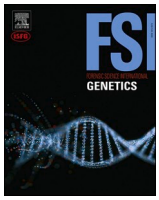
Forensic Science International: Genetics 71 (2024) 103059

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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



Simultaneous sequencing of 102 Y-STRs on Ion Torrent ™ GeneStudio ™ S5 System

Fei Guo [a,b,c,*], Guangxin Jing [d], Yubo Lang [e], Ze Liu [f], Shaobo Yu [f,**]

[a] Shenyang Medical College, Shenyang, Liaoning 110034, PR China
[b] Key Laboratory of Human Ethnic Specificity and Phenomics of Critical Illness in Liaoning Province, Shenyang, Liaoning 110034, PR China
[c] Key Laboratory of Phenomics in Shenyang City, Shenyang, Liaoning 110034, PR China
[d] Heyuan Public Security Forensic Science Center, Heyuan, Guangdong 517000, PR China
[e] School of Public Security Information Technology and Intelligence, Criminal Investigation Police University of China, Shenyang, Liaoning 110854, PR China
[f] DNA Laboratory of Forensic Science Center, Shenyang Public Security Bureau, Shenyang, Liaoning 110002, PR China

ARTICLE INFO

Keywords:
Next-generation sequencing
Ion Torrent ™ GeneStudio ™ S5 System
GrandFiler Y-STR Panel
ForeNGS Analysis Software
Y-chromosomal short tandem repeat

ABSTRACT

The Precision ID NGS System from Thermo Fisher Scientific is a mainstream next-generation sequencing (NGS) platform used in forensic laboratories to detect almost all commonly used forensic markers, except for Y-chromosomal short tandem repeats (Y-STRs). This study aimed to: 1) develop a Y-STR panel compatible with the automatic workflow of the NGS system using Ion AmpliSeq Technology, 2) evaluate the panel performance following the SWGDAM guidelines, and 3) explore the possibility of using a combination workflow to detect autosomal STRs and Y-STRs (AY-STR NGS workflow). The GrandFiler Y-STR Panel was successfully designed using the 'separating' and 'merging' strategies, including 102 Y-STRs and Amelogenin with an average amplicon length of 133 bp. It is a mega Y-STR multiplex system in which up to 16 samples can be sequenced simultaneously on an Ion 530 ™ Chip. Developmental validation studies of the performance of the NGS platform, species specificity, reproducibility, concordance, sensitivity, degraded samples, case-type samples, and mixtures were conducted to unequivocally determine whether the GrandFiler Y-STR Panel is suitable for real scenarios. The newly developed Y-STR panel showed compelling run metrics and NGS performance, including 92.47% bases with $\geq$ Q20, 91.80% effective reads, 2106 $\times$ depth of coverage (DoC), and 97.09% inter-locus balance. Additionally, it showed high specificity for human males and 99.40% methodological and bioinformatical concordance, generated complete profiles at $\geq$ 0.1 ng input DNA, and recovered more genetic information from severely degraded and diverse case samples. Although the outcome when used on mixtures was not as expected, more genetic information was obtained compared to that from capillary electrophoresis (CE) methods. The AY-STR NGS workflow was established by combining the GrandFiler Y-STR Panel with the Precision ID GlobalFiler ™ NGS STR Panel v2 at a 2:1 concentration ratio. The combination workflow on NGS performance, reproducibility, concordance, and sensitivity was as stable as the single Y-STR NGS workflow, providing more options for forensic scientists when dealing with different case scenarios. Overall, the GrandFiler Y-STR Panel was confirmed as the first to effectively detect a large number of Y-STR markers on the Precision ID NGS System, which is compatible with 51 Y-STRs in commercial CE kits and 51 Y-STRs in commercial NGS kits and the STRBase. The panel is as robust, reliable, and sensitive as current CE/NGS kits, and is suitable for solving real cases, especially for severely degraded samples (degradation index $>$ 10).

1. Introduction

Short tandem repeats (STRs) in the non-recombining region of the Y-chromosome are used as lineage markers that can only be detected in male individuals and are passed down from generation to generation, unless mutational events occur [1]. Y-chromosomal STR (Y-STR) markers have proved helpful in forensic science in numerous scenarios, for example: (1) to detect individual male DNA from mixtures with high

* Corresponding author at: Shenyang Medical College, Shenyang, Liaoning 110034, PR China.
** Corresponding author.
E-mail addresses: nevergfaye@gmail.com (F. Guo), 243816450@qq.com (S. Yu).

https://doi.org/10.1016/j.fsigen.2024.103059
Received 31 July 2023; Received in revised form 2 May 2024; Accepted 5 May 2024
Available online 9 May 2024

background of female DNA, (2) to confirm the sex of individuals with Amelogenin Y-deficiency, (3) to highlight multiple male perpetrators in a 'gang rape,' and (4) to identify the paternal lineage of male perpetrators [1].

Over 4000 Y-STRs have been identified since first reported in 1992 [2–6], but only a small proportion of them have been applied to forensic science. A set of 9 Y-STRs were selected as core loci and termed the 'European minimal haplotype' in 1997, including DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, and DYS385a/b [7,8]. Then, DYS438 and DYS439, combined with the 'minimal haplotype' core loci, were recommended for inclusion and termed the 'extended haplotype' by the Scientific Working Group on DNA Analysis Methods (SWGDAM) in 2003 [9]. Since then, based on capillary electrophoresis (CE) technology, manufacturers have launched commercial Y-STR kits varying in the total number of Y-STRs, although all of them include these core sets. The earliest commercial kits were released by Reliagene in 2003, i.e., Y-PLEX 6, multiplexing 6 Y-STRs [10], Promega in 2005, i.e., the PowerPlex ® Y System that incorporates 12 Y-STRs [11], and then, in 2006, i.e., the AmpFℓSTR Yfiler ® PCR Amplification Kit by Thermo Fisher Scientific including 17 Y-STRs [12]. As CE technology gradually improved, more commercial kits capable of detecting up to 20–27 Y-STRs in a single reaction became available in 2015 [13–22]. Thus, by 2023, at least six commercial kits were reported as capable of simultaneously detecting 32–54 Y-STRs [23–29]. However, as it becomes excessively crowded, arranging these Y-STRs within the range of a 550-bp window with six to eight fluorescent dyes becomes increasingly challenging. Thus, new technologies are urgently needed to examine additional loci in a single reaction.

Compared to CE, next-generation sequencing (NGS) technology can not only make STR amplicons as short as possible and simultaneously detect more loci but, additionally, NGS can reveal STR repeat region and flanking region variants. These advantages have driven forensic scientists to engage in systematic efforts to develop NGS Y-STR detection systems. Thus, during 2015–2016, Zhao *et al.* [30] established a 9-plex Y-STR panel on the Ion Torrent PGM platform, while Warshauer *et al.* [31] and Kwon *et al.* [32] developed 28-plex and 23-plex Y-STR panels on the MiSeq platform, respectively. Further, in 2021, the most extended Y-chromosomal NGS panel ('CSYseq') targeting 202 Y-STRs and 9014 Y-chromosomal single nucleotide polymorphisms (Y-SNPs) was developed on the MiSeq System [33]. More recently, 31- and 68-plex Y-STR panels on the MiSeq System have also been reported [34,35]. Among manufacturers, Verogen (now a part of Qiagen) was the first company to release a commercial NGS kit that included 26 Y-STR markers, along with other forensically relevant markers in the ForenSeq ™ DNA Signature Prep Kit [36]. In turn, Promega released the PowerSeq ™ Auto/Y System, which included 23 Y-STRs [37,38], and two more commercial kits that can detect over 40 Y-STRs were recently reported: 48 Y-STRs in the MGIEasy Signature Identification Library Prep Kit from MGI Tech [39] and 81 Y-STRs in the Forensic Analysis System Multiplecues SetB Kit from DeepReads Biotech [40]. Consequently, it has become increasingly convenient for forensic scientists to obtain sequence-based Y-STR data. Meanwhile, Thermo Fisher Scientific launched many commercial kits based on Ion AmpliSeq Technology on the Ion Torrent ™ PGM platform and/or Precision ID NGS System [41–47]. Custom-designed panels have also been developed in the Ion Community [48–51]. These commercial kits and custom-designed panels have covered commonly-used forensic DNA markers in most instances, such as autosomal STRs (A-STRs) [46,47], identity-informative SNPs (iiSNPs) [41,42], ancestry-informative SNPs (aiSNPs) [43,50,51], phenotypic-informative SNPs (piSNPs) [50,51], Y-SNPs [48], microhaplotypes [49], and mitochondrial DNA [44,45]. However, Y-STR markers are lacking in this semiconductor sequencing platform. If a Y-STR marker system compatible with the existing platform is developed, it would be a cost-efficient tool that laboratories can use to detect a greater variety of genetic markers without changes in the instrumentation required.

In this study, we developed a new panel ('GrandFiler Y-STR Panel') that allows co-amplification and detection of 102 Y-STRs and Amelogenin on the Precision ID NGS System. Specifically, 51 Y-STRs were used in commercial CE kits to facilitate compatible data sharing. Another 51 Y-STRs with relatively high gene diversity were selected using commercial NGS kits and the Short Tandem Repeat DNA Internet DataBase (STRBase; https://strbase.nist.gov/) to improve the discrimination capacity as a haplotype to minimize adventitious haplotype matches. Additionally, we validated this panel with respect to three significant aspects: 1) run performance (run metrics, sample-specific metrics, and coverage analyses), 2) NGS performance (sample-to-chip arrangement, depth of coverage, sequence coverage ratio, stutter ratio, and inter-locus balance), and 3) panel characteristics (species specificity, sensitivity, repeatability, concordance, degraded samples, case-type samples, and mixtures). By developing and validating the GrandFiler Y-STR Panel on the Precision ID NGS System, we aimed to answer the following questions: (1) Is the Ion AmpliSeq Technology applicable to multiplex Y-STRs? (2) Can the performance of this newly developed Y-STR panel meet the needs of forensic scientists in real-world scenarios? (3) Can the Y-STRs in this study and A-STRs in a commercial panel be jointly detected?

2. Materials and methods

2.1. Sampling, and DNA extraction and quantification

All samples were extracted, quantified, and stored as described previously [52]. Informed consent was obtained from all the participants. Briefly, samples were extracted using the Automatic 96 Channel Micro DNA Extraction Workstation (Bokun Biotech, Changchun, China) or the AutoMate Express ™ Forensic DNA Extraction System with the PrepFiler Express BTA ™ Forensic DNA Extraction Kit (Thermo Fisher Scientific, MA, USA), according to manufacturer recommendations [53]. The Applied Biosystems ® QuantStudio 5 Real-time PCR System and the Quantifiler ® Trio DNA Quantification Kit (Thermo Fisher Scientific) were used for human genomic DNA quantification according to manufacturer recommendations [54]. Concomitantly, the Qubit ™ 3.0 fluorometer and Qubit ™ dsDNA HS Assay Kit (Thermo Fisher Scientific) were used for nonhuman genomic DNA according to manufacturer recommendations [55]. All DNA was stored at – 20 °C.

Samples involved in the developmental validation studies were prepared as follows. The details of the sample-to-chip arrangement in this study can be found in Table S1.

2.1.1. NGS performance, reproducibility, and concordance

Eight standard samples were used, including the Standard Reference Material ® (SRM) 2395c Components A, B, C, D, and E (NIST, MD, USA), DNA Control 007 (Thermo Fisher Scientific), 2800 M Control DNA (Promega, WI, USA), and 9948 Male DNA (Promega). All standards were sequenced in duplicate with 1 ng of input DNA on two separate Ion 530 ™ chips along with other samples.

2.1.2. Species specificity

Eight swab samples (chimpanzee, lemur, cat, dog, cow, pig, sheep, and chicken) were donated by the Shenyang Forest Zoo (Shenyang, China). Non-human extracts were sequenced once with 1 ng of input DNA on one Ion 530 ™ chip, along with other samples. Additionally, eight human female swab samples were sequenced once with 1 ng of input DNA on one Ion 530 ™ Chip along with eight human male samples to evaluate human male specificity.

2.1.3. Sensitivity

Serial dilutions of 1, 0.5, 0.2, 0.1, and 0.05 ng were prepared with the quantified male DNA Control 007 and sequenced in duplicate on two separate Ion 530 ™ chips along with other samples.

2.1.4. Degraded samples

Four samples with 0.71, 4.52, 63.84, and 146.01 of degradation index (DI) values were adopted, as described previously [52]. DNA extracts were sequenced with 1 ng of input DNA for samples with a concentration > 0.067 ng/μL; otherwise, 15 μL of the extract was added to the target amplification.

2.1.5. Case-type samples

Seven samples were collected from routine casework, including muscle, buccal swab, hair, fingernail, old bloodstain, fresh bloodstain, and semen/female vaginal secretion mixture samples. The amount of input DNA was 1 ng for the case-type samples, except for the fingernail sample (0.045 ng).

2.1.6. Mixture

Male-male mixtures were prepared with the male DNA Control 007 and the male 2800 M Control DNA at different ratios (19:1, 9:1, 4:1, 1:1, 1:4, 1:9, and 1:19), while holding the total amount of input DNA constantly mixed at 1 ng to the target amplification [52]. In turn, male-female mixtures were prepared with the male DNA Control 007 and the female K562 Genomic DNA (Promega) at ratios of 1:1, 1:10, 1:100, and 1:1000, while holding the male input DNA at 0.1 ng. All mixtures were sequenced in duplicate on two Ion 530 ™ chips along with other samples.

In addition, *sensitivity* and *NGS performance, reproducibility, and concordance* were validated using the AY-STR NGS workflow.

2.2. Marker selection and panel design

The selection of markers for the GrandFiler Y-STR Panel was as follows: (1) common Y-STR loci incorporated in commercial CE kits, (2) additional Y-STR loci used in commercial NGS kits, and (3) available Y-STR loci with relatively high gene diversity included in the STRBase.

A core set was selected to cover all Y-STRs (except for DYS447 and DYS526II) in commercial CE kits [11–29], including DYF387S1a/b, DYF404S1a/b, DYS19, DYS385a/b, DYS388, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS443, DYS444, DYS446, DYS448, DYS449, DYS456, DYS458, DYS459a/b, DYS460, DYS481, DYS508, DYS510, DYS518, DYS520, DYS522, DYS526I, DYS527a/b, DYS531, DYS533, DYS549, DYS552, DYS557, DYS570, DYS576, DYS587, DYS593, DYS596, DYS617, DYS622, DYS626, DYS627, DYS630, DYS635, DYS643, DYS645, DYS713, Y-GATA-A10, and Y-GATA-H4. An expanded set was adopted to have the maximum compatibility with Y-STRs in commercial NGS kits [35,38–40] and STRBase. This expanded set included DYS426, DYS434, DYS435, DYS436, DYS441, DYS442, DYS445, DYS450, DYS452, DYS453, DYS454, DYS455, DYS461, DYS462, DYS472, DYS476, DYS490, DYS492, DYS495, DYS497, DYS502, DYS505, DYS511, DYS512, DYS513, DYS525, DYS530, DYS532, DYS534, DYS538, DYS540, DYS541, DYS556, DYS568, DYS572, DYS575, DYS578, DYS585, DYS590, DYS594, DYS607, DYS612, DYS613, DYS616, DYS632, DYS638, DYS640, DYS641, DYS644, DYS712, and DYS717. Approximately 100–150 bases on either side of the repeat motif were components of the reference sequences, which were obtained from Ensembl (https://asia.ensembl.org/Homo_sapiens/Info/Index) for the selected loci. The characteristics of each Y-STR marker are shown in Figure S1 and Table S2.

The multiplex (108 Y-STRs and Amelogenin) was designed using the AmpliSeq Custom Pipeline v7.62 in the Ion AmpliSeq Designer (https://ampliseq.com/; Thermo Fisher Scientific). Some default parameters were selected: 'DNA Gene designs (multi-pool)' for Application Type, 'cfDNA (140 bp)' to 'Standard DNA (375 bp)' for DNA Type, and 'Human (hg38)' for the genome to use. The CSV-formatted file containing the type (REGION), name, chromosome, start, and end of each target (repeat region) was submitted to AmpliSeq for the design with 5-bp exon padding. This pipeline filters the resulting primers, maintaining the optimal melting temperatures and 20%–80% GC of nucleotide compositions. Moreover, it avoids the presence of long homopolymers and known SNPs with > 5% minor allele frequency on primers and potential interactions between primers in a pool. In addition, the design considered the following specific aspects: (1) 100% overall coverage of each target, (2) shortest in silico amplicon size for each target, and (3) incorporation of all targets into one amplicon pool. Primers were modified and synthesized by the manufacturer using a proprietary method specifically for use in the AmpliSeq workflow.

2.3. Library construction

An automatic library construction method was used. Similarly, for Y-STR markers, the AmpliSeq workflow options for the Precision ID GlobalFiler NGS STR Panel v2 were used [56]. Briefly, libraries were constructed on the Ion Chef System (Thermo Fisher Scientific) using the Precision ID DL8 Kit (Thermo Fisher Scientific) and the GrandFiler Y-STR Panel: 150 μL of 2× GrandFiler Y-STR Panel was added to Positions A and B in the Precision ID DL8 Reagents cartridge, 15 μL of quantified DNA extracts was pipetted into wells A1 to H1 of the Precision ID DL8 IonCode ™ Barcode Adapters 1−32 for Chef DL8 in 96 Well PCR Plates, and the rest of all consumables and chemicals of the Precision ID DL8 Kit were loaded onto the Ion Chef System. The AmpliSeq workflow protocol included one pool of primers, 24 cycles of target amplification, and 4 min of annealing and extension. Comparisons between 8 and 16 samples containing sensitivity samples were performed to determine a more suitable sample-to-chip arrangement for the developmental stage. All samples described in Section 2.1 were validated with the GrandFiler Y-STR Panel, including NGS performance, reproducibility, concordance, species specificity, sensitivity, degraded and case-type samples, and mixture studies. The details of these samples, arranged in chips 1–7, are listed in Table S1.

As for the AY-STR markers, the performance of three combinations at different concentration ratios (1:1, 2:1, and 3:1) of the GrandFiler Y-STR Panel and Precision ID GlobalFiler ™ NGS STR Panel v2 (Thermo Fisher Scientific) were compared to determine the optimal experimental parameters at the initial stage. Briefly, 150 μL of the Y-STR panel was constantly added in Position A in the Precision ID DL8 Reagents cartridge, but 150 μL of the A-STR panel, 75 μL of the A-STR panel + 75 μL of H_2O, and 50 μL of the A-STR panel + 100 μL of H_2O were added in Position B, respectively. The AmpliSeq workflow protocol included two pools of primers, 24 cycles of target amplification, and 4 min of annealing and extension. Sample sensitivity, NGS performance, reproducibility, and concordance were validated for the AY-STR NGS workflow. The details of these samples, arranged in chips 8–13, are listed in Table S1.

Each combined library generated from the Ion Chef System was quantified on the Applied Biosystems QuantStudio 5 Real-time PCR System using the Ion Library TaqMan ® Quantitation Kit (Thermo Fisher Scientific) according to the user guide of the Precision ID GlobalFiler NGS STR Panel v2 [56].

2.4. Templating and sequencing

Templating and sequencing options for the Precision ID GlobalFiler NGS STR Panel v2 were mainly referenced [56], with some modifications for the GrandFiler Y-STR Panel. Templating and chip-loading procedures were automatically performed on the Ion Chef System using an Ion S5 ™ Precision ID Chef & Sequencing Kit (Thermo Fisher Scientific). Briefly, each quantified combined library was normalized to 50 pM and pooled according to the arrangements listed in Table S1. Then, 25 μL of the super-pooled libraries was added to Position (s) A (and B) of the Ion S5 ™ Precision ID Chef Reagents cartridge. The consumables and chemicals of the Ion S5 ™ Precision ID Chef & Sequencing Kit and Ion 530 ™ chip(s) were loaded onto the Ion Chef System. The default templating protocol for the Precision ID Chef

Reagents was used in this study. Subsequently, the loaded chip was sequenced on the initialized Ion GeneStudio S5 ™ Plus System (Thermo Fisher Scientific) using the Ion S5 ™ Precision ID Chef & Sequencing Kit. For this study, three parameters were modified: GrandFiler_Y-STR _Panel_Targets.bed (or GrandFiler_AY-STR_Panel_Targets.bed) for 'Target Regions,' GRCh38.fasta for 'Reference Library,' and 852 for 'Sequence Flows.' As mentioned in previous studies [46,52], the Ion samba HID2 flow order was used to improve end-to-end STR sequencing performance.

2.5. Data processing

Raw data were processed using Torrent Suite ™ Software version 5.12 (Thermo Fisher Scientific). The aligned BAM/BAI files were automatically transferred and genotyped using ForeNGS ™ Analysis Software (FNAS; Yuhua BioTech, Beijing, China) [52] at the default analysis settings (Y-STR markers and AY-STR markers in Table S3). In most instances, the fixed and floating analytical threshold (AT) and interpretation threshold (IT) were set to 100 × AT and 200 × IT in the minimum read depth and 2.0% AT and 5.0% IT in the maximum read depth, respectively; the maximum stutter ratio was less than 20% and 15% for Y-STRs and A-STRs, respectively. Further, the minimum allele coverage ratio was greater than 0.5 and 0.6 for multi-locus Y-STRs and for A-STRs, respectively. The STR nomenclature referenced the International Society for Forensic Genetics (ISFG) considerations [57,58], and the flanking-region SNP nomenclature was in accordance with the Database of Single Nucleotide Polymorphism (dbSNP).

2.6. CE typing

Degraded samples were examined using the Yfiler ™ Platinum Casework PCR Amplification Kit (Thermo Fisher Scientific) according to manufacturer recommendations [24]. Samples with concentrations higher than 0.067 ng/μL were amplified using 1 ng of input DNA with 29 thermal cycles in a volume of 25 μL on the ProFlex ™ PCR System (Thermo Fisher Scientific); otherwise, 15 μL extracts were used. Amplified products were separated and detected on the Applied Biosystems ® 3500XL Genetic Analyzer (Thermo Fisher Scientific). The.hid files were analyzed with GeneMapper ® ID-X Software v1.6 (Thermo Fisher Scientific) with a relative fluorescent unit (RFU) analytical threshold of 150.

2.7. Statistics

The depth of coverage (DoC) was defined as the sum total of all reads within the locus from which the reads were extracted from the NGS Genotype reports generated by FNAS. The sequence coverage ratio (SCR), including % allele, % stutter, and % noise, was calculated by dividing the reads for the true allele (also known as typed allele or effective reads), (N – 1) stutter, and noise by DoC for STR markers. The stutter ratio was estimated as the ratio of stutter to the true allele read. The allele coverage ratio (ACR), also known as heterozygote balance or intra-locus balance, was measured as the ratio of lower to higher allele coverage. Inter-locus balance was assessed as the proportion of loci with a DoC exceeding 20% of the average DoC across all loci.

The chi-squared test, Fisher's exact test, and one-way analysis of variance (ANOVA) were computed using R software version 4.0.5 [59], and figures were generated by the Package 'ggplot2' for R.

3. Results and discussion

3.1. GrandFiler Y-STR Panel

A single set of primers can produce two products at DYF387S1a/b, DYF404S1a/b, DYS385a/b, DYS459a/b, and DYS527a/b because these multilocus markers are located on palindromes or inverted repeats of the Y-chromosome (Figure S1). A similar trend was observed for Amelogenin X/Y. In addition, two products were generated at DYS389I/II and DYS526I/II because one of the primers bound to the flanking region of the two different repeat regions. Furthermore, DYS460 and upstream DYS461 were captured within a single amplicon. Thus, 102 pairs of primers had to be designed for 108 Y-STRs and Amelogenin; however, DYS526II was not considered because its amplicon was too long to be detected using NGS.

The panel design, which included 'separating' and 'merging' procedures, was a major challenge during the course of our study. First, the 102-plex was submitted for designs using 'cfDNA (140 bp),' 'FFPE (175 bp),' 'Standard DNA (275 bp),' and 'Standard DNA (375 bp)' for DNA Type. The outcome showed 43 pairs of primers designed by 'cfDNA (140 bp)' with 100% overall coverage, the shortest in-silico amplicon size, and each target incorporated into 1 pool, 33 pairs designed by 'FFPE (175 bp),' 24 pairs designed by 'Standard DNA (275 bp),' and 2 pairs by 'Standard DNA (375 bp).' Among them, a small number of primers contained repeat regions or had excessively long flanking regions, which were redesigned by changing the start and end positions of the targets.

Second, all successfully designed targets above were copied and incorporated into a new design using the 'merge' function in the Ion AmpliSeq Designer. The new design was again submitted using 'Standard DNA (375 bp)' for DNA Type and the 5-bp exon padding. Similarly, unsuccessful primers were redesigned and remerged until 102 amplicon targets were reached with 100% overall coverage and incorporation of all targets into one pool. Table S2 shows that the final design included 108 Y-STRs and Amelogenin with an average amplicon size of 133 bp, from a minimum size of 72 bp (DYS508 and DYS538) to a maximum size of 322 bp (DYS461 and DYS460 combined). The GrandFiler Y-STR Panel consists of 75 simple, 26 compound, and 7 complex repeat motif Y-STR markers having 16 trimer, 78 tetramer, 12 pentamer, and 2 hexamer repeat units.

Finally, we could not adjust the concentration of primers and customize the primer placement positions without the help of the White Glove Service from Thermo Fisher Scientific. Therefore, the pre-wetlab procedure was conducted using one tube of primer pools with an equimolar concentration of 100 nM. Primers that failed target amplification or sequencing were redesigned, remerged, and resynthesized until the full profile of DNA Control 007 was obtained (Figure S2). However, DYF387S1a/b, DYS449, DYS518, DYS627, and DYS713 were not analyzed in subsequent studies because of their stability or other technical/quality-related issues.

3.2. Run summary

The run metrics from chips 2 and 3, and the sample-specific metrics and coverage analyses of male samples with 1 ng of input DNA from chips 2–5 were evaluated. Details of the run metrics, sample-specific metrics, and coverage analyses for each chip are listed in Table S1.

As for run metrics, the mean ± standard deviation (SD) of total bases, Q20 bases, percentages of bases with a quality score of 20 or higher (%≥ Q20), and total reads were calculated as (810.50 ± 86.97) M, (749.50 ± 81.32) M, (92.47 ± 0.11) %, and (6.14 ± 0.54) M, respectively. Additionally, the mean ± SD of percentages of ISP loading, usable reads, polyclonal, test fragment, adapter dimer, and low quality were estimated as (60.50 ± 3.53) %, (28.00 ± 1.41) %, (35.50 ± 2.12) %, (1.00 ± 0.00) %, (1.00 ± 0.00) %, and (54.40 ± 3.53) %, respectively. As this is the first report on detecting a massive number of Y-STR markers using the Precision ID NGS System, neither articles nor user guides were available; hence, only the recommended values in the user guide of the Precision ID GlobalFiler NGS STR Panel v2 [56] were referenced in this study. Results showed that ISP loading, polyclonal, test fragment, and adapter dimer percentages were all within the recommended ranges (>50%, 25%–45%, 1%, and ≤ 1%, respectively). However, the percentage of usable reads was lower than 30% and that of low-quality

reads was higher than 50%. With the limited information on the results generated in our laboratory and public documents [45,56,60,61], the percentage of low-quality reads increased from 12% to 54% in ascending order, whereas that of useable reads decreased from 61% to 28% in descending order when sequencing SNPs (identity-informative SNPs, ancestry-informative SNPs, and Y-SNPs), whole mitochondrial genomes, A-STRs, combinations of A-STRs and SNPs, and Y-STRs (details found in Table S4). Unfortunately, comprehensive studies of run metrics for sequencing different categories of genetic markers are rare, and we could not determine the actual reason.

Furthermore, sample-specific metrics and coverage analyses were also computed in order to present a general performance of each sample with sufficient input DNA, where the mean ± SD of the reads, mean read length, on-target reads, mean depth, and uniformity were (401262.00 ± 93086.81) ×, (132.64 ± 8.34) bp, (80.03 ± 4.06) %, (2250.60 ± 540.63) ×, and (77.75 ± 4.89) %, respectively.

3.3. NGS performance

Lastly, 10 samples from sensitivity studies were used to determine the sample-to-chip arrangement, and 16 samples from the reproducibility and concordance studies were used to evaluate the performance of the GrandFiler Y-STR Panel. The details of the NGS performance metrics in this study and the default analysis settings in FNAS are listed in Table S3.

3.3.1. Sample-to-chip arrangement

Five sensitivity samples (1, 0.5, 0.2, 0.1, and 0.05 ng, respectively) and three or eleven other samples (all at 1 ng) were sequenced separately in two chips in this study. When sequencing eight samples on an

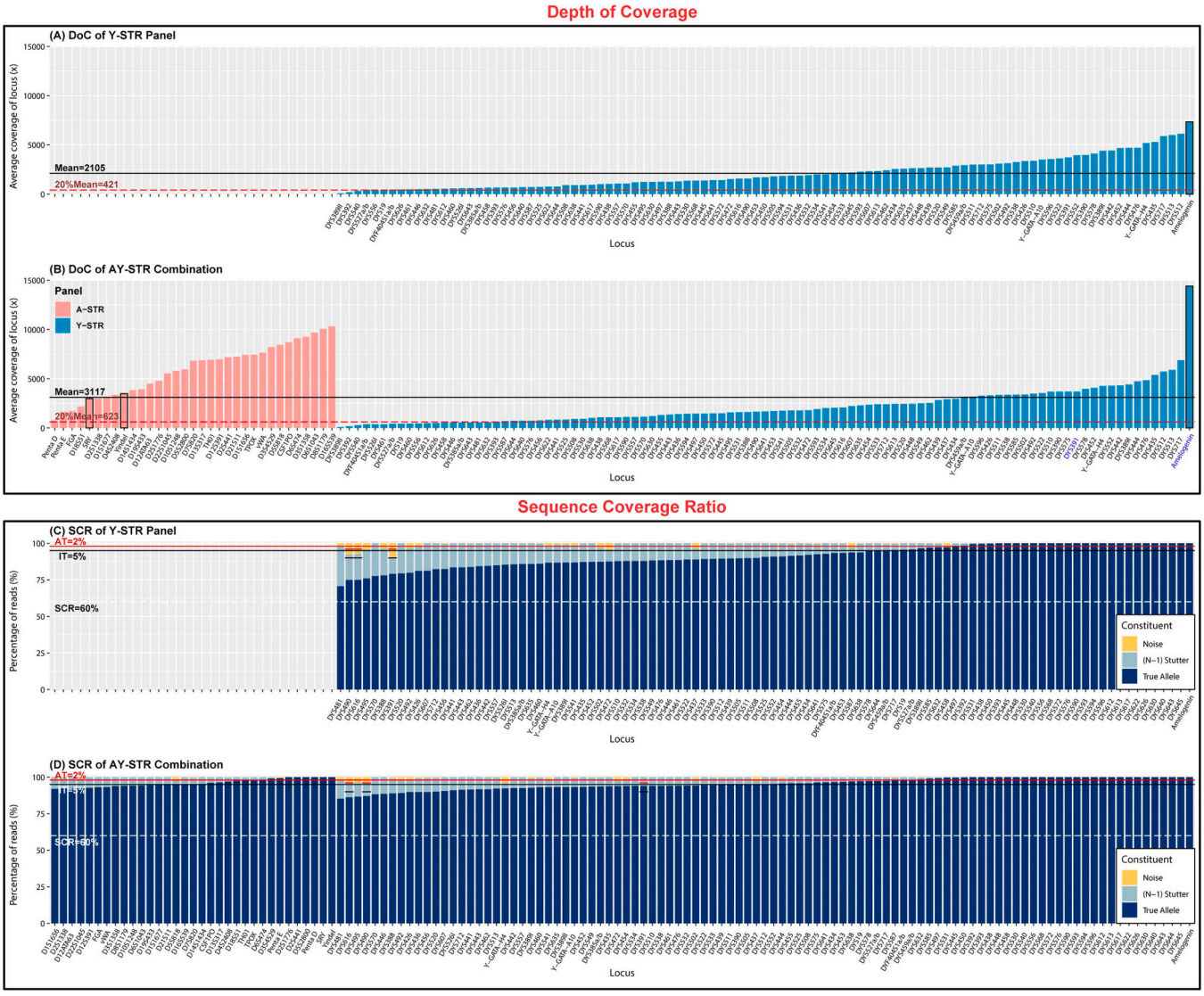


Fig. 1. NGS performance. (A) The depth of coverage (DoC) of the GrandFiler Y-STR Panel shows the average coverage at a locus from lowest to highest. The horizontal black solid line indicates the mean DoC = 2105 × across 102 Y-STRs and Amelogenin, and the horizontal red dash line indicates 20% of the mean DoC = 421 ×. (B) DoC of the AY-STR NGS workflow shows the average coverage distribution for the Precision ID GlobalFiler ™ NGS STR Panel v2 and the GrandFiler Y-STR Panel separately. The black line indicates the mean DoC across 136 loci = 3117 ×, and the red line indicates 20% of the mean DoC = 623 ×. The bar graph with a black outline in (A) and (B) demonstrates that the locus belongs to a sex-determining marker. (C) The sequence coverage ratio (SCR) of the GrandFiler Y-STR Panel displays % true allele, % (N – 1) stutter, and % noise at each locus, where loci are arranged with % true allele from the lowest to the highest. (D) SCR in the AY-STR NGS workflow displays the % true allele distribution for A-STR and Y-STR panels, respectively. In (C) and (D), horizontal red or black solid lines indicate the recommended analytical threshold (2%) or interpretation threshold (5%) in the ForeNGS Analysis Software, respectively. Occasionally, 4% / 10% are given at DYS391, DYS490, and DYS616. Horizontal white dash lines indicate SCR = 60%. In the AY-STR NGS workflow, Amelogenin and DYS391 are shared in two panels but assigned to the Y-STR panel for convenience. Details of DoC and SCR are listed in Table S3.

Ion 530 ™ Chip, DoCs of the sensitivity samples were averaged as 3635 ×, 3189 ×, 2605 ×, 2155 ×, and 1368 ×, respectively, and full profiles were generated from FNAS for all samples. When 16 samples were on a chip, the average DoCs were 1662 ×, 1476 ×, 1226 ×, 1007 ×, and 645 ×, respectively, and full profiles could be obtained for all, except for the 0.05 ng samples that showed DoCs of two Y-STRs below the fixed analytical threshold (100 ×). With respect to balancing cost and efficiency, we consider multiplexing 16 samples per chip as the optimal sample-to-chip arrangement.

3.3.2. Depth of coverage

The mean ± SD of DoC was calculated as (2105.39 ± 1561.28) × for the GrandFiler Y-STR Panel, ranging from the lowest observed at DYS389II (74.44 ×) to the highest at Amelogenin (7345.64 ×), as shown in Fig. 1A and Table S3. DYS389II was the only locus with an average DoC of < 100 × in this panel. This underperformance might be associated with primer concentration, allele amplicon spread, and/or read length. First, one pair of primers was shared by DYS389II and DYS389I, as demonstrated in Section 3.1, where the forward primer bound to the same upstream flanking region but the reverse primer bound to two different downstream flanking regions. Thus, the theoretical concentration was halved. Second, the large spread (121 bp) between a longer amplicon of DYS389II (211 bp) and a shorter amplicon of DYS389I (90 bp) easily caused preferential amplification using the same pair of primers, similar to the heterozygote imbalance of A-STR markers. Most importantly, the large difference between the number of reads for DYS389I and DYS389II was probably due to the short read length, which resulted in fewer reads for interpretation of the second repeat region of DYS398II, whereas the short read length also increased the number of reads for DYS389I. This DYS389I/II pattern is commonly observed using NGS technology [52,62].

3.3.3. Sequence coverage ratio

As shown in Fig. 1C and Table S3, the percentage of true alleles averaged 91.08%, from the lowest observed at DYS481 (70.62%) to the highest observed at 22 Y-STRs (100%). The average % (N – 1) stutter was 8.25%, with the lowest at DYS458 and 22 Y-STRs above-mentioned, and the highest at DYS481 (27.46%). The 100% true alleles and 0% (N – 1) stutters at 22 Y-STRs do not necessarily reflect an association with the natural structures of amplicons, which may be due to the strict filter settings of FNAS. However, vulnerable % true allele and % (N – 1) stutter at DYS481 have been reported previously [36,40,63], which is likely due to the simple trimer-repeat unit (CTT[*n*]) of this locus. Based on the limited data set used in this study, we observed that an extremely high % (N – 1) stutter would appear in the scenario of allele > 27 at DYS481, especially in a low-quality sample. The average % noise was calculated at 0.67%, from the highest at DYS391 (9.50%) to the lowest at 66 Y-STRs (0.00%), where 15 Y-STRs presented >1% noise. It should be noticed that, in this study, noise was defined as all sequences having the same length as true alleles and (N – 1) stutters but differing from them at least one base, and all those having a different length from true alleles or (N – 1) stutters, such as the reads from 1 to 5 nucleotide indels (x.1∼x.5 sequences) and other (N ± k) stutters, which would raise the percentage of noise in SCR. Overall, it was easy to distinguish true allele sequences from stutter and noise sequences when the percentage of true alleles was greater than 60%.

3.3.4. Stutter ratio

The (N – 1) stutter ratios from the Y-STRs within the GrandFiler Y-STR Panel are listed in Table S3. The average value was 8.15%. All (N – 1) stutter ratios were less than 20%, except for DYS495 (20.03%), DYS520 (20.63%), DYS570 (22.61%), and DYS481 (27.49%), which fell below the recommended stutter filter in the FNAS.

3.3.5. Inter-locus balance

The value equal to 20% of the average DoC was calculated as 421 × of the GrandFiler Y-STR Panel, where 100 loci (97.09%) exceeded this value; however, DYS389II, DYS392, and DYS540 fell below this value (Fig. 1A and Table S3). Thus, we concluded that the inter-locus balance of the GrandFiler Y-STR Panel is within an acceptable range (≥70%).

3.4. Developmental validation

Species specificity, sensitivity, reproducibility, concordance, degraded samples, case-type samples, and mixtures were studied to validate the GrandFiler Y-STR Panel according to the SWGDAM Validation Guidelines for DNA Analysis Methods (2016) [64] and the Guidance Document for the FBI Quality Assurance Standards for Forensic DNA Testing and DNA Databasing Laboratories (2023) [65].

3.4.1. Species specificity

Although the GrandFiler Y-STR Panel was designed for human male samples, its ability to detect genetic information from non-targeted female and non-human samples should be determined. Therefore, eight non-human samples, namely, chimpanzee, lemur, cat, dog, cow, pig, sheep, and chicken, and 10 human samples including, male DNA Control 007, female K562 Genomic DNA, and eight female extracted DNA, were used to evaluate species specificity (Figure S3). In the run summary from the Torrent Suite Software, the length distribution was mainly concentrated within 100 – 150 bp and 200 – 250 bp from the human male sample, whereas those for the human female and non-human samples were clearly differentiated. When alleles were called by FNAS, 44.44% (181163 out of 407644) of the reads were assigned to this panel for the human male, and a full profile could be obtained. Conversely, in the female samples, the average percentage of reads aligned to this panel was relatively low (8.95% ± 2.58%) compared to the male sample. Only one allele (AMELX) was detected and accounted for the majority of reads. Some loci with too few reads were detected but filtered out because they did not meet the AT requirements. Cross-reactivity can be population-based as well as subgroup-based. However, this would not have been detected in such a small size of samples. In non-human samples, the percentage of reads aligned was meager (< 1.3%), and no allele was called, except for AMELX from the female chimpanzee. Unfortunately, a male chimpanzee sample was unavailable for this study; therefore, we could not fully demonstrate the cross-reactivity of this panel with that species. Species specificity was studied in the Ion AmpliSeq HID Y-SNP Research Panel by Ralf *et al.* [48], and a similar pattern was observed in this study: all non-human samples with a high number of sample-specific reads in a run summary (> 50000), but a low percentage of reads aligned to the GrandFiler Y-STR Panel (< 2%) and a small number of alleles called (< 2). Based on our results, human female and non-human samples were effectively distinguished from human male samples in most cases.

3.4.2. Sensitivity

Sensitivity is defined as the ability to reproducibly generate full profiles from a range of input DNA, especially the lower limit of the assay, with DoCs above the analytical threshold on the NGS instrument using the specified conditions. Sensitivity studies were performed in duplicate using serial dilutions. Figure S4 shows that 100.00% profiles were obtained from samples with input DNA > 0.1 ng, and an average 95.15% of the profile was generated from 0.05 ng samples with two dropouts (DYS570 and DYS644) and eight dropouts (DYS389II, DYS446, DYS458, DYS508, DYS540, DYS556, DYS612, and DYS638) observed from each replicate. Overall, as expected, the average DoC for serial dilutions in duplicate decreased linearly as the amount of input DNA decreased: from 1620.48 × at 1 ng input DNA to 481.44 × at 0.05 ng. In other words, as the amount of input DNA decreased, the possibility of allele/locus dropout increased. The sensitivity of the GrandFiler Y-STR Panel is similar to that previously described in validation studies of CE Y-STR kits [11–29] and other NGS kits [36,38,52], but it can acquire a higher resolution of the haplotype in the Y-chromosome than other

panels from a sample once a test. In order to attain a balance between the generation of a full profile and performance stability, we recommend an input DNA > 0.1 ng as the optimal amount to use when 16 samples are sequenced on a chip.

3.4.3. Reproducibility and concordance

Reproducibility was evaluated using reliable and accurate sequence-based (SB) and length-based (LB) alleles obtained from the same operator and/or instrument. Specifically, eight standard control samples were adapted with 16 different barcodes and sequenced on chips 2 and 3. The results showed that all LB and SB alleles generated by FNAS were reproducible and accurate.

As alleles at many loci from these samples were unreported, the concordance was assessed to the utmost extent by allele comparisons between NGS and CE kits and among different NGS kits. In Table S5, a total of 359 alleles were compared: 225 LB alleles from DNA Control 007, 2800 M Control DNA, and 9948 Male DNA at 73 Y-STRs and Amelogenin were compared with those reported by Fan *et al.* [40], and 134 LB alleles and 127 SB alleles from SRM 2395c Components A, B, C, D, and E at 25 Y-STRs and Amelogenin were compared with those listed in STRBase. According to the ISFG guidelines [57,58], SB alleles were changed to forward strands if they were named as reverse strands in the original documents.

Compared with the LB alleles in Fan *et al.* [40], two discrepancies were found at one Y-STR. For DYS572 ($[AAAT]_n$), one concordance was observed from 2800 M (allele 10), but two discordances were found from 007 (allele 11 in this study *vs.* 10 in Fan *et al.* [40]) and 9948 (11 *vs.*10). However, we obtained all concordant alleles from the five SRM 2395c components at DYS572. As amplicon strings were not available in Fan *et al.* [40], we could not determine the real reasons for these discrepancies. Another 15 discrepancies were found at 5 Y-STRs due to the different motif designations marked in bold, such as DYS552 ($[TCTA]_n$ TCTG $[TCTA]_n$ in this study *vs.* $[TCTA]_n$ TCTG $[TCTA]_n$ **N_{40} [TCTA] n** in Fan *et al.* [40]), DYS587 ($[CAATA]_n$ *vs.* $[ATACA]_n$ **[(GTACA) (ATACA)] 3**), DYS613 (**[ATG] 8 ATA** $[ATG]_n$ *vs.* $[ATG]_n$), DYS616 ($[TAT]_n$ **CAT [TAT] 3** *vs.* $[TAT]_n$), and DYS626 ($[GAAA]_n$ **N_{24}** $[GAAA]_3$ gagaca **[GAAA] 5** aaa **[GAAA] n** gaag **[GAAA] 3** *vs.* $[AAAG]_n$ $[AGAA]_2$ AGAG $[GAAG]_3$ $[AAAG]_3$). Fortunately, these 15 discrepancies were resolved by adjusting the configuration file in FNAS because they had sequences compatible with those in Fan *et al.* [40].

Compared with LB and SB alleles in the STRBase, the exceptionally high (N – 1) stutter percentage that almost approached the read depth of the parent allele was observed at DYS481 in SRM 2395c C and E, which was discussed in Section 3.3.3. However, we could still identify the true alleles 28 for both with caution. An actual discrepancy was observed at DYS527a/b in SRM 2395c Component E, where LB alleles (22,23) were the same, but one SB allele, namely, 22, was different ($[GAAA]_{16}$ $[GGAA]_6$ in this study *vs.* $[GAAA]_{15}$ $[GGAA]_7$ in STRBase; https://strbase-archive.nist.gov/srm2395.htm), and the other SB allele, i.e., 23, was absent from STRBase. The reason for this SB discrepancy is unknown. Except for this discrepancy, all LB and SB alleles were in agreement with those listed in the STRBase.

Further, the bioinformatic concordance was determined using STRait Razor v2s [66] and STRinNGS v2 [67]. The results showed that concordance among different software packages for both LB and SB alleles reached 100%. In summary, 99.64% (829/832) of alleles from the GrandFiler Y-STR Panel called by FNAS were concordant with the different CE, NGS, and/or bioinformatics interpretations, whereas minor discordance (0.36%) resulted from two LB alleles at DYS572 and one SB allele at DYS527a/b.

3.4.4. Degraded samples

Many environmental factors and delay in testing evidence samples can result in DNA degradation or damage at random locations. To determine the detection efficiency, samples soaked in tap water at room temperature were analyzed and categorized as intact (control sample), slightly to moderately degraded (Day 3 sample), and severely degraded (Day 6 and Day 9 samples). Table 1 shows the number of alleles and percentage of profiles obtained using the GrandFiler Y-STR Panel and Yfiler Platinum Casework PCR Amplification Kit, respectively. Control samples contained 104 and 41 alleles with two assays, respectively. In degraded samples, the difference in profile percentages between the two assays was not noticeable when detecting the sample that suffered slight to moderate degradation ($p = 0.1795$, Fisher's exact test), but was significant in severely degraded samples ($p = 0.0063$ for Day 6, and $p = 7.76 \times 10^{-5}$ for Day 9 samples, as per Fisher's exact test). Furthermore, when we combined the results of Y-STRs and A-STRs in a previous study [52], the number of alleles increased by 452.38%, from 21/76 alleles (27.63%) with CE to 116/156 alleles (74.36%) with NGS in the case of Day 9 samples. Fig. 2 illustrates the length distribution of fragments (amplicons + primers) against the samples obtained using the NGS and CE assays. As expected, the GrandFiler Y-STR Panel showed an overwhelming advantage in detecting degraded samples, which benefited from the Ion AmpliSeq Technology that can shorten amplicons and integrate more loci when designing the panel, as demonstrated in Section 3.1.

3.4.5. Case-type samples

The ability to obtain reliable genotypes should be evaluated using samples that are representative of those typically encountered in forensic laboratories, such as fresh bloodstains, old bloodstains, buccal swabs, semen/female vaginal secretion mixtures, and muscle, hair, and fingernail samples. Table S6 shows that full profiles were obtained using the GrandFiler Y-STR Panel from all case-type samples, except for the fingernail (83.65% profile). The fingernail was a challenging sample with low input DNA (0.045 ng), beyond the lower limits (0.1 ng) of both CE and NGS assays to obtain full profiles [13,14,24,36,42,46,52]. Although this study had limited data, the results indicate that the GrandFiler Y-STR Panel is suitable for detecting case-type samples in a real scenario.

3.4.6. Mixtures

Evidence samples containing body fluids and/or tissues originating from more than one individual are commonly encountered in forensic casework. Mixture studies may assist in determining the range of contributor ratios within which a full profile (or 100% unique alleles) can be obtained from a minor contributor with a known reference profile. In this study, seven male-male mixtures (19:1, 9:1, 4:1, 1:1, 1:4, 1:9, and 1:19) and four male-female mixtures (1:1, 1:10, 1:100, and 1:1000) were sequenced in duplicate to evaluate their ability to detect minor contributors using the GrandFiler Y-STR Panel. NGS detection provides not only more genetic information, but also more uncertainty. The fixed analytical threshold was set at 20× for interpretation at all loci.

For male-male mixtures, 60 unique SB alleles for the male DNA Control 007 and 59 unique SB alleles for the male 2800 M Control DNA are shown in Table S7. As the mixture ratio increased, the percentage of unique alleles identified from the minor contributors decreased. Table S8 and Figure S5 show that 100% of the unique alleles from the minor contributors could be obtained only at a 1:1 ratio. When moving to 4:1 and 1:4 ratios, an average of 98.31% unique alleles from the minor contributor 2800 M was identified along with dropouts at DYS389II or DYS527b, and an average of 90.00% unique alleles from the minor contributor DNA 007 was identified with dropouts at DYF404S1a/b, DYS19, DYS385a/b, DYS389II, DYS495, DYS497, DYS527a/b, DYS616, or DYS626. However, when moving forward to 19:1 and 1:19 ratios, only 76.27% and 59.17% unique alleles were observed from minor contributors, respectively. Although these unique allele dropouts were observed at loci that were underperformed as demonstrated in Section 3.3, had > 200 bp of amplicon length, or presented multi-copy alleles, the result of the two-sample mixture deconvolution was not as specific as expected.

As for male-female mixtures, only one SB allele (AMELX) from

Table 1
The number of alleles and the percentage of profiles recovered from degraded samples with the GrandFiler Y-STR Panel and the Yfiler ™ Platinum Casework PCR Amplification Kit.

Sample	Degradation Index	Concentration (ng/μL)	Input DNA (ng)	GrandFiler Y-STR Panel		Yfiler Platinum Casework PCR Amplification Kit	
				# of Allele	% profile	# of Allele	% profile
Control	0.71	7.82	1	104	100.00%	41	100.00%
Day 3	4.52	2.82	1	93	89.42%	40	97.56%
Day 6	63.84	0.14	1	83	79.81%	23	56.10%
Day 9	146.01	0.02	0.3	67	64.42%	11	26.83%

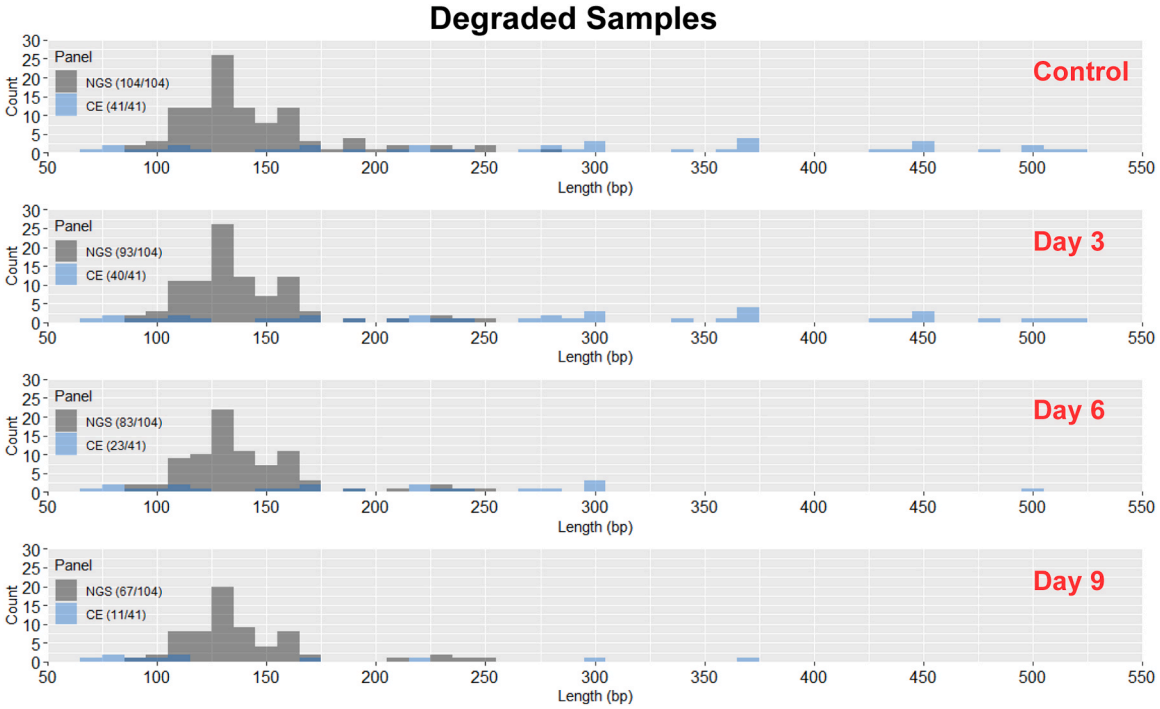


Fig. 2. Degraded samples. The frequency distribution graphs show true alleles are counted against the length of fragments (amplicons + primers) from the control sample and three time-period (Day 3, 6, and 9) degraded samples using NGS and CE assays, where the number of true alleles is annotated.

female K562 Genomic DNA was found using the GrandFiler Y-STR Panel, which indicated that the primers used were specific to the human Y-chromosome, as previously demonstrated in Section 3.4.1. However, the recovery performance in the lower limit amount of male DNA (0.1 ng) mixed with an excess of female DNA (up to 100 ng) was not significant. As shown in Table S9 and Figure S5, 100% of the profiles from the male DNA Control 007 were detected in duplicate only at a 1:1 ratio. When 0.1 ng of DNA 007 was mixed with 1 ng of K562, an average of 96.15% of the male profile was obtained, which was similar to the percentage obtained using 81 Y-STRs [40]. Furthermore, 83.17% of the male profile was obtained with 0.1 ng mixed with 10 ng (1:100), which was higher than that obtained using 859 Y-SNPs at a 1:50 ratio [48]. However, an abrupt decrease was observed at the 1:1000 ratio, in which case, only 22.60% of the male profile remained. The full male profile was recovered at a >1:1000 ratio using the Yfiler Plus PCR Amplification Kit and the Yfiler Platinum Casework PCR Amplification Kit [14,24]. The reason for the discouraging performance with NGS is alluded to in Ralf *et al.* [48]. In our study, all male-female mixtures also showed a high number of sample-specific reads (e.g., 346770 at a 1:1000 ratio) and a high percentage of off-target reads (e.g., 97.46%) in a run summary, but a low percentage of reads aligned to the Y-chromosome (e.g., 1.43%) in FNAS. This may be explained by the target amplification products being amplified again during the templating step (i.e., emulsion PCR), which is unavailable in the CE workflow.

3.5. Autosomal STR and Y-STR combination workflow

The GrandFiler Y-STR Panel and the Precision ID GlobalFiler NGS STR Panel v2 were combined into a single workflow to maximize the genetic information obtained from forensic samples. The completeness of the profiles and DoCs of the combinations (1:1, 2:1, and 3:1) were compared to determine the optimal experimental parameters for the AY-STR NGS workflow (Table S10). The results showed that 100% profiles could be obtained from eight standard control samples when the combination was mixed at a 2:1 concentration ratio, which was significantly ($p = 9.46 \times 10^{-5}$, Fisher's exact test) higher than that at a 1:1 ratio containing 34 dropouts at 1 A-STR and 33 Y-STRs. Although 5 dropouts from A-STRs were observed at a 3:1 ratio, the difference was not statistically significant ($p = 0.0595$, Fisher's exact test). This finding demonstrates that a decrease in the A-STR panel concentration benefits the AY-STR NGS workflow. Additionally, the DoC of the combinations was compared, and there was no statistically significant difference between the 2:1 and 1:1 ratios ($p = 0.8530$, one-way ANOVA); however, the DoC of the 2:1 ratio was significantly ($p = 0.0228$, one-way ANOVA) higher than that of the 3:1 ratio, indicating that the overall DoC was impeded when the A-STR panel concentration was below a specific limit. In summary, we believe that the completeness of genetic information and overall NGS performance can be balanced using a combination of Y-STR and A-STR panels at a 2:1 concentration ratio.

The run summary of the AY-STR NGS workflow was evaluated with

run metrics from chips 9 and 11–13, and sample-specific metrics and coverage analyses of males with 1 ng of input DNA from these four chips. Details of each chip and sample are listed in Table S1. First, the mean ± SD of run metrics (i.e., total bases, Q20 bases, %≥ Q20, and total reads) were calculated as (895.50 ± 43.28) M, (833.50 ± 43.56) M, (93.06 ± 0.39) %, and (6.47 ± 0.65) M, respectively. The mean ± SD of percentages of ISP loading, usable reads, polyclonal, test fragment, adapter dimer, and low quality were also calculated as (62.00 ± 2.58) %, (30.50 ± 3.00) %, (39.00 ± 0.82) %, (1.00 ± 0.00) %, (1.75 ± 0.96) %, and (48.00 ± 5.23) %, respectively. All results were within the recommended ranges, except for the adapter dimer. The higher percentage of adapter dimers was presumably due to the excessive number of samples with lower input DNA in chips 11 and 12. That is, the chips contained sensitivity samples as previously explained [41]. Notably, the percentage of usable reads and low quality of the AY-STR NGS workflow were better than those of the single Y-STR workflow. The combination workflow of Y-STRs with easily detected markers may be a way to enhance run performance. Second, sample-specific metrics and coverage analyses were also computed, where the mean ± SD of the reads, mean read length, on-target reads, mean depth, and uniformity were (869312.83 ± 147183.02) ×, (132.79 ± 3.73) bp, (83.26 ± 4.06) %, (3398.42 ± 561.44) ×, and (79.47 ± 3.09) %, respectively.

The AY-STR NGS workflow was validated for several aspects, including NGS performance, reproducibility, concordance, and sensitivity. Notably, two genetic markers (Amelogenin and DYS391) were shared in the two panels but were assigned to the Y-STR panel for convenience in this study, as shown in Figs. 2B and 2D and Tables S3, S5, and S11.

First, the average DoC was calculated as 3117.13 ×, 5912.25 ×, and 2185.43 × for all 136 markers from both panels, 33 from the A-STR panel, and 102 Y-STRs and Amelogenin from the Y-STR panel, respectively. The average DoC of the Y-STRs in the AY-STR NGS workflow was close to that in the single Y-STR workflow, and both workflows shared a similar DoC distribution trend, with the lowest DoC at the same 22 Y-STRs (Fig. 1B and Table S3). This demonstrated that the GrandFiler Y-STR Panel is stable when used alone or in combination with the Precision ID GlobalFiler NGS STR Panel. Surprisingly, the percentage of non-alleles, that is, % (N – 1) stutter and % noise, at Y-STRs in the AY-STR NGS workflow significantly ($p = 6.41 \times 10^{-7}$, one-way ANOVA) decreased compared to that in the single Y-STR workflow (Fig. 1D and Table S3). This pattern was observed in each sample, and there were no extreme outliers in the Y-STR panel alone or in the AY-STR NGS workflow. These findings suggest that co-sequencing Y-STRs with easily detected markers can enhance the run metrics abovementioned and benefit FNAS genotyping because more reads are assigned to the true alleles.

Second, Figure S6 shows an FNAS profile from DNA Control 007, and Table S5 shows that all LB and SB alleles from eight standard control samples were effectively obtained from replicates, indicating that the AY-STR NGS workflow is reproducible and accurate. In addition, all SB alleles at 102 Y-STRs and Amelogenin for eight samples from the AY-STR NGS workflow were concordant with those from the single Y-STR workflow described in Section 3.4.3. In addition, all SB alleles at 31 A-STRs, Yindel, and SRY for DNA 007, 2800 M, and 9948 achieved 100% concordance with those reported in a previous study [52]. These results demonstrate that combining the two panels did not produce adverse effects in terms of reproducibility or concordance.

Finally, as shown in Table S11, full profiles were obtained from samples with ≥ 0.1 ng of input DNA and an average of 92.70% from samples with input DNA at 0.05 ng. These results demonstrate that the AY-STR NGS workflow is as sensitive as the single Y-STR workflow and can help forensic scientists achieve the goal of generating more genetic information while consuming fewer DNA extracts.

In summary, the combination workflow of the GrandFiler Y-STR Panel and the Precision ID GlobalFiler NGS STR Panel v2 is as robust as that of the GrandFiler Y-STR Panel, and researchers can choose either workflow freely without any concern with regard to the consequences and variances caused by cross-reactions between primer sets in the two panels.

4. Conclusions

This article presents the GrandFiler Y-STR Panel, which is the first massive Y-chromosomal short tandem repeat (Y-STR) panel on the Precision ID NGS System targeting Amelogenin, 51 core Y-STRs, and 51 expanded Y-STRs in a single reaction with automatic data interpretation using ForeNGS Analysis Software. Our data allow us to answer the questions stated in the Introduction. First, the panel was successfully developed with the Ion AmpliSeq Technology by our 'separating' and 'merging' strategies, suitable for automatic targeted next-generation sequencing (NGS). The average amplicon length in the GrandFiler Y-STR Panel was 133 bp; even with primers, it was approximately 175 bp. This demonstrates the advantages of Ion AmpliSeq Technology. Second, the GrandFiler Y-STR Panel presents competent run metrics and NGS performance when 16 samples are sequenced in a chip, such as ≥ 60% bases with ≥ Q20, > 60% of effective reads, > 2000× of the depth of coverage (DoC), and ≥ 70% of the inter-locus balance. Furthermore, the preliminary developmental validation conducted herein revealed favorable results. Specifically, the system shows high specificity for human males, generating complete profiles at ≥ 0.1 ng input DNA, and recovering more genetic information from severely degraded samples and various case samples, which can meet forensic needs in real scenarios. Most importantly, we combined the GrandFiler Y-STR Panel and Precision ID GlobalFiler NGS STR Panel v2 into a single NGS workflow. Similarly, the AY-STR NGS workflow is reliable, robust, and sensitive whereby, 136 genetic markers can be simultaneously obtained from a sample without primer cross-reactions.

To the best of our knowledge, we achieved promising developments in an attempt to solve real cases using the GrandFiler Y-STR Panel. The two cases were obtained from previous population genetic studies [68]. In one case, one haplotype (HT00004) was shared by two individuals with the same surname but were excluded as close relatives. We could not tell the patrilineal difference between two samples using 26 sequence-based Y-STRs in the ForenSeq DNA Signature Prep Kit or 41 length-based Y-STRs in the Goldeneye ® DNA ID Y Plus Kit, but mismatches were observed at three Y-STRs using the GrandFiler Y-STR Panel. Whether having more Y-STR loci is better for differentiating unrelated individuals depends on the application. Amplicon size can affect the success of challenged samples to some degree. Under specific circumstances, a higher resolution of haplotypes may be helpful. In another case, we identified null alleles at DYS505 due to an Asian-specific large deletion (rs2051848397) stretching across its forward primer-binding region and the upstream locus DYS526I. In addition, simultaneous duplications and deletions of DYS448 and DYS626 located on the u2 segment between segments t1 and t2 (i.e., the P3 palindrome spacer) were identified using the GrandFiler Y-STR Panel. As shown in Figure S1, the dense loci in Yp11.32–Yq11.23 were incorporated into this panel, which is a valuable tool for investigating the mechanisms of duplications and/or deletions on the Y-chromosome, such as sequence-tagged site (STS) markers. In this regard, the GrandFiler Y-STR Panel has prospective applications in the field of male infertility [69].

However, the GrandFiler Y-STR Panel is not a finished product, as some loci (DYF387S1a/b, DYS449, DYS518, DYS627, and DYS713) proved to be vulnerable and were not analyzed in this study. Primer concentrations are among the most significant factors in target amplification that determine the overall performance of each amplicon, and repeated tests and primer titrations are commonly conducted to achieve optimal performance [70]. Although we tried to redesign the panel numerous times, an equimolar concentration of primers directly from the manufacturer was used without mixing to prevent potential contamination. Additionally, some expanded Y-STRs require more information regarding allelic polymorphisms and mutation rates in

Chinese populations. However, whether such expanded Y-STRs maximize haplotype diversity remains unclear.

In future experiments, we plan to adjust the concentration of each primer separately and investigate the population genetic information to further optimize NGS performance and locus selection of the GrandFiler Y-STR Panel. In addition, we will update the combination workflow for the GrandFiler Y-STR Panel and Ion AmpliSeq HID Y-SNP Research Panel v1 because the AY-STR NGS workflow is practical.

CRediT authorship contribution statement

Shaobo Yu: Writing – review & editing, Conceptualization. **Ze Liu:** Methodology. **Guangxin Jing:** Validation. **Fei Guo:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Yubo Lang:** Software, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank Yishu Zhou, Yu Liang, and Fangdong Lai at Thermo Fisher Scientific and Ziran Wang at Yuhua BioTech for their technical support.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fsigen.2024.103059.

References

[1] M. Kayser, Forensic use of Y-chromosome DNA: a general overview, Hum. Genet. 136 (2017) 621–635, https://doi.org/10.1007/s00439-017-1776-9.
[2] L. Roewer, J. Arnemann, N.K. Spurr, K.H. Grzeschik, J.T. Epplen, Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts, Hum. Genet. 89 (1992) 389–394, https://doi.org/10.1007/BF00194309.
[3] D.L. Daniels, A.M. Hall, J. Ballantyne, SWGDAM developmental validation of a 19-locus Y-STR system for forensic casework, J. Forensic Sci. 49 (2004) 668–683, https://doi.org/10.1520/JFS2003134.
[4] E.K. Hanson, P.N. Berdos, J. Ballantyne, Testing and evaluation of 43 "noncore" Y chromosome markers for forensic casework applications, J. Forensic Sci. 51 (2006) 1298–1314, https://doi.org/10.1111/j.1556-4029.2006.00263.x.
[5] E.K. Hanson, J. Ballantyne, Comprehensive annotated STR physical map of the human Y chromosome: Forensic implications, Leg. Med. (Tokyo) 8 (2006) 110–120, https://doi.org/10.1016/j.legalmed.2005.10.001.
[6] T. Willems, M. Gymrek, G.D. Poznik, C. Tyler-Smith, , 1000 Genomes Project Chromosome Y Group, Y. Erlich, Population-Scale Sequencing Data Enable Precise Estimates of Y-STR Mutation Rates, Am. J. Hum. Genet. 98 (2016) 919–933, https://doi.org/10.1016/j.ajhg.2016.04.001.
[7] M. Kayser, A. Caglià, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, F. Heidorn, S. Herrmann, B. Herzog, M. Hidding, K. Honda, M. Jobling, M. Krawczak, K. Leim, S. Meuser, E. Meyer, W. Oesterreich, A. Pandya, W. Parson, G. Penacino, A. Perez-Lezaun, A. Piccinini, M. Prinz, C. Schmitt, L. Roewer, Evaluation of Y-chromosomal STRs: a multicenter study, Int. J. Leg. Med. 110 (125–133) (1997) 141–149, https://doi.org/10.1007/s004140050051.
[8] V.L. Pascali, M. Dobosz, B. Brinkmann, Coordinating Y-chromosomal STR research for the Courts, Int. J. Leg. Med. 112 (1998) 1, https://doi.org/10.1007/s004140050188.
[9] J.M. Butler, Recent Developments in Y-Short Tandem Repeat and Y-Single Nucleotide Polymorphism Analysis, Forensic Sci. Rev. 15 (2003) 91–111.
[10] S.K. Sinha, B. Budowle, S.S. Arcot, S.L. Richey, R. Chakrabor, M.D. Jones, P. W. Wojtkiewicz, D.A. Schoenbauer, A.M. Gross, S.K. Sinha, J.G. Shewale, Development and validation of a multiplexed Y-chromosome STR genotyping system, Y-PLEX 6, for forensic casework, J. Forensic Sci. 48 (2003) 93–103.
[11] B.E. Krenke, L. Viculis, M.L. Richard, M. Prinz, S.C. Milne, C. Ladd, A.M. Gross, T. Gornall, J.R.H. Frappier, A.J. Eisenberg, C. Barna, X.G. Aranda, M. S. Adamowicz, B. Budowle, Validation of male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex, Forensic Sci. Int. 151 (2005) 111–124, https://doi.org/10.1016/j.forsciint.2005.02.008.
[12] J.J. Mulero, C.W. Chang, L.M. Calandro, R.L. Green, Y. Li, C.L. Johnson, L. K. Hennessy, Development and validation of the AmpFlSTR Yfiler PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system, J. Forensic Sci. 51 (2006) 64–75, https://doi.org/10.1111/j.1556-4029.2005.00016.x.
[13] J.M. Thompson, M.M. Ewing, W.E. Frank, J.J. Pogemiller, C.A. Nolde, D.J. Koehler, A.M. Shaffer, D.R. Rabbach, P.M. Fulmer, C.J. Sprecher, D.R. Storts, Developmental validation of the PowerPlex® Y23 System: a single multiplex Y-STR analysis system for casework and database samples, Forensic Sci. Int. Genet. 7 (2013) 240–250, https://doi.org/10.1016/j.fsigen.2012.10.013.
[14] S. Gopinath, C. Zhong, V. Nguyen, J. Ge, R.E. Lagacé, M.L. Short, J.J. Mulero, Developmental validation of the Yfiler(®) Plus PCR Amplification Kit: An enhanced Y-STR multiplex for casework and database applications, Forensic Sci. Int. Genet. 24 (2016) 164–175, https://doi.org/10.1016/j.fsigen.2016.07.006.
[15] S. Zhang, H. Tian, Z. Wang, S. Zhao, Z. Hu, C. Li, C. Ji, Development of a new 26plex Y-STRs typing system for forensic application, Forensic Sci. Int. Genet. 13 (2014) 112–120, https://doi.org/10.1016/j.fsigen.2014.06.015.
[16] S. Li, C. Liu, H. Liu, J. Ge, B. Budowle, C. Liu, W. Zheng, F. Li, B. Ge, Developmental validation of the EX20+4 system, Forensic Sci. Int. Genet. 11 (2014) 207–213, https://doi.org/10.1016/j.fsigen.2014.03.001.
[17] S. Zhang, Y. Bian, H. Tian, Z. Wang, Z. Hu, C. Li, Development and validation of a new STR 25-plex typing system, Forensic Sci. Int. Genet. 17 (2015) 61–69, https://doi.org/10.1016/j.fsigen.2015.03.008.
[18] W. Du, L. Chen, H. Liu, P. Qiu, F. Li, J. Gao, Y. Zhou, B. Wang, C. Liu, Developmental validation of the HomyGene19+14Y System, Int. J. Leg. Med. 131 (2017) 605–620, https://doi.org/10.1007/s00414-016-1505-2.
[19] S. Li, L. Chen, Y. Wang, Q. Xu, H. Liu, Y. Li, C. Liu, Developmental validation of a 6-dye STR kit with 27 loci, Int. J. Leg. Med. 132 (2018) 335–342, https://doi.org/10.1007/s00414-017-1586-6.
[20] X. Bai, Y. Yao, C. Wang, W. Li, Y. Wang, W. Song, X. Mo, W. Ma, J. Ye, Development of a new 25plex STRs typing system for forensic application, Electrophoresis 40 (2019) 1662–1676, https://doi.org/10.1002/elps.201900021.
[21] X.T. Mo, J. Zhang, W.H. Ma, X. Bai, W.S. Li, X.C. Zhao, J. Ye, Developmental validation of the DNATyper™Y26 PCR amplification kit: An enhanced Y-STR multiplex for familial searching, Forensic Sci. Int. Genet. 38 (2019) 113–120, https://doi.org/10.1016/j.fsigen.2018.10.008.
[22] H. Meng, Y. Guo, X. Jin, C. Chen, W. Cui, J. Shi, X. Wang, R. Liu, B. Zhu, Internal validation study of a newly developed 24-plex Y-STRs genotyping system for forensic application, Int. J. Leg. Med. 133 (2019) 733–743, https://doi.org/10.1007/s00414-019-02028-x.
[23] W. Du, P. Feng, H. Huang, W. Wu, L. Zhang, Y. Guo, C. Liu, H. Liu, C. Liu, L. Chen, Technical note: developmental validation of a novel 6-dye typing system with 36 Y-STR loci, Int. J. Leg. Med. 133 (2019) 1015–1027, https://doi.org/10.1007/s00414-018-1864-y.
[24] S. Lee, P.K. McAnany, C.W. Chang, W. Norona, M.L. Short, J.J. Mulero, C. Zhong, Developmental validation of Applied Biosystems YFiler Platinum Casework PCR Amplification Kit, Sci. Rep. 13 (2023) 14539, https://doi.org/10.1038/s41598-023-41788-w.
[25] Y. Zhao, X. Yao, Y. Li, C. Chen, D. Guo, Z. Jiao, Z. Yang, Z. Yu, W. Ma, Q. Zhang, Developmental validation of the microreader™ Y Prime Plus ID system: an advanced Y-STR 38-plex system for forensic applications, Sci. Justice 61 (2021) 260–270, https://doi.org/10.1016/j.scijus.2021.01.003.
[26] G. Fan, L. Pan, P. Tang, Y. Zhou, M. Liu, X. Luo, Technical note: developmental validation of a novel 41-plex Y-STR system for the direct amplification of reference samples, Int. J. Leg. Med. 135 (2021) 409–419, https://doi.org/10.1007/s00414-020-02326-9.
[27] C. Xiong, C. Yang, W. Wu, Y. Zeng, T. Lin, L. Chen, H. Liu, C. Liu, W. Du, M. Wang, H. Sun, C. Liu, Development and validation of a multiplex typing system with 32 Y-STRs for forensic application, Forensic Sci. Int. 339 (2022) 111409, https://doi.org/10.1016/j.forsciint.2022.111409.
[28] Y. Zeng, L. Chen, M. Wang, C. Yang, H. Liu, C. Xiao, C. Liu, Y. Li, Q. Xu, W. Du, C. Liu, The Validation of a Single Multiplex Typing System With 45 Y-STR Markers for Familial Searching and Database Construction, Front. Genet. 13 (2022) 842004, https://doi.org/10.3389/fgene.2022.842004.
[29] M. Song, L. Jiang, Q. Lyu, J. Ying, Z. Wang, Y. Zhou, F. Song, H. Luo, X. Song, B. Ying, Developmental validation of the Microreader Group Y Direct ID System: A novel six-dye typing system with 54 Y-chromosomal loci for forensic application, Electrophoresis 43 (2022) 2023–2032, https://doi.org/10.1002/elps.202200109.
[30] X. Zhao, K. Ma, H. Li, Y. Cao, W. Liu, H. Zhou, Y. Ping, Multiplex Y-STRs analysis using the ion torrent personal genome machine (PGM), Forensic Sci. Int. Genet. 19 (2015) 192–196, https://doi.org/10.1016/j.fsigen.2015.06.012.
[31] D.H. Warshauer, J.D. Churchill, N. Novroski, J.L. King, B. Budowle, Novel Y-chromosome Short Tandem Repeat Variants Detected Through the Use of Massively Parallel Sequencing, Genom. Proteom. Bioinforma. 13 (2015) 250–257, https://doi.org/10.1016/j.gpb.2015.08.001.
[32] S.Y. Kwon, H.Y. Lee, E.H. Kim, E.Y. Lee, K. Shin, Investigation into the sequence structure of 23 Y chromosomal STR loci using massively parallel sequencing, Forensic Sci. Int. Genet. 25 (2016) 132–141, https://doi.org/10.1016/j.fsigen.2016.08.010.
[33] S. Claerhout, P. Verstraete, L. Warnez, S. Vanpaemel, M. Larmuseau, R. Decorte, CSYseq: The first Y-chromosome sequencing tool typing a large number of Y-SNPs and Y-STRs to unravel worldwide human population genetics, PLoS Genet 17 (2021) e1009758, https://doi.org/10.1371/journal.pgen.1009758.
[34] M.H. Moon, S.R. Hong, K. Shin, Sequence Variations of 31 Y-Chromosomal Short Tandem Repeats Analyzed by Massively Parallel Sequencing in Three U.S.

Population Groups and Korean Population, J. Korean Med. Sci. 37 (2022) e40, https://doi.org/10.3346/jkms.2022.37.e40.
[35] K.R. Yang, L. Miao, K.L. Kang, Y.S. Feng, A.Q. Ji, C. Zhang, L.L. Guo, Y. Gao, M. T. Wei, J. Ye, J. Wu, L. Wang, Sequence polymorphisms of forensic Y-STRs revealed by a 68-plex in-house massively parallel sequencing panel, Forensic Sci. Int. Genet. 59 (2022) 102727, https://doi.org/10.1016/j.fsigen.2022.102727.
[36] F. Guo, J. Yu, L. Zhang, J. Li, Massively parallel sequencing of forensic STRs and SNPs using the Illumina® ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™ Forensic Genomics System, Forensic Sci. Int. Genet. 31 (2017) 135–148, https://doi.org/10.1016/j.fsigen.2017.09.003.
[37] T.I. Huszar, M.A. Jobling, J.H. Wetton, A phylogenetic framework facilitates Y-STR variant discovery and classification via massively parallel sequencing, Forensic Sci. Int. Genet. 35 (2018) 97–106, https://doi.org/10.1016/j.fsigen.2018.03.012.
[38] R. Moura-Neto, J.L. King, I. Mello, V. Dias, B. Crysup, A.E. Woerner, B. Budowle, R. Silva, Evaluation of Promega PowerSeq™ Auto/Y systems prototype on an admixed sample of Rio de Janeiro, Brazil: Population data, sensitivity, stutter and mixture studies, Forensic Sci. Int. Genet. 53 (2021) 102516, https://doi.org/10.1016/j.fsigen.2021.102516.
[39] R. Li, X. Shen, H. Chen, D. Peng, R. Wu, H. Sun, Developmental validation of the MGIEasy Signature Identification Library Prep Kit, an all-in-one multiplex system for forensic applications, Int. J. Leg. Med. 135 (2021) 739–753, https://doi.org/10.1007/s00414-021-02507-0.
[40] H. Fan, L. Wang, C. Liu, X. Lu, X. Xu, K. Ru, P. Qiu, C. Liu, S. Wen, Development and validation of a novel 133-plex forensic STR panel (52 STRs and 81 Y-STRs) using single-end 400 bp massive parallel sequencing, Int. J. Leg. Med. 136 (2022) 447–464, https://doi.org/10.1007/s00414-021-02738-1.
[41] M. Eduardoff, C. Santos, M. de la Puente, T.E. Gross, M. Fondevila, C. Strobl, B. Sobrino, D. Ballard, P.M. Schneider, Á. 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, https://doi.org/10.1016/j.fsigen.2015.04.007.
[42] F. Guo, Y. Zhou, H. Song, J. Zhao, H. Shen, B. Zhao, F. Liu, X. Jiang, Next generation sequencing of SNPs using the HID-Ion AmpliSeq™ identity panel on the ion torrent PGM™ platform, Forensic Sci. Int. Genet. 25 (2016) 73–84, https://doi.org/10.1016/j.fsigen.2016.07.021.
[43] G. Espregueira Themudo, H. Smidt Mogensen, C. Børsting, N. Morling, Frequencies of HID-ion ampliseq ancestry panel markers among Greenlanders, Forensic Sci. Int. Genet. 24 (2016) 60–64, https://doi.org/10.1016/j.fsigen.2016.06.001.
[44] M.T.A. Ta, N.N. Nguyen, D.M. Tran, T.H. Nguyen, T.A. Vu, D.T. Le, P.T. Le, T.T. H. Do, H. Hoang, H.H. Chu, Massively parallel sequencing of human skeletal remains in Vietnam using the precision ID mtDNA control region panel on the Ion S5™ system, Int. J. Leg. Med. 135 (2021) 2285–2294, https://doi.org/10.1007/s00414-021-02649-1.
[45] J.C. Cihlar, C. Amory, R. Lagacé, C. Roth, W. Parson, B. Budowle, Developmental Validation of a MPS Workflow with a PCR-Based Short Amplicon Whole Mitochondrial Genome Panel, Genes (Basel) 11 (2020) 1345, https://doi.org/10.3390/genes11111345.
[46] 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, https://doi.org/10.1016/j.fsigen.2016.04.004.
[47] R. Tao, W. Qi, C. Chen, J. Zhang, Z. Yang, W. Song, S. Zhang, C. Li, Pilot study for forensic evaluations of the Precision ID GlobalFiler™ NGS STR Panel v2 with the Ion S5™ system, Forensic Sci. Int. Genet. 43 (2019) 102147, https://doi.org/10.1016/j.fsigen.2019.102147.
[48] A. Ralf, M. van Oven, D.M. González, P. de Knijff, K. van der Beek, S. Wootton, R. Lagacé, M. Kayser, Forensic Y-SNP analysis beyond SNaPshot: High-resolution Y-chromosomal haplogrouping from low quality and quantity DNA using Ion AmpliSeq and targeted massively parallel sequencing, Forensic Sci. Int. Genet. 41 (2019) 93–106, https://doi.org/10.1016/j.fsigen.2019.04.001.
[49] F. Oldoni, D. Bader, C. Fantinato, S.C. Wootton, R. Lagacé, K.K. Kidd, D. Podini, A sequence-based 74plex microhaplotype assay for analysis of forensic DNA mixtures, Forensic Sci. Int. Genet. 49 (2020) 102367, https://doi.org/10.1016/j.fsigen.2020.102367.
[50] M. Diepenbroek, B. Bayer, K. Schwender, R. Schiller, J. Lim, R. Lagacé, K. Anslinger, Evaluation of the Ion AmpliSeq™ PhenoTrivium Panel: MPS-Based Assay for Ancestry and Phenotype Predictions Challenged by Casework Samples, Genes (Basel) 11 (2020) 1398, https://doi.org/10.3390/genes11121398.
[51] C. Xavier, M. de la Puente, A. Mosquera-Miguel, A. Freire-Aradas, V. Kalamara, A. Vidaki, T. E. Gross, A. Revoir, E. Pośpiech, E. Kartasińska, M. Spólnicka, W. Branicki, C. E. Ames, P. M. Schneider, C. Hohoff, M. Kayser, C. Phillips, W. Parson, VISAGE Consortium, Development and validation of the VISAGE AmpliSeq basic tool to predict appearance and ancestry from DNA, Forensic Sci. Int. Genet. 48 (2020) 102336, https://doi.org/10.1016/j.fsigen.2020.102336.
[52] F. Guo, Y. Lang, G. Long, Z. Liu, G. Jing, Y. Zhou, B. Zhang, S. Yu, Ion Torrent ™ Genexus ™ Integrated Sequencer and ForeNGS Analysis Software-An automatic NGS-STR workflow from DNA to profile for forensic science, Forensic Sci. Int. Genet. 61 (2022) 102753, https://doi.org/10.1016/j.fsigen.2022.102753.
[53] Thermo Fisher Scientific, AutoMate Express™ Instrument, P/N 4441982, Rev.G, Thermo Fisher Scientific, Waltham, 2018. Available at: (⟨https://assets.thermofisher.cn/TFS-Assets/LSG/manuals/4441982_AutoMateExp_UG.pdf⟩).
[54] Thermo Fisher Scientific, Quantifiler™ HP and Trio DNA Quantification Kits, P/N 4485354, Rev.H, Thermo Fisher Scientific, Waltham, 2018. Available at: (⟨https://assets.thermofisher.cn/TFS-Assets/LSG/manuals/4485354.pdf⟩).
[55] Thermo Fisher Scientific, Qubit® dsDNA HS Assay Kits, P/N MAN0002326, Rev. B.0, Thermo Fisher Scientific, Waltham, 2015. Available at: (⟨https://assets.thermofisher.cn/TFS-Assets/LSG/manuals/Qubit_dsDNA_HS_Assay_UG.pdf⟩).
[56] Thermo Fisher Scientific, Precision ID GlobalFilerTM NGS STR Panel v2 with the HID Ion S5TM/HID Ion GeneStudioTM S5 System, P/N MAN0016129, Rev.B.0, Thermo Fisher Scientific, Waltham, 2018. Available at: (⟨https://assets.thermofisher.cn/TFS-Assets/LSG/manuals/MAN0016129_PrecisionIDSTRIonS5_UG.pdf⟩).
[57] W. Parson, D. Ballard, B. Budowle, J.M. Butler, K.B. Gettings, P. Gill, L. Gusmão, D. R. Hares, J.A. Irwin, J.L. King, P. Knijff, N. Morling, M. Prinz, P.M. Schneider, C. Van Neste, S. Willuweit, C. Phillips, Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements, Forensic Sci. Int. Genet. 22 (2016) 54–63, https://doi.org/10.1016/j.fsigen.2016.01.009.
[58] K.B. Gettings, M. Bodner, L.A. Borsuk, J.L. King, D. Ballard, W. Parson, C.C. G. Benschop, C. Børsting, B. Budowle, J.M. Butler, K.J. van der Gaag, P. Gill, L. Gusmão, D.R. Hares, J. Hoogenboom, J. Irwin, L. Prieto, P.M. Schneider, M. Vennemann, C. Phillips, Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on short tandem repeat sequence nomenclature, Forensic Sci. Int. Genet. 68 (2024) 102946, https://doi.org/10.1016/j.fsigen.2023.102946.
[59] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, 2015. ⟨https://www.r-project.org⟩ (Available at:).
[60] Thermo Fisher Scientific, Precision ID SNP Panels with the HID Ion S5TM/HID Ion GeneStudioTM S5 System, P/N MAN0017767, Rev.C.0, Thermo Fisher Scientific, Waltham, 2019. Available at: (⟨https://assets.thermofisher.cn/TFS-Assets/LSG/manuals/MAN0017767_PrecisionID_SNP_Panels_S5_UG.pdf⟩).
[61] Thermo Fisher Scientific, Maximize Forensic Genetic Analysis Potential Using an STR + SNP Combination NGS Workflow, Thermo Fisher Scientific, Waltham, 2019. Available at: (⟨https://assets.thermofisher.com/TFS-Assets/GSD/Technical-Notes/precision_id_str_snp_combo_2019_technical_note.pdf⟩).
[62] Verogen, ForenSeq DNA Signature Prep Reference Guide, Document # VD2018005 Rev. A. Verogen, San Diego, 2022. Available at: (⟨https://verogen.com/wp-content/uploads/2022/01/forenseq-dna-signature-prep-reference-guide-PCR1-vd2018005-d.pdf⟩).
[63] C. Xavier, W. Parson, Evaluation of the Illumina ForenSeq™ DNA Signature Prep Kit - MPS forensic application for the MiSeq FGx™ benchtop sequencer, Forensic Sci. Int. Genet. 28 (2017) 188–194, https://doi.org/10.1016/j.fsigen.2017.02.018.
[64] Scientific Working Group on DNA Analysis Methods, SWGDAM Validation Guidelines for DNA Analysis Methods. SWGDAM, Washington, 2016. Available at: (⟨https://www.swgdam.org/_files/ugd/4344b0_813b241e8944497e99b9c45b163b76bd.pdf⟩).
[65] Scientific Working Group on DNA Analysis Methods, The Guidance Document for the FBI Quality Assurance Standards for Forensic DNA Testing and DNA Databasing Laboratories. SWGDAM, Washington, 2023. Available at: (⟨https://www.swgdam.org/_files/ugd/4344b0_04e7c4147b3b4a00bda10ba1c9a874bd.pdf⟩).
[66] J.L. King, F.R. Wendt, J. Sun, B. Budowle, STRait Razor v2s: Advancing sequence-based STR allele reporting and beyond to other marker systems, Forensic Sci. Int. Genet. 29 (2017) 21–28, https://doi.org/10.1016/j.fsigen.2017.03.013.
[67] C.G. Jønck, X. Qian, H. Simayijiang, C. Børsting, STRinNGS v2.0: Improved tool for analysis and reporting of STR sequencing data, Forensic Sci. Int. Genet. 48 (2020) 102331, https://doi.org/10.1016/j.fsigen.2020.102331.
[68] F. Guo, Z. Liu, G. Long, B. Zhang, X. Dong, D. Liu, S. Yu, High-resolution genotyping of 58 STRs in 635 Northern Han Chinese with MiSeq FGx ® Forensic Genomics System, Forensic Sci. Int. Genet. 65 (2023) 102879, https://doi.org/10.1016/j.fsigen.2023.102879.
[69] Y. Xu, Q. Pang, Repetitive DNA sequences in the human Y chromosome and male infertility, Front. Cell Dev. Biol. 10 (2022) 831338, https://doi.org/10.3389/fcell.2022.831338.
[70] R. Schoske, P.M. Vallone, C.M. Ruitberg, J.M. Butler, Multiplex PCR design strategy used for the simultaneous amplification of 10 Y chromosome short tandem repeat (STR) loci, Anal. Bioanal. Chem. 375 (2003) 333–343, https://doi.org/10.1007/s00216-002-1683-2.