Forensic Science International: Genetics 31 (2017) 155–159

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Short communication

Applying massively parallel sequencing to paternity testing on the Ion Torrent Personal Genome Machine

Hui Li[a,*], Xueying Zhao[a], Ke Ma[a], Yu Cao[b], Huaigu Zhou[b], Yuan Ping[b], Chengchen Shao[c], Jianhui Xie[c], Wenbin Liu[a,*]

[a] Shanghai Key Laboratory of Crime Scene Evidence, Shanghai Research Institute of Criminal Science and Technology, Shanghai 200083, China
[b] Key Laboratory of Forensic Evidence and Science Technology, Ministry of Public Security, Institute of Forensic Science, Shanghai Public Security Bureau, Shanghai 200083, China
[c] Department of Forensic Medicine, Shanghai Medical College of Fudan University, Shanghai 200032, China

ARTICLE INFO

Keywords:
Forensic genetics
Massively parallel sequencing (MPS)
Paternity testing
Ion Torrent Personal Genome Machine (PGM)

ABSTRACT

Massively parallel sequencing (MPS) is a promising supplementary method for forensic genetics and has gradually been applied to forensic casework. In this study, we applied MPS to forensic casework on an Ion Torrent Personal Genome Machine to evaluate its performance in paternity testing with mismatched STR loci. A total of 15 samples from seven cases containing one mismatched locus by capillary electrophoresis typing were analyzed. Combined paternity index (CPI) and relative chance of paternity were calculated according to the International Society for Forensic Genetics guidelines and the Chinese national standards recommended for paternity testing. With simultaneous analysis of enough STR loci, the results support the certainty of paternity, and the mismatched alleles were considered to be mutations (CPI > 10,000). With the detection of allele sequence structures, the origins of the mutations were inferred in some cases. Meanwhile, nine STRs (CSF1PO, D1S1656, D2S441, D2S1338, D3S1358, D8S1179, D12S391, D21S11 and D4S2408) were found in an increased number of unique alleles and three new alleles in three STRs (D2S441, D21S11, and FGA) that have not been reported before were detected. Therefore, MPS can provide valuable information for forensic genetics research and play a promising role in paternity testing.

1. Introduction

Over the past few years, massively parallel sequencing (MPS), also known as next generation sequencing (NGS), has become a promising supplementary method for forensic genetics [1,2]. With its advantages such as high-throughput, lower cost, and more genetic information, MPS is now expected to make an excellent addition to capillary electrophoresis (CE), the conventional Short tandem repeat (STR) typing method, which has been widely recognized in criminal investigations and court trials for the past two decades [3–5].

Of the two current MPS platforms that are widely used in forensic genetics, the Miseq (Illumina, San Diego, CA, USA) and Ion Torrent Personal Genome Machine (PGM™, Thermo Fisher, Foster City, CA, USA), several studies [6–14] have already revealed the potential application of MPS for STR typing in forensic casework, including individual identification and paternity testing. In individual identification, valuable genetic information can be obtained from degraded, low-quantity DNA samples or mixture samples by MPS when CE-based typing cannot provide enough clues [7–12]. In paternity testing, MPS is more conducive for distinguishing between mutations and exclusions with more marker information. Indeed, Ma et al. has evaluated the improvement of resolution for paternity testing on the Miseq [14].

In the work presented here, we applied MPS to forensic casework using the PGM to evaluate its performance in paternity testing with mismatched STR loci. The MPS panel used in this study was the Precision ID GlobalFiler™ NGS STR panel (Thermo Fisher), which contained 20 Expanded U.S. Core Loci (CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, D22S1045, and Amelogenin), one Y-STR (DYS391), and nine additional STRs (D1S1677, D2S1776, D3S4529, D4S2408, D5S2800, D6S474, D6S1043, D12ATA63, and D14S1434).

* Corresponding authors.
E-mail addresses: lihui860225@163.com (H. Li), wbliu1981@163.com (W. Liu).

http://dx.doi.org/10.1016/j.fsigen.2017.09.007
Received 12 July 2017; Received in revised form 6 September 2017; Accepted 11 September 2017
Available online 12 September 2017

2. Materials and methods

2.1. Sample preparation

A total of 15 samples from seven cases (one trio and six duos) were analyzed. All of these samples were provided by the Department of Forensic Medicine, Shanghai Medical College of Fudan University from routine casework, and each case contained one mismatched locus in 20 autosomal STRs as determined by CE. All participants provided their written informed consent for the collection of blood samples. Genomic DNA was extracted using the BioRobotEZ1 Advanced XL and EZ1 DNA Investigator kits (Qiagen, Hilden, Germany) according to the manufacturer's protocols. The quantity of recovered DNA was estimated using a Qubit 2.0 Fluorometer with a Qubit dsDNA HS Assay Kit (Thermo Fisher). This study was approved by the Ethical Committee of Fudan University.

2.2. Library preparation

DNA libraries were constructed following the application guide of the Precision ID Panels with Ion PGM™ System provided by Thermo Fisher, which is available online: https://www.thermofisher.com/order/catalog/product/A30939. Briefly, the library preparation process involved a 20 μL PCR, containing 4 μL 5 × Ion Ampliseq™ HiFi Mix, 10 μL Precision ID STR panel, an 1 ng gDNA. Amplification conditions were 99 °C for 1 min, followed by 23 cycles of 99 °C for 15 s and 60 °C for 4 min, with a hold at 10 °C. Primers were partially digested by adding 2 μL of FuPa reagent (Thermo Fisher) to the reaction system, mixing, and thermal cycling at 50 °C for 10 min, 55 °C for 10 min, and 60 °C for 20 min, with a hold at 4 °C. Adaptors with barcodes were ligated to the amplicons by adding 0.5 μL Ion P1 Adapter, 0.5 μL Ion Xpress™ Barcode X (a different Ion Xpress Barcoded adaptor for each sample), 1 μL nuclease-free water, 4 μL Switch Solution, and 2 μL DNA Ligase to the reaction system, mixing, and thermal cycling at 22 °C for 30 min, and 68 °C for 10 min with a hold at 10 °C for up to 1 h. The Agencourt AMPure XP PCR purification system (Beckman Coulter, Fullerton, CA, USA) was used to purify the libraries as recommended by the manufacturer. All libraries were quantified on a LightCycler® 96 Real-time PCR System (Roche Diagnostics, Rotkreuz, Switzerland) with the Ion Library TaqMan® Quantitation Kit (Thermo Fisher) and then normalized and pooled according to the manufacturer's recommendations.

2.3. Emulsion PCR and ion PGM™ sequencing

The pooled libraries were processed for emulsion PCR on an Ion OneTouch™ 2 instrument (Thermo Fisher) with the Ion PGM™ Hi-Q™ OT2 Kit (Thermo Fisher) following the manufacturer's protocol. Emulsion PCR products were enriched for template positive Ion Sphere™ Particles (ISPs) using the Ion OneTouch™ Enrichment System (ES) (Thermo Fisher). Sequencing was accomplished on the Ion Torrent™ PGM System using the Ion PGM™ Sequencing Hi-Q™ Kit (Thermo Fisher) following the manufacturer's sequencing protocol. In this study, all 15 libraries were run on one Ion 318 v2 Chip (Thermo Fisher).

2.4. Data analysis

Sequencing data were generated by the Torrent Suite software (v5.0). The HID STR Genotyper (v3.0) plugin (Thermo Fisher) was used to analyze the sequencing results. The length and sequence variants per locus and per allele were compiled and counted. Allele calls with a minimum of 50× coverage were involved for further analysis [15]. The sequences that were one repeat motif shorter or longer than the corresponding allele were considered as stutters.

2.5. Confirmation by CE-based typing

To confirm the STR typing results obtained by MPS, two commercial CE-based STR kits, the GlobalFiler™ PCR Amplification Kit (Thermo Fisher) and AGCU 21 + 1 STR kit (AGCU, Wuxi, China) were used in, which all STR loci of the Precision ID GlobalFiler™ NGS STR panel were included. PCR products were separated and detected on an ABI 3500XL Genetic Analyzer (Thermo Fisher) according to the manufacturer's instructions. The electrophoretic results were analyzed using GeneMapper® ID-X software v1.4 (Thermo Fisher). Sanger-type sequencing was performed to verify the new alleles that have not yet been reported, using a BigDye1 Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher).

2.6. Calculation of paternity testing parameters

The paternity index (PI) was calculated following the Chinese national standards recommended for paternity testing (**Supplementary** Table 1) on the basis of Bayesian mathematics [16]. The combined paternity index (CPI) was calculated by the equation $CPI = \prod_{i=1}^{n} PI_i$, where n is the number of STR loci. The relative chance of paternity (RCP) was calculated by the equation RCP = CPI/(CPI + 1).

3. Results and discussion

Currently, MPS is a hot research method. Two MPS platforms, the Miseq and Ion Torrent, are mainly used in forensic genetics. The work by Ma et al. has demonstrates that MPS provides improved resolution for paternity testing on the Miseq [11]. In this study, we applied MPS to forensic casework on the PGM to evaluate its performance in paternity testing with mismatched STR loci by using the Precision ID GlobalFiler™ NGS STR panel. A total of 15 samples from seven cases were run on one Ion 318 v2 Chip. On average, each sample was represented with 65,800 sequence reads (range: 16693–147156 reads). The sequencing reads of each sample are shown in **Supplementary** Fig. 1 . Allele calls were involved for further analysis with a coverage threshold of 50× [12]. The coverage of stutters was all < 20% of the corresponding allele. All profiles of the 15 samples were obtained without missing loci.

According to the International Society for Forensic Genetics (ISFG) guidelines [17,18] and the Chinese national standards recommended for paternity testing, CPI and RCP should be calculated in paternity tests. When some loci do not conform to genetic rules, the PI calculated for all STR loci including the mismatched alleles should been considered. When the CPI > 10,000, the certainty of paternity should be supported, and the mismatched alleles are considered to be mutations. Otherwise, more loci need to be detected to distinguish between mutations and exclusions. With the CE method, because similarly sized amplicons must be labeled with different fluorescent markers, the number of loci multiplexed together is limited [7]. Two or more STR kits are often needed. By contrast, because MPS does not require size separation between amplicons, a large number of loci can be simultaneously analyzed. Further, multiple samples can be analyzed in a single reaction through the use of barcoding. Therefore, MPS overcomes the limitations of CE method and improves efficiency in paternity testing, especially when there are mismatched loci. In this study, the Precision ID GlobalFiler™ NGS STR panel contained 30 loci (including Amelogenin and DYS391). Given lack of mass MPS sequence data in Chinese Han population, especially in part of loci (D1S1677, D2S1776, D3S4529, D4S2408, D5S2800, D6S474, D12ATA63, and D14S1434), the PI was calculated based on the frequency of CE-based alleles from the Chinese Han population provided by the Chinese national standards recommended for paternity testing. The paternity testing parameters of each case are represented in Table 1. All loci including the mismatched loci was taken into account in the PI calculation. In all seven cases, the