Forensic Science International: Genetics 20 (2016) 140–146

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Characterization of microRNA expression profiles in blood and saliva using the Ion Personal Genome Machine® System (Ion PGM™ System)

Zheng Wang[a], Di Zhou[b], Yandong Cao[b], Zhen Hu[a], Suhua Zhang[a], Yingnan Bian[a], Yiping Hou[c], Chengtao Li[a,*]

[a] Shanghai Key Laboratory of Forensic Medicine, Institute of Forensic Science, Ministry of Justice, P.R. China, Shanghai 200063, China
[b] Thermo Fisher Scientific, Shanghai, China
[c] Department of Forensic Genetics, West China School of Preclinical and Forensic Medicine, Sichuan University (West China University of Medical Sciences), Chengdu, China

ARTICLE INFO

Article history:
Received 17 May 2015
Received in revised form 27 October 2015
Accepted 27 October 2015
Available online 1 November 2015

Keywords:
miRNA
Massively parallel sequencing (MPS)
Ion PGM™ System
Body fluid identification
Forensic science

ABSTRACT

MicroRNA (miRNA) expression profiling is gaining interest in the forensic community because the intrinsically short fragment and tissue-specific expression pattern enable miRNAs as a useful biomarker for body fluid identification. Measuring the quantity of miRNAs in forensically relevant body fluids is an important step to screen specific miRNAs for body fluid identification. The recent introduction of massively parallel sequencing (MPS) has the potential for screening miRNA biomarkers at the genome-wide level, which allows both the detection of expression pattern and miRNA sequences. In this study, we employed the Ion Personal Genome Machine® System (Ion PGM™ System, Thermo Fisher) to characterize the distribution and expression of 2588 human mature miRNAs (miRBase v21) in 5 blood samples and 5 saliva samples. An average of 1,885,000 and 1,356,000 sequence reads were generated in blood and saliva respectively. Based on miRDong, a Perl-based tool developed for semi-automated miRNA distribution designations, and manually ascertained, 6 and 19 miRNAs were identified respectively as potentially blood and saliva-specific biomarkers. Herein, this study describes a complete and reliable miRNA workflow solution based on Ion PGM™ System, starting from efficient RNA extraction, followed by small RNA library construction and sequencing. With this workflow solution and miRDong analysis it will be possible to measure miRNA expression pattern at the genome-wide level in other forensically relevant body fluids.

1. Introduction

MicroRNAs (miRNAs, ~22 nucleotides) belong to a class of small noncoding RNA molecules that regulate diverse physiological and pathological processes by post-transcriptional mechanisms [1–3]. Several studies have examined the relative abundance of miRNAs in human tissue, and reported that some miRNAs are expressed in a tissue-specific manner [4–6]. The intrinsically shorter fragment makes miRNA less susceptible to degradation, even feasible and valid miRNA profiling results could be obtained from formalin-fixed paraffin-embedded (FFPE) tissue samples [7,8]. This has led to considerable interest in the development of miRNAs as novel biomarkers for forensic purposes [9–19]. In several forensic laboratories, differentially expressed miRNAs have been investigated as a potential means for body fluid identification [9–15], but most of the suggested miRNAs were not independently evaluated and there are few agreements among them. Additionally, previous studies screening body fluid-specific miRNAs mainly based on microarray techniques [10–13], which restricted to a limited number of miRNAs and imperfect specificity in some cases for miRNAs that are closely related in sequence. Therefore, microarray outputs need to be verified by other methods. Currently, there are more than 2500 known human miRNAs in miRBase v21 (http://www.mirbase.org) [20], and measuring the quantity of miRNAs in forensically relevant body fluids at the genome-wide level is an important step to screen novel miRNA biomarkers for body fluid identification.

Massively parallel sequencing (MPS), also termed next generation sequencing (NGS), has the capability to sequence many targeted regions of multiple samples simultaneously with high coverage [21]. To date, multiple studies have been conducted to sequence DNA markers using MPS, including genetic investigations

* Corresponding authors at: West Guangfu Road 1347, Shanghai 200063, China. Fax: +86 21 52352959.
E-mail address: lichengtaohla@163.com (C. Li).

http://dx.doi.org/10.1016/j.fsigen.2015.10.008
1872-4973/

of STR [22–24] and SNP [25,26], and mtDNA validation [27–29]. All these studies indicate that DNA typing by MPS holds promise for forensic applications. However, little work has been reported on the use of the MPS to analyze the miRNA expression pattern in forensically relevant body fluids. The Ion Personal Genome Machine® System (Ion PGM™ System), released in 2012, is a non-optical sequencing platform that exploits CMOS integrated circuits to detect pH changes by release of protons during incorporation of nucleotides into growing strands [30]. The Ion PGM™ System would be an ideal candidate for characterization of microRNA expression profiles in body fluids due to its running costs, sequencing time (<2 h, 160 flows) and scalability (three different chip sizes: 314™, 316™ and 318™ Chip).

In continuation to our previous studies [13–15], the present study characterizes the distribution and expression of miRNAs at the genome-wide level in blood and saliva samples using the Ion PGM™ System. Here we describe a complete miRNA workflow solution, starting from efficient extraction of RNA from two kinds of body fluids followed by small RNA library construction and sequencing. In addition, miRDong, a Perl-based tool, was developed for facilitating semi-automated miRNA distribution designations.

2. Materials and methods

2.1. Collection of body fluid samples

Human blood and saliva samples were collected with the approval of the Ethics Committee of the Institute of Forensic Sciences, Ministry of Justice, P.R. China. Body fluid samples were collected from 5 unrelated healthy individual donors after receiving written informed consents. Peripheral blood was collected by venipuncture without anticoagulation treatment, and 200 μL was used to extract total RNA. Unstimulated (at least 1 h after eating or drinking) saliva was collected in plastic tubes, and 300 μL was used to extract total RNA. In addition, triplicate samples (200 μL blood per replicate) were collected from a given individual for assessing the reliability of the workflow solution.

2.2. RNA isolation and quantification, and small RNA enrichment

Total RNA including miRNA was isolated using the mirVana™ miRNA Isolation kit (Thermo Fisher, Foster City, CA, USA) according to the manufacturer's instructions. The amount of RNA in the sample was estimated by the NanoDrop® Spectrophotometer (Thermo Fisher).

The small RNA content in the RNA sample was determined with an Agilent 2100 Bioanalyzer® instrument (Agilent Technologies, Santa Clara, CA, USA). The RNA 6000 Nano chip (Agilent Technologies) was used for determining the quality (RNA integrity number, RIN) and concentration of total RNA, and then the Small RNA chip (Agilent Technologies) was used for gaining a more detailed view of small RNAs in the 10 to 40-nucleotide size range. The miRNA content in the RNA sample was calculated using the formula: % miRNA = (mass of miRNA ÷ mass of total RNA) × 100. According to the manufacturer's instructions, if miRNA content ≥ 0.5%, small RNA enrichment is not needed. However, for optimal results, small RNA was enriched from total RNA samples using the Magnetic Bead Cleanup Module (Thermo Fisher) in this study. The quantity of the small RNA-enriched sample was assessed again with the Agilent® Small RNA Kit following the manufacturer's instructions.

2.3. The small RNA library construction

Small RNA libraries were prepared using the Ion Total RNA-Seq Kit v2 protocol (Revision E, http://ioncommunity.lifetechnologies.com/docs/DOC-2999) and materials with minor modifications.

For the ligation of small RNAs with adaptor, 3 μL of small RNA sample (~2 ng) was combined with the hybridization reagents (5 μL) and incubated at 65 °C for 10 min and 16 °C for 5 min. Ligation reagents (12 μL) were then added and the samples (20 μL) were incubated at 16 °C for 16 h.

After ligation, reverse transcription (RT) was performed on a GeneAmpPCR System 9700: 16 μL RT master mix was added to the ligated RNA samples (20 μL), tubes were incubated at 70 °C for 10 min, and then snap-cooled on ice, the RT enzyme (4 μL of 10X SuperScript® III Enzyme Mix) was added, and reactions (40 μL) were incubated at 42 °C for 30 min. cDNA from the RT reaction was purified and size-selected using the Magnetic Bead Cleanup Module and eluted in 12 μL of nuclease-free water.

For amplifying the cDNA, 6 μL of the purified cDNA was combined with 0.5 μL of Ion Xpress™ RNA 3′ Barcode Primer, 0.5 μL of Ion Xpress™ RNA-Seq Barcode BC primer (choose from BC01–BC16 for different samples), 45 μL of Platinum PCR SuperMix High Fidelity and 1 μL of nuclease-free water, reaction mix (53 μL) was then placed in a GeneAmp PCR System 9700 and amplified using the following protocol: 94 °C for 2 min; (94 °C for 30 s, 50 °C for 30 s, and 68 °C for 30 s) for 2 cycles; (94 °C for 30 s, 62 °C for 30 s, and 68 °C for 30 s) for 14 cycles; 68 °C for 5 min. The amplified DNA for each sample was purified using the Magnetic Bead Cleanup Module and eluted in 15 μL of nuclease-free water.

To assess the yield and size distribution, 1 μL of the barcoded libraries were assessed by Agilent 2100 Bioanalyzer® instrument with High Sensitivity DNA Kit (Agilent Technologies) following the manufacturer's recommendations. The ratio of mRNA ligation products in total ligation products was calculated using the formula for Barcoded libraries: [Area (94–114 bp)] ÷ [Area (50–300 bp)]. The molar concentration of the barcoded library was determined using size range 50–300 bp, then diluted to a final concentration of ~20 pM following the manufacturer's protocol. Equal volumes of the two diluted libraries were combined for the next steps.

2.4. Template preparation and Ion PGM™ sequencing

The diluted library (15 μL) was used to generate template positive Ion Sphere™ Particles (ISPs) containing clonally amplified DNA. Emulsion PCR (emPCR) was conducted in the Ion OneTouch™ 2 System (Thermo Fisher) with the Ion PGM™ Template OT2 200 Kit (Thermo Fisher) following the recommended protocol (Revision A.0, http://ioncommunity.lifetechnologies.com/docs/DOC-6772). Template-positive ISPs were enriched with the Ion OneTouch™ ES (Thermo Fisher) following the manufacturer's recommendations.

Sequencing was performed on the Ion Torrent PGM™ instrument using the Ion 318™ Chip v2 with 160 flows following the protocol of Ion PGM™ Sequencing 200 Kit v2 (Revision 3.0, http://ioncommunity.lifetechnologies.com/docs/DOC-6775). Sequencing primer and Control Ion Spheres™ Particles of the Ion PGM sequencing kit were added to the enriched, template-positive ISPs. After annealing the sequencing primer, sequencing polymerase was added and a final volume of 30 μL was loaded onto the chips.

2.5. Sequence data acquisition and analysis

All sequence data generated in this study were analyzed using the Torrent Suite v4.2 (Thermo Fisher) with the default settings. The generated BAM (Binary Alignment Map) files were used for secondary analysis. For the purposes of this study, the following criteria were used for miRNAs calling: the 2588 human mature miRNA sequences (miRBase v21) as alignment reference; a perfect matching required, i.e., no longer, no shorter and no mismatch; and sequence reads ≥ 100 counts. The abundance of a given miRNA in a