. Extracted DNA was quantified using the Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific) and a Qubit® 2.0 Fluorometer (Thermo Fisher Scientific). Samples were normalized to 0.2 ng/μl when possible.

2.2. Library preparation and MPS sequencing

Libraries were prepared using the ForenSeq™ DNA Signature Prep Kit as described in Churchill et al. [5,33] with the following exceptions: 10 μl of pooled, normalized libraries were diluted in hybridization buffer; primer mix A was used for two runs of low quantity DNA samples (ASN); and runs included 32–34 samples each, including controls. Primer mix A, which only includes primers for the STRs and iiSNPs, was used for two sequencing runs that included low quantity DNA (ASN) samples in an effort to improve the likelihood of obtaining a full profile for these samples. Therefore, the total number of samples analyzed for the aiSNPs and piSNPs is less than for the iiSNPs and STRs for the ASN data. Massively parallel sequencing was performed on the MiSeq desktop sequencer (Illumina) with a MiSeq FGx Reagent Kit (Illumina) following the manufacturer's recommended protocols [34].

2.3. Data analysis

Data analysis was performed using the ForenSeq Universal Analysis Software v1.2.16173 (UAS, Illumina) [35]. This data analysis was supplemented by additional analyses of the ForenSeqRunStatistics XML file with in-house Excel-based workbooks where marker performance (i.e., relative locus performance and allele coverage ratios) was evaluated. Read depth was used to calculate normalized relative locus performance at each locus to account for variability between runs (e.g., differing cluster densities) by dividing read depth for the locus by the sample's total read depth for the run. Allele coverage ratios were calculated for all heterozygous genotypes for each autosomal and X (for females) locus by dividing the lower read depth allele by the higher read depth allele at that locus (e.g., 300X/400X = 0.75; 1.0 indicating equal read depth/balanced alleles). For genotype calls, a minimum read depth threshold of 40X was used for all loci, and a minimum allele coverage ratio threshold of 0.25 was used for allele calling for autosomal loci and X-loci. Bioinformatic concordance for genotype calls was determined for the STR genotypes using a modified version of STRait Razor 2.0 and in-house Excel-based workbooks [personal communication with Jonathan King; [36,37]]. Allele frequencies were determined by the counting method.

Standard population genetic analyses, including heterozygosity and tests for Hardy Weinberg Equilibrium, F_{ST}, and linkage disequilibrium, were performed using Genetic Data Analysis (GDA) [38]. Random match probability calculations were performed with in-house Excel-based workbooks. Principal component analysis of the aiSNPs was completed on a per sample basis with the ForenSeq UAS software by projecting the sample on the first two principal component analysis components, which were pre-trained with 1000 Genomes data [35,39,40]. Additionally, a separate principal component analysis for all aiSNP genotypes generated in this data set was completed with RStudio® [41] using the 'princomp' function. Phenotype estimates were performed on the ForenSeq UAS software which bases prediction probabilities on the HIrisPlex model [35,42–44].

3. Results and discussion

3.1. Identity informative SNPs

The 94 iiSNPs included in both primer mix A and primer mix B of the ForenSeq DNA Signature Prep Kit were typed on 725 samples from four population groups. Allele counts and allele frequencies were calculated for each population at each locus and are listed in Supplementary Table 1. Tests for Hardy Weinberg Equilibrium identified eight loci (rs1493232, rs2111980, rs1490413, rs159606, rs729172, rs7041158, rs214955, and rs9905977), five loci (rs1382387, rs1335873, rs993934, rs6811238, and rs321198), six loci (rs2269355, rs2342747, rs917118, rs993934, rs4606077, and rs9905977), and one locus (rs2920816) in the ASN, AFA, CAU, and HIS populations, respectively, that deviated from expectations ($p < 0.05$). After adjusting for multiple tests with a Bonferroni correction [45], no significant deviations from Hardy Weinberg Equilibrium ($p < 0.00053$; Supplementary Table 2) were detected. Wright's F_{ST} estimates [46–48] were calculated for the 94 iiSNPs in the four major population groups (Supplementary Table 6). The F_{ST} estimate over all 94 iiSNP loci was 0.04897 and ranged from −0.00231 for rs321198 to 0.24149 for rs1335873. Tests for linkage disequilibrium were performed solely on the iiSNPs for all four populations with an α of 0.05. After Bonferroni correction, three pairwise comparisons, with one pairwise comparison observed in the AFA population and two pairwise comparisons observed in the CAU population, generated a significant *p*-value (<0.00001144; Supplementary Table 3). All three pairwise comparisons with significant linkage disequilibrium were not syntenic. The observed heterozygosity and expected heterozygosity of the iiSNPs were compared (Supplementary Table 4). The two iiSNP loci with the greatest difference in observed heterozygosity versus expected heterozygosity values were rs1886510 and rs1028528 (with a decrease of 0.1039 and 0.1006, respectively, of observed heterozygosity over expected heterozygosity). The allele frequencies reported in Supplementary Table 1 were used to determine random match probabilities for each iiSNP locus and a combined iiSNP random match probability for each of the four populations (Supplementary Table 5). The average single-locus random match probability value across all populations ranged from 0.38 ± 0.000406 for rs1498553 to 0.59 ± 0.049539 for rs2056277.

The combined iiSNP random match probability for each of the four populations was 1.03E-35, 1.15E-35, 4.49E-38, and 2.09E-37 for the ASN, AFA, CAU, and HIS populations, respectively.

Performance metrics such as profile completeness, read depth, relative locus performance, and allele coverage ratios were calculated to evaluate the performance of the ForenSeq multiplex. There were 67,558 locus genotype calls made out of a possible 68,150 locus genotype calls for typing 94 iiSNPs in 725 samples with the thresholds set for analysis of this data set. Therefore, the individual locus genotyping success rate was 99.13%. Seventy-five individual locus genotypes did not meet the allele coverage ratio threshold, and 517 individual locus genotypes did not meet the read depth threshold. Of the 725 samples typed for iiSNPs, 504 (69.52%) generated complete profiles, and 718 (99.03%) generated profiles that were 90% complete or greater. All 725 iiSNP profiles were 73.40% complete or greater. Thirty of the 94 iiSNP loci (31.91%) were typed successfully for all 725 samples in the data set. Rs7041158 and rs2920816 were the lowest performing loci. Genotyping failed at the rs7041158 locus for 109 of the 725 samples (15.03%), and rs2920816 genotyping failed for 93 of the 725 samples (12.83%) with the thresholds set for the analysis of this data set. These two loci were some of the lowest performing loci in terms of read depth/relative locus performance (Fig. 1). Average read depth for the iiSNPs ranged from 81.09 reads (X) (±35.98X) to 4919.27X (±2046.17X), with similar spreads observed across each population analyzed. Average relative locus performance ranged from 0.00031 (±0.00010) to 0.01940 (±0.00588) across all 94 iiSNPs (Fig. 1; Supplementary Fig. 1). Average allele coverage ratios ranged from 0.42 (±0.10) to 0.89 (±0.09) across the 94 iiSNPs (Fig. 2; Supplementary Fig. 2).

3.2. Ancestry informative SNPs

The 56 aiSNPs included in primer mix B of the ForenSeq DNA Signature Prep Kit include two SNPs, rs16891982 and rs12913832, that overlap with the 24 piSNPs also included in primer mix B. While these two SNPs were included in the principal component analysis and population genetic analyses of the aiSNPs, their allele frequencies and performance evaluation are included below in the discussion of the piSNPs. The aiSNPs and piSNPs were typed on 676 samples that included 112 Chinese, 167 African Americans, 208 US Caucasians, and 189 Southwest Hispanics. Allele counts and allele frequencies for the aiSNPs were calculated for each population at each locus and are listed in Supplementary Table 7. Tests for departures from Hardy Weinberg Equilibrium identified five loci (rs1426654, rs3827760, rs3811801, rs1229984, and rs671), three loci (rs192655, rs7657799, and rs9522149), two loci (rs6990312 and rs200354), and two loci (rs1572018 and rs4918664) in the ASN, AFA, CAU, and HIS populations, respectively, that deviated from expectations ($p < 0.05$). After adjusting for multiple tests with a Bonferroni correction, no significant deviations from Hardy Weinberg Equilibrium ($p < 0.00089$; Supplementary Table 8) were detected. Tests for linkage disequilibrium were completed for each population with an α of 0.05. After Bonferroni correction, 57 pairwise comparisons (33, 13, 3, and 8 pairwise comparisons observed in the ASN, AFA, CAU, and HIS populations, respectively) generated a significant *p*-value (<0.00003247; Supplementary Table 9). Twenty of these 57 pairwise comparisons with significant linkage disequilibrium included syntenic loci and would be expected to demonstrate linkage disequilibrium. Ancestry estimates, completed with the ForenSeq UAS software, were produced for all 676 samples. Examples of these ancestry predictions can be seen in Fig. 3. The ancestry predictions generated with the aiSNP genotypes from all 676 samples are displayed in a principal component analysis plot shown in Fig. 4. This principal component analysis plot illustrates the clustering of the four major populations analyzed and provides an overall visual summary of the ancestry predictions for the 676 samples evaluated in this study. Generally, samples cluster by population affiliation as self-declared. As expected, HIS population samples do overlap with other populations. Since samples are anonymized, actual population affiliation could not be verified by any other means, and some of the overlap in the clustering of populations seen in Fig. 4 could be due to inaccurately self-declared population affiliations. Some

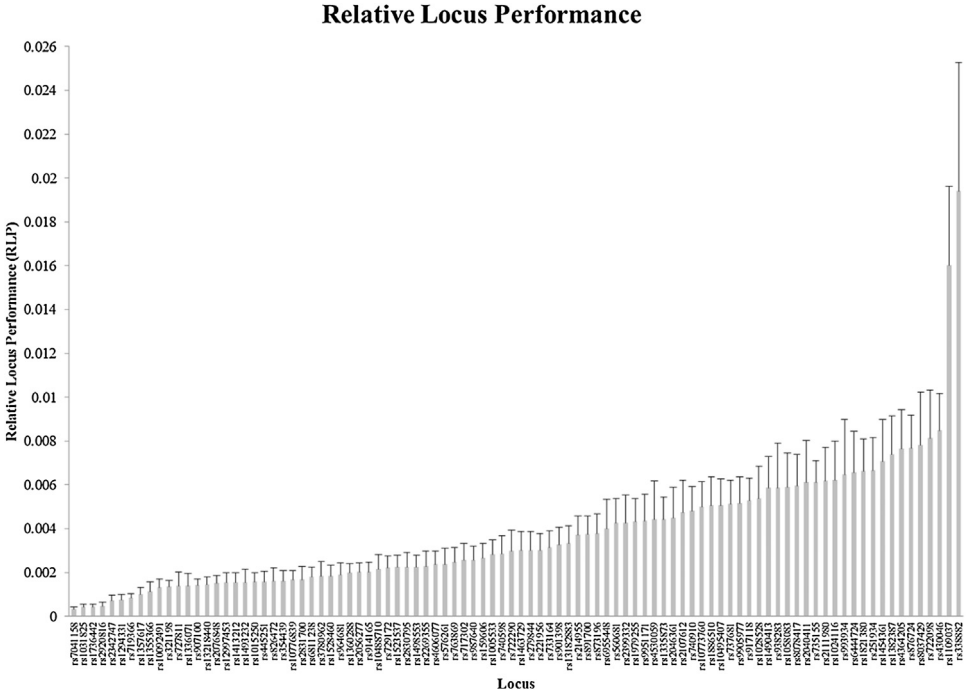

Fig. 1. Average relative locus performances of the iiSNPs for the entire sample set (N = 725) ranged from 0.00031 (±0.00010) to 0.01940 (±0.00588). Error bars represent standard deviation.

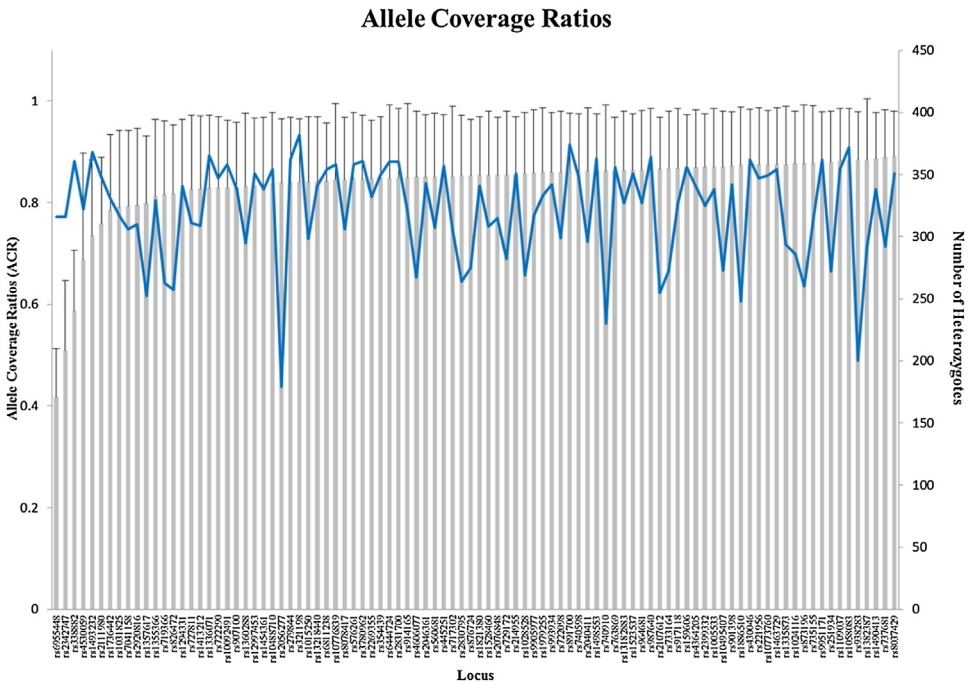

Fig. 2. Average allele coverage ratios of the iiSNPs for the entire sample set (N = 725) ranged from 0.42 (±0.10) to 0.89 (±0.09), where an allele coverage ratio of 1.0 indicated alleles were equal in read depth. Error bars represent standard deviation. Horizontal line denotes number of heterozygous individuals.

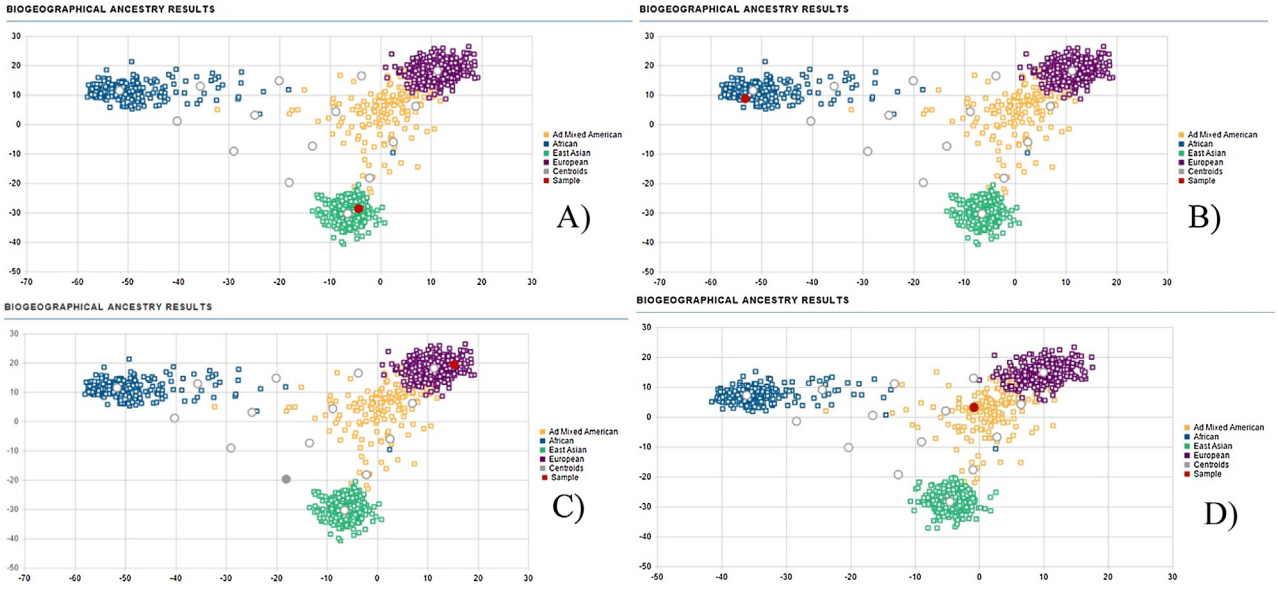

Fig. 3. Examples of ancestry predictions using the ForenSeq UAS software for each population included in the study. The red circle represents the unknown sample. A) Principal component analysis plot for a Chinese sample included in the study. B) Principal component analysis plot for an African American sample included in the study. C) Principal component analysis plot for a Caucasian sample included in the study. D) Principal component analysis plot for a Southwest Hispanic sample included in the study. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of the overlap in the population clusters also could be the result of admixture. Lastly, some of the overlap in the population clusters could be a limitation of the aiSNPs included in this panel to sufficiently cluster each population.

Performance evaluations were completed on all 54 aiSNP loci using metrics such as profile completeness, read depth, relative locus performance, and allele coverage ratios to evaluate the performance of the ForenSeq multiplex. There were 36,355 individual locus genotype calls made out of the possible 36,504 locus genotype for typing 54 aiSNPs on 676 samples with the thresholds set for analysis of this data set. Therefore, the individual locus genotyping success rate was 99.59%. Eight individual locus genotypes did not meet the allele coverage ratio threshold, and 141 individual locus genotypes did not meet the read depth threshold. Of the 676 samples typed for aiSNPs, 608 (89.94%) generated complete profiles, and 673 samples (99.56%) generated profiles that were 90% complete or greater. All 676 aiSNP profiles were 62.96% complete or greater. Twenty-seven of the 54 aiSNPs (50.0%) were typed in all 676 samples in the data set. The rs1572018 locus failed to type in 30 of the 676 samples (4.44%), and rs3737576 failed to type in 25 of the 676 samples (3.70%) based on the analytical threshold set for the analysis of this data set. The rs1572018 locus

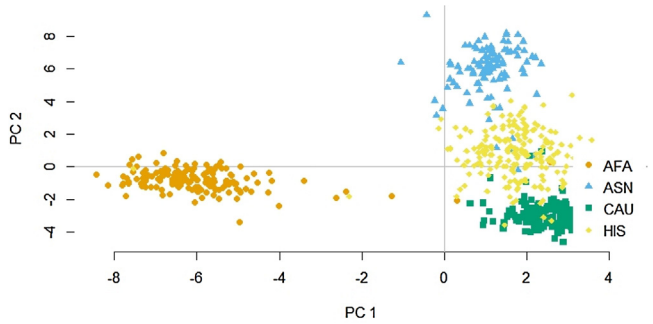

Fig. 4. Principal component analysis plot generated with the aiSNP genotypes from all 676 samples in RStudio®.

was the lowest performing aiSNP in terms of relative locus performance (Fig. 5) and was the second lowest aiSNP in terms of allele coverage ratio (Fig. 6). The rs3737576 locus was the third lowest performing aiSNP in terms of relative locus performance (Fig. 5) and the fourth lowest performing aiSNP in terms of allele coverage ratio (Fig. 6). Average read depth for the aiSNPs ranged from 130.67X (±55.63X) to 2573.35X (±978.25X), with similar spreads observed across the populations analyzed. Average relative locus performance ranged from 0.00050 (±0.00015) to 0.00973 (±0.00220) across the 54 aiSNPs (Fig. 5; Supplementary Fig. 3). Average allele coverage ratios ranged from 0.68 (±0.15) to 0.90 (±0.10) across the 54 aiSNPs (Fig. 6; Supplementary Fig. 4).

3.3. Phenotype informative SNPs

The 24 piSNPs included in primer mix B of the ForenSeq DNA Signature Prep Kit were typed on a total of 676 samples. Allele counts and allele frequencies for the piSNPs were calculated for each population at each locus and are listed in Supplementary Table 10. Tests for departures from Hardy Weinberg Equilibrium identified one locus, rs4959270 in the AFA population, that deviated from expectations ($p < 0.05$). After adjusting for multiple tests with a Bonferroni correction, no significant deviations from Hardy Weinberg Equilibrium ($p < 0.0021$; Supplementary Table 11) were detected. Tests for linkage disequilibrium were performed for each population with an α of 0.05. A total of 13, 12, 10, and 15 pairwise comparisons in the ASN, AFA, CAU, and HIS populations, respectively, demonstrated significant linkage disequilibrium ($p < 0.05$). After Bonferroni correction, seven pairwise comparisons still were significant ($p < 0.00018$; Supplementary Table 12). All seven of the pairwise comparisons with significant linkage disequilibrium were between syntenic loci (Supplementary Table 12). These results were expected considering the 24 piSNPs are located within 11 different genes [44]. Phenotype estimates, completed with the ForenSeq UAS software, require genotype data from all 24 piSNPs [35]. Therefore, phenotype estimates were produced for 662 of the samples out of the 676 samples typed for the 24 piSNPs. Examples of these phenotype estimates are shown in Fig. 7. Probabilities for each sample's phenotype estimate are not available in a downloadable format from the UAS [35], and averages for hair and eye color estimates across populations could not be provided. However, phenotype estimations were generally concordant with the population with US Caucasians showing the greatest distribution of results (Fig. 7), as expected.

Performance metrics, including profile completeness, read depth, relative locus performance, and allele coverage ratios, were used to evaluate the performance of the ForenSeq multiplex. There were 16,202 individual locus genotype calls made out of the possible 16,224 locus genotype calls for typing 24 piSNPs on 676 samples with the analytical thresholds set for analysis of this data set. Therefore, the individual locus genotyping success rate was 99.86%. Two individual locus genotypes did not meet the allele coverage ratio threshold, and 20 individual locus genotypes did not meet the read depth threshold. Of the 676 samples typed for piSNPs, 662 (97.93%) generated complete profiles. All 676 piSNP profiles were 79.17% complete or greater. Eighteen of the 24 piSNPs (75.0%) were typed in all 676 samples in the data set. The rs683 locus failed to type in eight of the 676 samples (1.18%), and the rs12821256 locus failed to type in six of the 676 samples (0.89%)

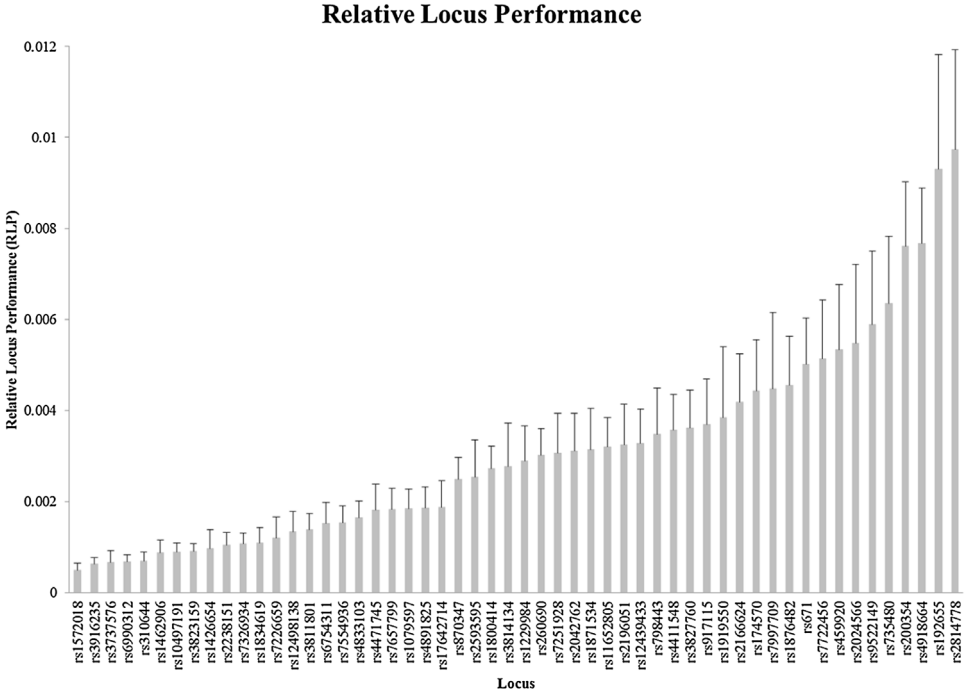

Fig. 5. Average relative locus performances of the aiSNPs for the entire sample set (N = 676) ranged from 0.00050 (±0.00015) to 0.00973 (±0.00220). Error bars represent standard deviation.

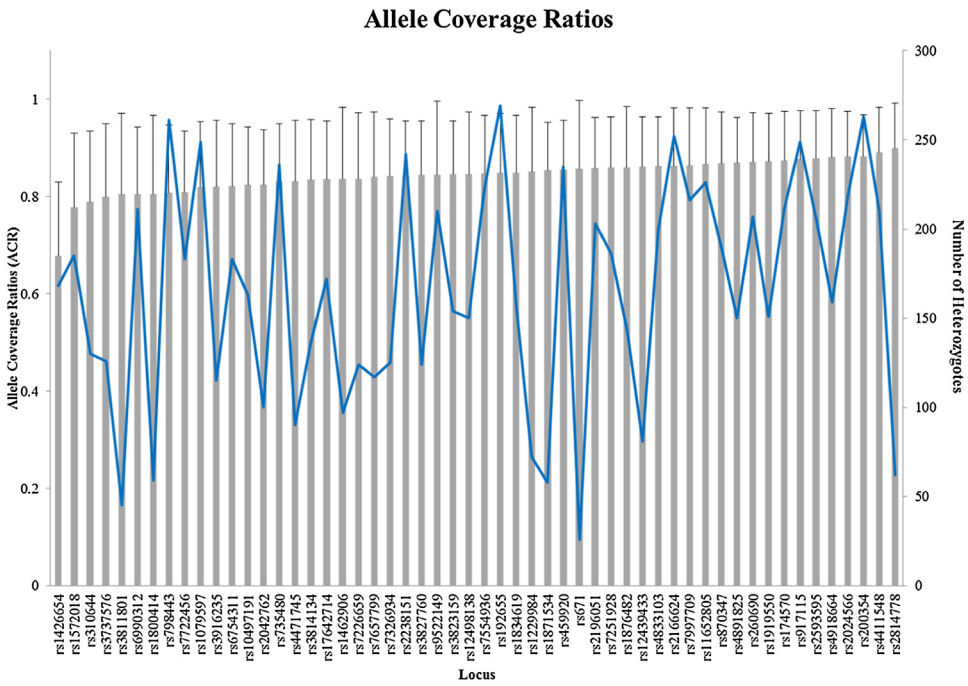

Fig. 6. Average allele coverage ratios of the aiSNPs for the entire sample set (N = 676) ranged from 0.68 (±0.15) to 0.90 (±0.10), where an allele coverage ratio of 1.0 indicated alleles were balanced. Error bars represent standard deviation. Horizontal line denotes number of heterozygous individuals.

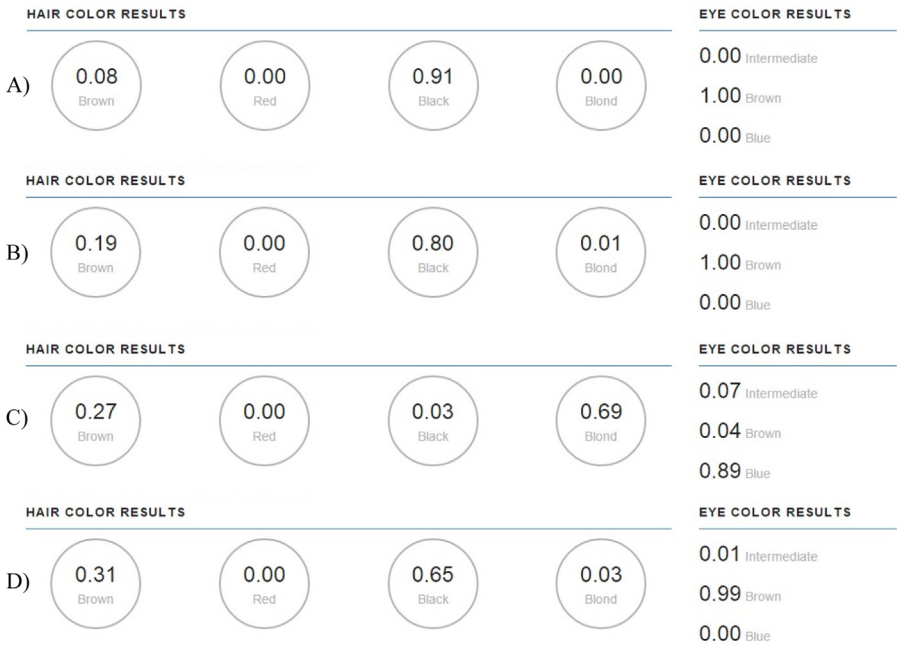

Fig. 7. Examples of phenotype estimates using the ForenSeq UAS software for each population included in the study. A) Phenotype estimate for a Chinese sample included in the study. B) Phenotype estimate for an African American sample included in the study. C) Phenotype estimate for a Caucasian sample included in the study. D) Phenotype estimate for a Southwest Hispanic sample included in the study.

with the thresholds set for the analysis of this data set. The rs12821256 locus was also the lowest performing piSNP in terms of relative locus performance (Fig. 8). Average read depth for the piSNPs ranged from 173.80X (±71.56X) to 3379.94X (±1361.93X), with similar spreads observed across all populations analyzed. Average relative locus performance ranged from 0.00066 (±0.00018) to 0.01273 (±0.00333) across the 24 piSNPs (Fig. 8; Supplementary Fig. 5). Average allele coverage ratios ranged from 0.67 (±0.07) to 0.93 (±0.09) across the 24 piSNPs (Fig. 9; Supplementary Fig. 6).

3.4. STRs

The 59 STRs (28 autosomal STRs, 24 Y-STRs, seven X-STRs, and Amelogenin) included in both primer mix A and primer mix B of the ForenSeq DNA Signature Prep Kit were typed on 725 samples.

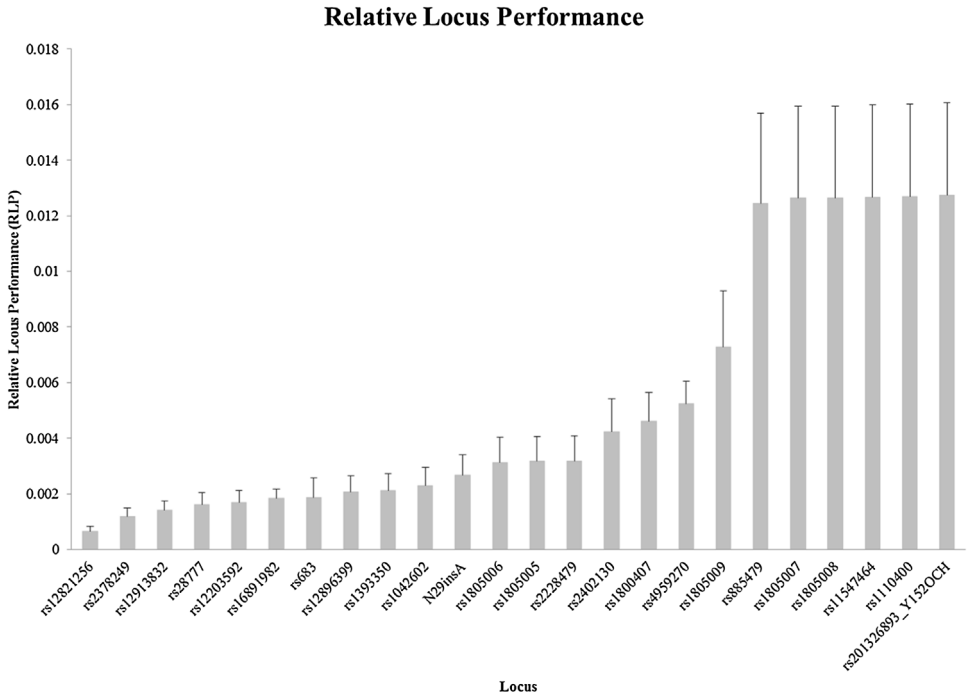

Fig. 8. Average relative locus performances of the piSNPs for the entire sample set (N = 676) ranged from 0.00066 (±0.00018) to 0.01273 (±0.00333). Error bars represent standard deviation.

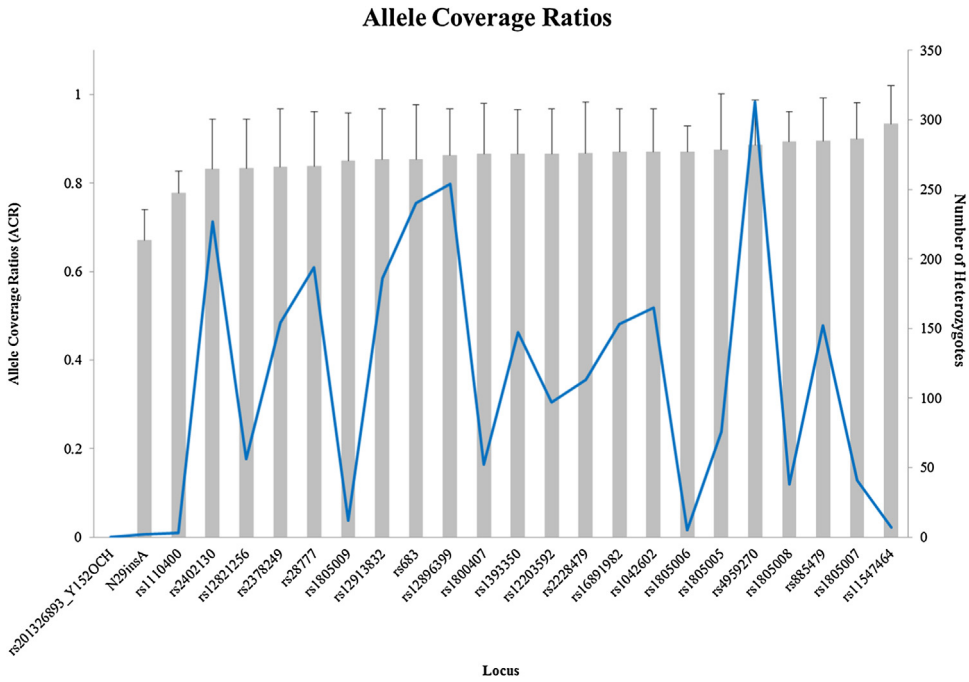

Fig. 9. Average allele coverage ratios of the piSNPs for the entire sample set (N = 676) ranged from 0.67 (±0.07) to 0.93 (±0.09), where an allele coverage ratio of 1.0 indicated alleles were balanced. Error bars represent standard deviation. Horizontal line denotes number of heterozygous individuals.

Allele counts, allele frequencies, and population genetic analyses for the STR loci with these samples are described by Novroski et al. [6]. Since the STR profiles discussed in Novroski et al. [6] were generated with STRait Razor [36,37], bioinformatic concordance between the STR profiles generated by STRait Razor and the STR profiles generated by the UAS software was assessed for a subset of the samples. The bioinformatic concordance evaluation identified a few inconsistencies between the two software options. The majority of these inconsistencies stemmed from the modified STRait Razor's ability to incorporate the flanking regions surrounding the STRs for analysis while the UAS analyses only the repeat region of the STR. Thus, while the repeat region allele calls were the same, the allele designations provided by STRait Razor when analyzing the repeat region and the flanking region differed. For the PentaD locus, one CAU sample was designated an 11/14 by the UAS and an 11/13.4 by STRait Razor. A deletion was identified in the

flanking region of the 14 allele when using STRait Razor, which is the reason for the differing allele designations. Three samples (two AFA samples and one CAU sample) received differing allele designations from STRait Razor and the UAS for the DXS10074 locus due to a two base-pair deletion in the immediate flanking region of DXS10074 identified by STRait Razor. Thus, the UAS assigned a 14/16, 14/17, and 8/17 to the three samples while STRait Razor assigned a 14.2/16, 14.2/17, 8/16.2 to the same three samples. Additionally, STRait Razor and the UAS produced differing genotype calls for the DYS612 locus with the entire sample set. These inconsistencies were described previously in Novroski et al. [6] and were not labeled as discordant as the reported regions differ between the two methods in each instance.

The bioinformatic concordance evaluation highlighted how the D22S1045 locus' performance affected accurate genotyping (detailed below). Apparent "bioinformatics" dropout did occur, however, at two loci (DXS7132 and HPRTB) seemingly associated with flanking region variation. The dropout observed in DXS7132 seems to be associated with a T → C substitution in the first base after the motif (hg38:X-65435703) in four alleles in one population (ASN). The alleles for this locus were present in both STRait Razor data (>2000X read depth) and Nextera Rapid Capture data (unpublished data). The allelic dropout at locus HPRTB appears to be in relation to a six nucleotide deletion in the left flank reported by Novroski et al. [6] as HPRTB [CE10.2]-ChrX-GRCh38 134481429-134481588 (ATCT)12 134481484-134481489 DEL found in one CAU sample. These discordances have been communicated to the manufacturer. These examples, while consistent within a method, further indicate the need for orthogonal testing with bioinformatics pipelines.

Tests for linkage disequilibrium were performed between the iiSNP genotypes and autosomal STRs with an α of 0.5. After Bonferroni correction, five, 134, 147, and 108 pairs of loci in the ASN, AFA, CAU, and HIS populations, respectively, generated significant linkage disequilibrium ($p < 0.00000689$; Supplementary Table 13). A few loci previously identified as departing from Hardy Weinberg Equilibrium in their respective populations (i.e., D7S820 in the ASN population; D13S317 and D16S539 in the AFA population; D13S317, D5S818, and D7S820 in the CAU population; D16S539, D20S482, and D7S820 in the HIS population) [6] were associated with 372 of the 394 pairwise comparisons with significant linkage disequilibrium. This effect of a locus significantly departing Hardy Weinberg Equilibrium expectations and causing apparent linkage disequilibrium has been described by Falush et al. [49] and Chakraborty [50]. None of the 22 remaining pairwise comparisons include syntenic loci. The allele frequencies reported above and in Novroski et al. [6] for the iiSNP and autosomal STR loci (length-based genotype calls and sequence-based genotype calls) were used to calculate an overall identity marker random match probability value (under the assumption of independence) for each of the four populations (Table 1).

The performance metrics of the STR loci were profile completeness, read depth, relative locus performance, and allele coverage ratios. There were 42,273 individual locus genotype calls made out of the possible 42,775 locus genotype calls typing the 59 STRs on 725 samples with the thresholds set for analysis of this data set. Therefore, the individual locus genotyping success rate was 98.83%. A total of 303 individual locus genotypes did not meet the allele coverage ratio threshold, and 178 individual locus genotypes did not meet the read depth threshold. An additional 21 genotype calls were not made for the DYS392 locus due to excessive stutter peaks which did not allow for accurate genotyping. Of the 725 samples typed for STRs, 426 (58.76%) generated complete profiles, and 720 (99.31%) generated profiles that were 90% complete or greater. All 725 STR profiles were 79.66% complete or greater. Seventeen of the 59 STRs (28.81%) had genotype calls for all 725 samples in the data set. Genotypes for the D6S1043 and DYS385a-b loci were not generated for 12 (1.66%) and 31 (4.28%) samples, respectively, with the thresholds set for the analysis of this data set. The DXS10103, DYS392, and D22S1045 loci were the poorest performing STRs with genotypes missing from 84 (11.59%), 88 (12.14%), and 172 (23.72%) samples, respectively, with the thresholds set for the analysis of this data set. The DXS10103 locus was the lowest performing X-STR in terms of read depth and relative locus performance (Fig. 11) and allele coverage ratio in female samples (Fig. 14). The D22S1045 locus was the lowest performing autosomal STR in terms of allele coverage ratio (Fig. 13).

Similarly, both D22S1045 and DYS392 varied greatly across the data set as evidenced by the substantial standard deviation in read depth for both loci (Figs. 10 and 12). To further elucidate the nature of this variation, read depth of each individual allele was plotted for each locus (Fig. 15). The steep reduction in read depth (slope = −526X and −841X, respectively) across the allelic range may complicate mixture interpretation and should be strongly considered prior to interpretation. This allele-specific imbalance presented as locus dropout in DYS392 and severe heterozygote imbalance with the potential for allele dropout at locus D22S1045. While a low allele coverage ratio allows for interpretation, validation studies are necessary for developing thresholds that account for loci with severe heterozygote imbalance such as the D22S045 locus. These loci share a common motif of ATT and ATA, respectively. Thus, the observed variance may be due to some sequence-specific limitation or secondary structure with regards to massively parallel sequencing as alluded to in Novroski et al. [6].

Average read depth for the autosomal STRs ranged from 638.54X (±243.20X) to 6956.39X (±2880.82X) with similar spreads observed across the populations analyzed. Average read depth for the X-STRs ranged from 144.01X (±56.43X) to 5307.02X (±1853.46X) in female samples and from 58.92X (±27.60X) to 2133.63X (±1009.82X) in male samples, with similar spreads observed in all four populations. The approximate two-fold difference in read depth of the X-STRs between male and female samples is most likely due to the difference in the number of X-chromosomes in males and females. Average read depth for the Y-STRs in the male samples ranged from 431.99X (±188.38X) to 3739.53X (±1652.61X), with similar spreads observed in all four populations. Average relative locus performance for the autosomal STRs ranged from 0.00251 (±0.00059) to 0.02735 (±0.00752) (Fig. 10; Supplementary Fig. 7). Average relative locus performance

Table 1
Combined random match probability values for the iiSNP and autosomal STR loci.

Population	Random Match Probability	
	iiSNP and Length-based STR Genotype Calls	iiSNP and Sequence-based STR Genotype Calls
ASN	6.56E-67	8.92E-71
AFA	1.02E-68	1.57E-74
CAU	2.82E-69	1.63E-73
HIS	3.16E-68	2.57E-72

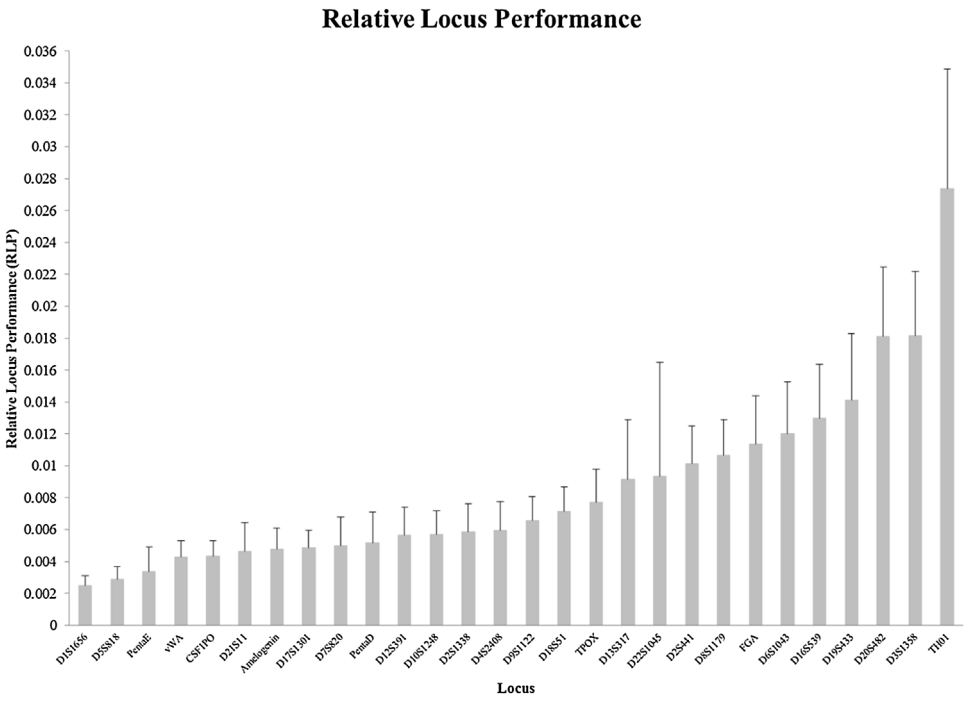

Fig. 10. Average relative locus performances of the autosomal STRs for the entire sample set (N = 725) ranged from 0.00251 (±0.00059) to 0.02735 (±0.00752). Error bars represent standard deviation.

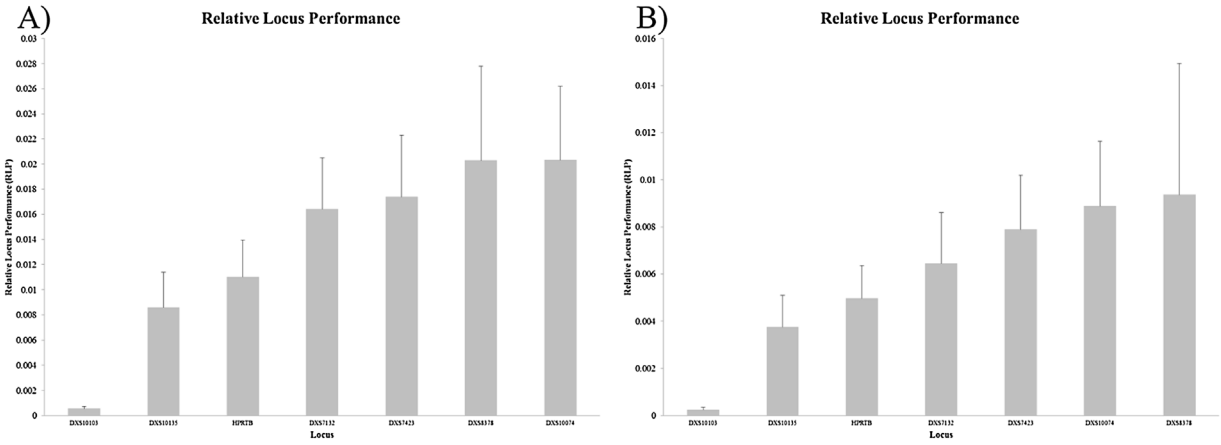

Fig. 11. Average relative locus performances of the X-STRs ranged from 0.00055 (±0.00017) to 0.02032 (±0.00588) for females (A; N = 453) and from 0.00025 (±0.00011) to 0.00937 (±0.00558) for males (B; N = 272). Error bars represent standard deviation.

for the X-STRs ranged from 0.00055 (±0.00017) to 0.02032 (±0.00588) in females and from 0.00025 (±0.00011) to 0.00937 (±0.00558) in males (Fig. 11; Supplementary Fig. 8 and 9). Average relative locus performance for the Y-STRs ranged from 0.00191 (±0.00084) to 0.01561 (±0.00421) in male samples (Fig. 12; Supplementary Fig. 10). Average allele coverage ratios for the autosomal STRs ranged from 0.33 (±0.19) to 0.88 (±0.12) (Fig. 13; Supplementary Fig. 11). Average allele coverage ratios for the X-STRs in females ranged from 0.79 (±0.15) to 0.88 (±0.10) (Fig. 14; Supplementary Fig. 12).

4. Conclusion

Prior to using the larger, mixed-marker panels that massively parallel sequencing affords, detailed performance evaluations regarding performance metrics are needed for each of the genetic markers included in the MiSeq FGx Forensic Genomics System. The performance metrics evaluated in this study indicated that quality and robust data were produced, which support the potential future use of this kit in forensic genetic laboratories. Most of the loci were typable. However, one should be aware that a few loci were under performers. The list of performance metrics for all 231 genetic markers across four separate populations also offers valuable information and guidance for establishing interpretation guidelines on data generated by massively parallel sequencing for forensic genetics applications.

Bioinformatic discordance was observed in two loci using UAS software version 1.2.16173 and while these alleles were low-frequency (i.e., below the minimum allele frequency), it supports the need for validating software. Allele frequencies from relevant populations, such as the ones reported herein, are necessary to enable calculation of the statistical weight of the massively parallel sequencing-generated DNA results. Allele frequencies for the 172 SNPs included in the MiSeq FGx Forensic Genomics kit were

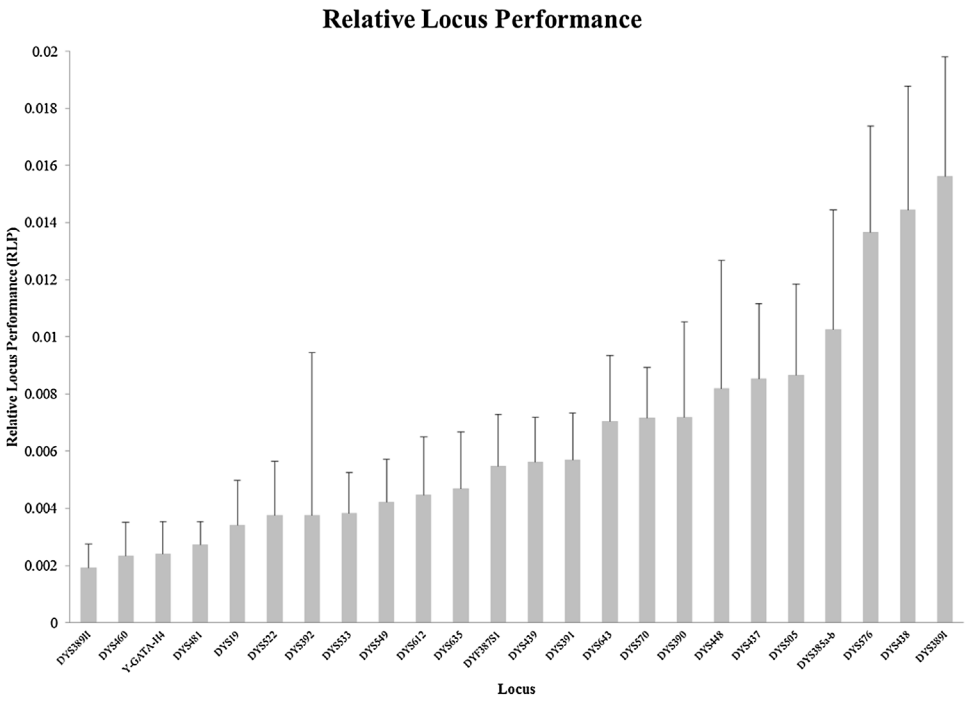

Fig. 12. Average relative locus performances of the Y-STRs ranged from 0.00191 (±0.00084) to 0.01561 (±0.00421) for males (N = 272). Error bars represent standard deviation.

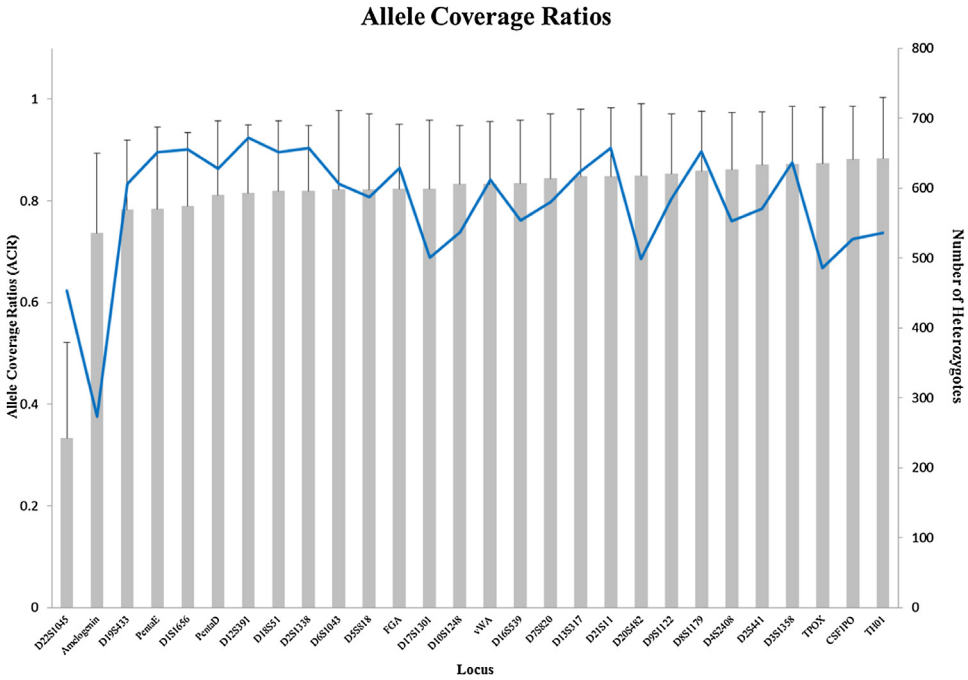

Fig. 13. Average allele coverage ratios of the autosomal STRs for the entire sample set (N = 725) ranged from 0.33 (±0.19) to 0.88 (±0.12), where an allele coverage ratio of 1.0 indicated alleles were balanced. Error bars represent standard deviation. Horizontal line denotes number of heterozygous individuals.

reported for 725 Chinese, African American, US Caucasian, and Southwest Hispanic samples. STR allele frequencies for the 59 STRs included in the MiSeq FGx Forensic Genomics kit were reported previously [6].

The combined iiSNP random match probability values ranged from 4.49E-38 to 1.03E-35 for the four populations evaluated. When including STRs, the combined random match probability values for the iiSNP and sequence-based STR genotype calls ranged from 1.57E-74 to 8.92E-71 for the four populations. Ancestry and phenotype predictions for each individual sample were possible for most samples and thus provide additional information that can

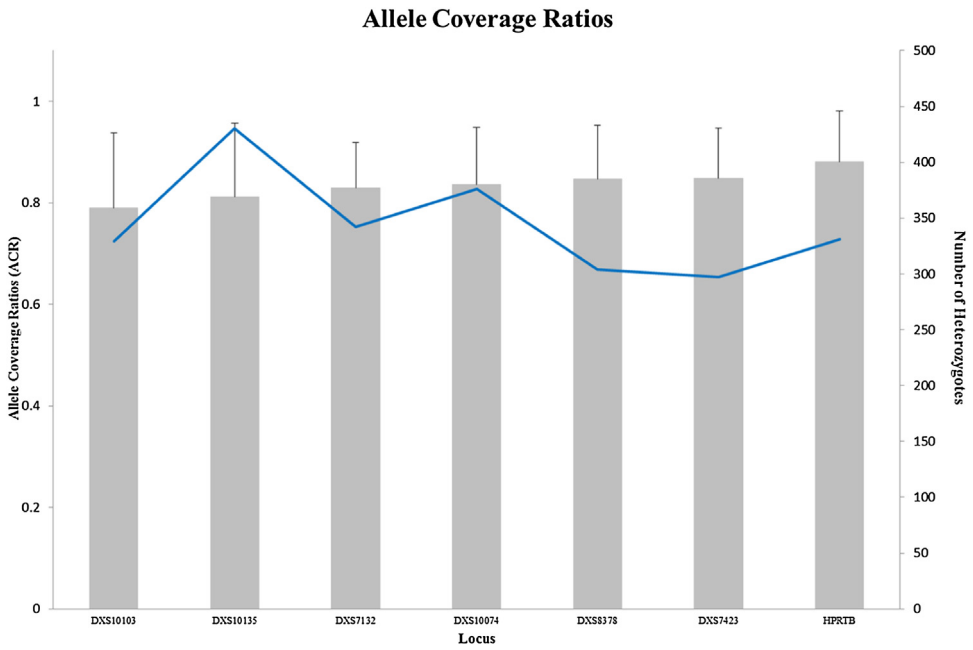

Fig. 14. Average allele coverage ratios of the X-STRs for the female samples (N = 453) ranged from 0.79 (±0.15) to 0.88 (±0.10), where an allele coverage ratio of 1.0 indicated alleles were balanced. Error bars represent standard deviation. Horizontal line denotes number of heterozygous individuals.

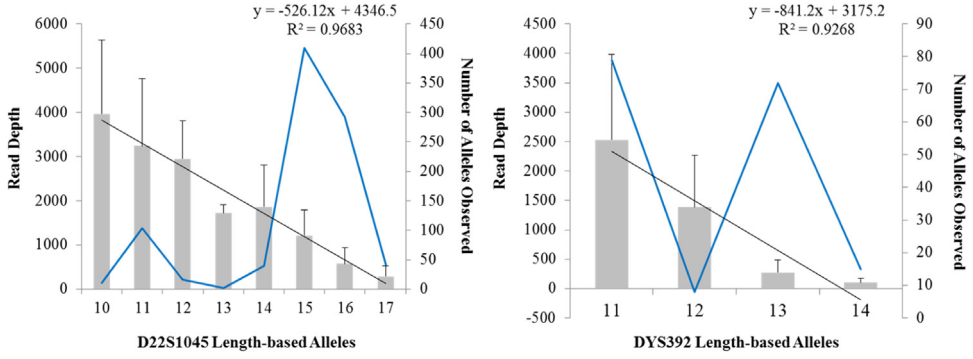

Fig. 15. Average read depth by length-based allele for N = 459 and N = 174 individuals for D22S1045 and DYS392, respectively. Error bars represent standard deviation. Horizontal line denotes number of heterozygous individuals.

be provided for investigative leads in cases where such data are useful.

Acknowledgements

The authors would like to thank Illumina, in particular Al Bodota, Cydne Holt, and Joe Varlaro, for study support. This work was supported in part by award no. 2015-DN-BX- K067, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the U.S. Department of Justice.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2017.06.004.

References

[1] C. Børsting, N. Morling, Next generation sequencing and its applications in forensic genetics, Forensic Sci. Int. Genet. 18 (2015) 78–89.
[2] C. Børsting, S.L. Fordyce, J. Olofsson, H.S. Mogensen, N. Morling, Evaluation of the Ion Torrent™ HID SNP 169-plex: a SNP typing assay developed for human identification by second generation sequencing, Forensic Sci. Int. Genet. 12 (2014) 144–154.
[3] C.G. Bottino, C.W. Chang, S. Wootton, N. Rajagopalan, R. Langit, R.E. Lagace, R. Silva, R.S. Moura-Neto, STR genotyping using ion torrent PGM and STR 24-plex system: performance and data interpretation, Forensic Sci. Int. Genet. SS5 (2015) e325–e326.
[4] J.D. Churchill, J. Chang, J. Ge, N. Rajagopalan, S.C. Wootton, C.W. Chang, R. Lagacé, W. Liao, J.L. King, B. Budowle, Blind study evaluation illustrates utility of the Ion PGM™ system for use in human identity DNA typing, Croat. Med. J. 56 (2015) 218–229.
[5] J.D. Churchill, S.E. Schmedes, J.L. King, B. Budowle, Evaluation of the Illumina® Beta Version ForenSeq™ DNA Signature Prep Kit for use in genetic profiling, Forensic Sci. Int. Genet. 20 (2016) 20–29.
[6] N.M.M. Novroski, J.L. King, J.D. Churchill, S.L. Hong, B. Budowle, Characterization of genetic sequence variation of 58 STR Loci in four major population groups, Forensic Sci. Int. Genet. 25 (2016) 214–226.
[7] M. Eduardoff, C. Santos, M. de la Puente, T.E. Gross, M. Fondevila, C. Strobl, B. Sobrino, D. Ballard, P.M. Schneider, A. Carracedo, M.V. Lareu, W. Parson, C. Phillips, Inter-laboratory evaluation of SNP-based forensic identification by massively parallel sequencing using the Ion PGM™, Forensic Sci. Int. Genet. 17 (2015) 110–121.

[8] M. Eduardoff, T.E. Gross, C. Santos, M. De la Puente, D. Ballard, C. Strobl, C. Børsting, N. Morling, L. Fusco, C. Hussing, B. Egyed, L. Souto, J. Uacyisrael, D. Syndercombe Court, Á. Carracedo, M.V. Lareu, P.M. Schneider, W. Parson, C. Phillips, The EUROFORGEN-NoE Consortium, W. Parson, C. Phillips, Inter-laboratory evaluation of the EUROFOGEN Global ancestry-informative SNP panel by massively parallel sequencing using the Ion PGM™, Forensic Sci Int Genet. 23 (2016) 178–189.

[9] F.R. Wendt, J.D. Churchill, N.M.M. Novroski, J.L. King, J. Ng, R.F. Oldt, K. McCulloh, J.A. Weise, D.G. Smith, S. Kanthaswamy, B. Budowle, Genetic analysis of the Yavapai Native Americans from West-Central Arizona using the Illumina MiSeq FGx™ forensic genomics system, Forensic Sci. Int. Genet. 24 (2016) 18–23.

[10] F.E. Wendt, X. Zeng, J.D. Churchill, J.L. King, B. Budowle, Analysis of short tandem repeat (STR) and single nucleotide polymorphism (SNP) loci from single source samples using a custom HaloPlex Target Enrichment System panel, Am. J. Forensic Med. Pathol. 37 (2) (2016) 99–107.

[11] X. Zeng, J. King, S. Hermanson, J. Patel, D.R. Storts, B. Budowle, An evaluation of the PowerSeq™ Auto System: a multiplex short tandem repeat marker kit compatible with massively parallel sequencing, Forensic Sci. Int. Genet. 19 (2015) 172–179.

[12] D.H. Warshauer, C.P. Davis, C. Holt, Y. Han, P. Walichiewicz, T. Richardson, K. Stephens, A. Jager, J. King, B. Bruce Budowle, Massively parallel sequencing of forensically relevant single nucleotide polymorphisms using TruSeq™ forensic amplicon, Int. J. Legal Med. 129 (2015) 31–36.

[13] C.X. Li, A.J. Pakstis, L. Jiang, Y.L. Wei, Q.F. Sun, H. Wu, O. Bulbul, P. Wang, L.L. Kang, J.R. Kidd, K.K. Kidd, A panel of 74 AISNPs: improved ancestry inference within eastern asia, Forensic Sci. Int. Genet. 23 (2016) 101–110.

[14] F. Guo, Y. Zhou, F. Liu, J. Yu, H. Song, H. Shen, B. Zhao, F. Jia, G. Hou, X. Jiang, Evaluation of the early access STR Kit v1 on the ion torrent PGM™ platform, Forensic Sci. Int. Genet. 23 (2016) 111–120.

[15] I. Grandell, R. Samara, A.O. Tillmar, A SNP panel for identity and kinship testing using massive parallel sequencing, Int. J. Legal Med. 130 (2016) 904–914.

[16] K.B. Gettings, K.M. Kiesler, S.A. Faith, E. Montano, C.H. Baker, B.A. Young, R.A. Guerrieri, P.M. Vallone, Sequence variation of 22 autosomal STR loci detected by next generation sequencing, Forensic Sci. Int. Genet. 21 (2016) 15–21.

[17] S.L. Fordyce, H.S. Mogensen, C. Borsting, R.E. Lagace, C.W. Chang, N. Rajagopalon, N. Morling, Second-generation sequencing of forensic STRs using he Ion Torrent™ HID STR 10-plex and the Ion PGM™, Forensic Sci. Int. Genet. 14 (2015) 132–140.

[18] A.L. Silvia, N. Shugarts, J. Smith, A preliminary assessment of the ForenSeq™ FGx System: next generation sequencing of an STR and SNP multiplex, Int. J. Legal Med. 131 (2016) 73–86.

[19] B. Budowle, A. van Daal, Forensically relevant SNP classes, BioTechniques 44 (2008) 603–610.

[20] K. Butler, M. Peck, J. Hart, M. Schanfield, D. Daniele Podini, Molecular “eyewitness”: Forensic prediction of phenotype and ancestry, Forensic Sci. Int. Genet. SS3 (2011) e498–e499.

[21] L. Chaitanya, I. Zupanic Pajnic, S. Walsh, J. Balazic, T. Zupanc, M. Kayser, Bringing colour back after 70 years: predicting eye and hair colour from skeletal remains of World War II victims using the HIrisPlex system, Forensic Sci Int Genet. 26 (2017) 48–57.

[22] J. Draus-Barini, S. Walsh, E. Pośpiechb, T. Kupiec, H. Głąb, W. Wojciech Branicki, M. Kayser, Bona fide colour: DNA prediction of human eye and hair colour from ancient and contemporary skeletal remains, Invest. Genet. 4 (2013) 1–15.

[23] M. Kayser, Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes, Forensic Sci. Int. Genet. 18 (2015) 33–48.

[24] C. Hollard, C. Keyser, T. Delabarde, A. Gonzalez, C.V. Lamego, V. Zvénigorosky, B. Ludes, Case report: on the use of the HID-Ion AmpliSeq™ Ancestry Panel in a real forensic case, Int. J. Legal Med. 131 (2017) 351–358.

[25] H. Boonyarit, S. Mahasirimongkol, N. Chavalvechakul, M. Aoki, H. Amitani, N. Hosono, N. Kamatani, M. Kubo, P. Lertrit, Development of a SNP set for human identification: a set with high powers of discrimination which yields high genetic information from naturally degraded DNA samples in the Thai population, Forensic Sci. Int. Genet. 11 (2014) 166–173.

[26] C. Børsting, H.S. Mogensen, N. Morling, Forensic genetic SNP typing of low-template DNA and highly degraded DNA from crime case samples, Forensic Sci. Int. Genet. 7 (2013) 345–352.

[27] K.B. Gettings, R.A. Aponte, P.M. Vallone, J.M. Butler, STR allele sequence variation: current knowledge and future issues, Forensic Sci. Int. Genet. 18 (2015) 118–130.

[28] K.B. Gettings, R.A. Aponte, K.M. Kiesler, P.M. Vallone, The next dimension in STR sequencing: polymorphisms in flanking regions and their allelic associations, Forensic Sci. Int. Genet. SS5 (2015) e121–e123.

[29] Illumina, ForenSeq™ DNA Signature Prep Reference Guide, September 2015.

[30] A. Ambers, J.D. Churchill, J.L. King, M. Stoljarova, H. Gill-King, M. Assidi, M. Abu-Elmagd, A. Buhmeida, B. Budowle, More comprehensive forensic genetic marker analyses for accurate human remains identification using massively parallel sequencing, BMC Genom. 17 (2016) 21–30.

[31] F. Calafell, R. Anglada, N. Bonet, M. Gonzalez-Ruiz, G. Prats-Munoz, R. Rasal, C. Lalueza-Fox, J. Bertranpetit, A. Malgosa, F. Casals, An assessment of a massively parallel sequencing approach for the identification of individuals from mass graves of the Spanish Civil War (1936–1939), Electrophoresis 00 (2016) 1–7.

[32] Qiagen, QIAamp® DNA Mini and Blood Mini Handbook, June, 2012.

[33] Illumina, ForenSeq™ DNA Signature Prep Protocol Guide, September 2015.

[34] Illumina, MiSeq FGx™ Reagent Kit Reference Guide, January 2015.

[35] Illumina, ForenSeq™ Universal Analysis Software Guide, June 2015.

[36] D.H. Warshauer, J.L. King, B. Budowle, STRait Razor v2.0: the improved STR allele identification tool—Razor, Forensic Sci. Int. Genet. 14 (2015) 182–186.

[37] D.H. Warshauer, D. Lin, K. Hari, R. Jain, C. Davis, B. Larue, J.L. King, B. Budowle, STRait Razor: a length-based forensic STR allele-calling tool for use with second generation sequencing data, Forensic Sci. Int. Genet. 7 (2013) 409–417.

[38] P.O. Lewis, D. Zaykin, Genetic Data Analysis Software, (1999) . URL http://en.bio-soft.net/dna/gda.html.

[39] 1000 Genomes Project Consortium, A. Auton, L.D. Brooks, R.M. Durbin, E.P. Garrison, H.M. Kang, J.O. Korbel, J.L. Marchini, S. McCarthy, G.A. McVean, G.R. Abecasis, A global reference for human genetic variation, Nature 526 (2015) 68–74.

[40] 1000 Genomes Project Consortium, G.R. AutonAbecasis, D. Altshuler, A. Auton, L.D. Brooks, R.M. Durbin, R.A. Gibbs, M.E. Hurles, G.A. McVean, A map of human genome variation from population-scale sequencing, Nature 467 (2010) 1061–1073.

[41] RStudio Team, RStudio: Integrated Development for R.R Studio, Inc., Boston MA 2015; URL http://www.rstudio.com/.

[42] S. Walsh, L. Chaitanya, L. Clarisse, L. Wirken, J. Draus-Barini, L. Kovatsi, H. Maeda, T. Ishikawa, T. Sijen, P. de Knijff, W. Branicki, F. Liu, M. Kayser, Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage, Forensic Sci. Int. Genet. 9 (2014) 150–161.

[43] F. Liu, K. vanDuijn, J.R. Vingerling, A. Hofman, A.G. Uitterlinden, A.C.J.W. Janssens, M. Kayser, Eye color and the prediction of complex phenotypes from genotypes, Curr. Biol. 19 (2009) R192–R193.

[44] S. Walsh, F. Liu, A. Wollstein, L. Kovatsi, A. Ralf, A. Kosiniak-Kamysz, W. Branicki, M. Kayser, The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA, Forensic Sci. Int. Genet. 7 (2013) 98–115.

[45] B.S. Weir, Genetic Data Analysis II: Methods for Discrete Population Genetic Data, vol. 2, Sinauer Associates, Sunderland, 2016, pp. 1996.

[46] S. Wright, Coefficients of inbreeding and relationship, Am. Nat. 56 (1922) 330–338.

[47] S. Wright, The interpretation of population structure by F-statistics with special regard to systems of mating, Evolution 19 (1965) 395–420.

[48] B. Weir, C.C. Cockerham, Estimating F-statistics for the analysis of population structure, Evolution 38 (1984) 1358–1370.

[49] D. Falush, M. Stephens, J.K. Pritchard, Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies, Genetics 164 (2003) 1567–1587.

[50] R. Chakraborty, Linkage Disequilibrium: Concept, Utility and Evolutionary Dynamics in the Context of the Human Genome Variation, DESTOBIO, West Lafayette, Indiana, 2000.