ORIGINAL ARTICLE

Comparison of the performance of Ion Torrent chips in noninvasive prenatal trisomy detection

Yanlin Wang^{1,2,7}, Zujia Wen^{3,7}, Jiawei Shen^{3,7}, Weiwei Cheng¹, Jun Li⁴, Xiaolan Qin⁵, Duan Ma^{2,6} and Yongyong Shi³

Semiconductor high-throughput sequencing, represented by Ion Torrent PGM/Proton, proves to be feasible in the noninvasive prenatal diagnosis of fetal aneuploidies. It is commendable that, with less data and relevant cost also, an accurate result can be achieved owing to the high sensitivity and specificity of such kind of technology. We conducted a comparative analysis of the performance of four different Ion chips in detecting fetal chromosomal aneuploidies. Eight maternal plasma DNA samples, including four pregnancies with normal fetuses and four with trisomy 21 fetuses, were sequenced on Ion Torrent 314/316/318/PI chips, respectively. Results such as read mapped ratio, correlation coefficient and phred quality score were calculated and parallelly compared. All samples were correctly classified even with low-throughput chip, and, among the four chips, the 316 chip had the highest read mapped ratio, correlation coefficient, mean read length and phred quality score. All chips were well consistent with each other. Our results showed that all Ion chips are applicable in noninvasive prenatal fetal aneuploidy diagnosis. We recommend researchers or clinicians to use the appropriate chip with barcoding technology on the basis of the sample number.

Journal of Human Genetics (2014) 59, 393–396; doi:10.1038/jhg.2014.40; published online 12 June 2014

INTRODUCTION

Chromosomal aneuploidy disease represented by trisomy 21, which affects 1 in 800 neonates,¹ is a common class of genetic disorder. As cell-free fetal DNA was discovered to be present in maternal circulation,² noninvasive prenatal detection of fetal aneuploidies by massively parallel sequencing has become an important means for reducing this kind of birth defects. Compared with invasive detection methods, namely chorionic villi sampling and amniocentesis, which are the gold standards of prenatal diagnosis, plasma DNA deep sequencing decreases the risk of miscarriage or stillbirth. Also, with the rapid development of sequencing technology, diagnostic accuracy of the noninvasive method can be more than 99%,³ which is even higher than that of chorionic villi sampling and amniocentesis whose accuracies are 98 and 99%, respectively.⁴

Next-generation sequencing was first demonstrated to be feasible in Noninvasive prenatal diagnosis of fetal trisomy in 2008 by two research groups.^{5,6} Different from other noninvasive diagnosis methods targeted toward fetal material, it is a comprehensive deep sequencing of all genetic material existing in maternal plasma. As we know, cell-free fetal DNA coexists with a large background of maternal DNA in the circulation during pregnancy. Recent study

has shown that it almost amounted to 10–20% of the total material.⁷ Therefore, the key problem in next-generation sequencing data analysis is whether the equipment can ‘read’ the difference in chromosome dosage and whether the analysis method has sufficient sensitivity to differentiate between an extra and a missing chromosome. According to previous studies, with most frequently used next-generation sequencing platforms and the z-score analysis method, it is possible to measure the proportion of chromosome 21 and to test fetal chromosome aneuploidy with sufficient sequencing reads.⁵

High-throughput sequencing is a suitable way for noninvasive prenatal diagnosis owing to its easy operation, high success ratio and remarkable accuracy. Moreover, it is a better method than the PCR amplification of a particular region, as we obtain data of the whole genome through a sequencing run, which allow multiple examinations other than trisomy detection, such as fetal sex determination, to be performed at the same time.

To meet increasing demands in research and clinical application, a more flexible and efficient sequencing approach is urgently needed. Meanwhile, the newly developed semiconductor sequencing technology, Ion Torrent PGM and Proton, seem to have the

¹Prenatal Diagnosis Center, International Peace Maternity and Child Health Hospital, Shanghai Jiao Tong University, Shanghai, China; ²Key Laboratory of Molecular Medicine, Ministry of Education, Department of Biochemistry and Molecular Biology, Institutes of Biomedical Sciences, Shanghai Medical College, Fudan University, Shanghai, China; ³Bio-X Institutes, Shanghai Jiao Tong University, Shanghai, China; ⁴Department of General Surgery, Shanghai Changning Center Hospital, Shanghai, China; ⁵Shanghai genomePilot Institutes of Genomics and Human Health, Shanghai, China and ⁶Shanghai Key Laboratory of Prevention and Intervention of Birth Defects, Shanghai, China

⁷These authors contributed equally to this work, and all three should be considered first authors.

Correspondence: Professor D Ma, Mailbox 238, 138 Medical College Road, Shanghai, China.

E-mail: duanma@shmu.edu.cn

or Professor Y Shi, Bio-X Institutes, Shanghai Jiao Tong University, The Central Little White House, 1954 Huashan Road, Shanghai 200030, China.

E-mail: 13764928889@163.com

Received 7 January 2014; revised 21 March 2014; accepted 16 April 2014; published online 12 June 2014

ability to solve many problems of the sequencing methods in current use.

Unlike traditional next-generation sequencing platforms, which are based on the theory of fluorescence imaging, the semiconductor sequencer reads nucleotides by sensing the pH value variation when the synthetic sequencing reaction is proceeding, which greatly shortens the sequencing time while maintaining the same or even higher accuracy. These bench-top sequencers also sharply decrease the experiment cost, making sequencing or detection affordable for normal people, and extending its application range. With Ion 314/316/318 chips of PGM, raw data are available from 70 Mbp to 2 Gbp, equal to 0.5–7 M reads of 200–400 bp, and 85% of the data can reach Q20. With PI chip of Proton, nearly 10G raw data are available, equal to 80M reads of 200 bp, and 75% of the data on average can reach Q20. Recent work by other groups has shown that it is feasible for PGM to perform prenatal diagnosis.⁸ To test the performance of different Ion chips in noninvasive fetal trisomy detection, especially the newly released Proton system, and estimate appropriate data size for analysis, we conducted a chip comparison experiment in this study.

MATERIALS AND METHODS

We selected maternal plasma DNA of eight pregnant women to perform sequencing analysis using the Ion 314/316/318 and PI chips. As the capacities of PGM chips are much smaller than those of the PI chips, we intended to test one sample on the 314/316/318 chips and several samples were pooled on a PI chip to be sequenced simultaneously; nearly 1G data were accessible for one sample.

Recruitment criteria

Participants of this study were recruited from two hospitals in China between 2011 and 2012. All of them had received full karyotyping results by chorionic villus sampling or amniocentesis before our noninvasive diagnosis. We recorded their ages and gestational weeks at blood sampling. To eliminate the influence of a too-short gestational period on the accuracy of detection results, the tested pregnancies must have a gestational duration of more than 12 weeks⁹ (average 18.875 weeks).

Sample processing

Fresh maternal peripheral blood was collected into EDTA-anticoagulant tubes and plasma was immediately separated by a double centrifugation method (1600 g, 10 min at 4 °C followed by 16000 g, 10 min at 4 °C). Plasma could be

stored at –80 °C and freezing-thawing was performed only once before further processing. Cell-free DNA was extracted from 0.6 ml plasma with the TIANamp Micro DNA Kit (TIANGEN, Beijing, China) according to the manufacturer's instructions. DNA samples had been quantified with the Qubit fluorometer before being constructed into sequencing libraries.

Semiconductor sequencing

DNA samples of 10–50 ng were used to construct the Ion Torrent fragment library according to the manufacturer's protocol. The libraries were universal for PGM and Proton systems. Each library had been analyzed with Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) to assess its length distribution and all of them were eligible for sequencing analysis (for more details, please refer to Supplementary Information). Each sample had been sequenced four times on 314/316/318 chips of personal genome machine (PGM) and PI chips of Proton, respectively, and all data were summarized and analyzed parallelly to assess the quality. In this way, noises could be reduced to minima so that a more objective result should be acquired.

Data analysis

The basic statistics were generated using SAMtools (samtools.sourceforge.net/) and BamUtil (<http://genome.sph.umich.edu/wiki/BamUtil>). Chromosomal aneuploidy detection was performed by in-house scripts.

RESULTS

Analysis results

Among these eight samples, four were normal and four were of trisomy 21. The karyotyping analysis results were shown in the Figure 1. All the samples were correctly classified. That is to say, even chips with low throughputs can provide significant results.

Total reads

Different chips have different throughputs. Although the total reads that every chip can obtain varies, we can evaluate the quality of the chips by comparing the stability of the chips. Eight identical samples were sequenced by the four chips and the average read number of the eight samples sequenced by the corresponding chip was calculated. In addition, another statistic, the standard deviation of the read number divided by the average read number was also provided to show the stability of the four chips (Table 1). From the table, we can see that the 316 chip has the lowest value of s.d./average, which means that, among the eight samples sequenced, the total reads of the sample acquired by the 316 chip floated the least so that the 316 chip is the most stable.

Percentage of reads mapped to genome

The ion torrent PGM generates a great number of reads. How many of these reads can be mapped to the genome is also a factor for evaluating the chips. The results (Table 1) showed that all the four chips performed well and that, on average, more than 98% of the reads could be mapped to genome. However, the 316 chip has the highest average ratio of reads mapped to genome (99.197%), whereas the PI chip has the lowest standard deviation (0.10908).

Correlation coefficient

As the whole genome was sequenced for every sample, the longer the length of the chromosome, the more reads it contained. Here, we used the correlation coefficient between the length of the chromosomes and the total reads of the corresponding chromosome to quantify it and evaluated the sequencing quality. Data with high quality should have a correlation coefficient close to 1. The 316 chip had the highest average correlation coefficient, whereas the 318 chip had the lowest standard deviation (Table 1).

Figure 1 Karyotyping result of chromosome 21 (chr21). Samples 1–4 were normal and samples 5–8 were of trisomy 21. All samples were sequenced by the 314, 316, 318 and PI chips and they were all correctly classified. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

Table 1 Comparison of five different parameters reflecting the sequencing quality of four Ion chips

Chip	Total reads		Read mapped ratio		Correlation coefficient		Mean read length		Phred quality score		
	Average	s.d./average	Average	s.d.	Average	s.d.	Average	s.d.	Q10	Q20	Q30
314	486706	0.1914	98.641	0.64959	0.98836	0.000896	148.328	10.1313	0.99415	0.889997	0.582045
316	3343196	0.0932	99.197	0.17872	0.98869	0.000865	158.3969	4.1036	0.995323	0.892954	0.583297
318	5898378	0.1504	98.358	1.64827	0.98868	0.000779	155.8881	10.0112	0.991687	0.878238	0.586263
PI	6246703 ^a	0.1393	98.469	0.10908	0.98843	0.00107	132.8963	4.1367	0.979589	0.808553	0.064222

^aIn our study, ten samples are pooled on a single PI chip, therefore this value is the average of total reads from the ten samples. Numbers in bold indicate the best results among these four chips.

	314	316	318	PI
314		0.973148	0.955434	0.948002
316			0.985339	0.968675
318				0.991552
PI				

Figure 2 Correlation of chromosome 21 (chr21)% between chips. The percentage of chr21 sequenced by the 314, 316, 318 and PI chips was compared. This figure showed that the results from different chips were highly correlated.

Mean read length

The length of the reads contributes a lot to the mapping quality. As the read length grows, the number of reads that uniquely map to the genome also increases. Among the four chips, the 316 chip had the longest mean read length and the lowest standard deviation (Table 1).

Phred quality score

The phred quality score is a widely accepted criterion to characterize the quality of DNA sequences. A phred quality score of 20 means that the chances of this base being called incorrectly are 1 in 100 and a phred quality score of 30 means 1 in 1000 and so on.

We calculated the Q10, Q20 and Q30 of every sample, and, for every chip, we averaged them among the eight samples sequenced on the same chip. From the results, we found that the sequencing quality of the 316 chip was higher than that of the 314, 318 and PI chips, for 316 had the highest Q10 and Q20, whereas the Q30 was slightly lower than 318 (Table 1).

Consistency between chips

In order to see the consistency, we gave the correlation coefficient between chips. Figure 2 indicated that the amount of chromosome 21 from samples was highly correlated.

Chip capacity for trisomy diagnosis

Our study has already shown that even a chip with low throughput, such as the 314 chip, could also provide right results. Assuming the minimum data amount for trisomy diagnosis to be the throughput of 314 chips, namely 0.5M reads, theoretically, 6 samples can be pooled on a 316 chip, 12 samples on a 318 chip and 150 samples on a PI chip, respectively. However, as the quality control step may reduce the usable total reads and as more data amounts always ensure a lower false positive rate, pooling fewer samples than that theoretical value will be better.

DISCUSSION

In this paper, we performed a chip comparison study of Ion Torrent systems for noninvasive prenatal fetal aneuploidy diagnosis. Remarkably, Proton I was introduced to this study and compared with PGM. Through the output data we could see that the PI system was still not fully mature in its sequencing capability. The sequencing length on average was shorter than that of PGM and overall data quality was also not quite as good as that from PGM. In general, the 316 chip proved to be better than the others in most aspects, including longest mean read length, highest mapping rate, best correlation coefficient and highest Q10 and Q20. On the other hand, although the PI chip had the highest throughput, the sequencing quality still needed to be improved, especially for the read length and Q30, which were far behind the other three chips.

Although the throughputs varied among chips, all the samples were correctly classified as T21 or not. For quantitative study, such as noninvasive prenatal fetal aneuploidy diagnosis, the throughput seems not to be a challenge. Chip with low throughput, such as the 314 chip, can also give the right result. This enables us to perform such kind of diagnosis at a low cost.

The sequence quality is an important factor contributing to the right diagnosis. Efforts can be made to improve the quality of raw data, in the expense of the amount of data. The raw data should be subjected to quality control, such as filtering reads with low score or only retaining reads that can be uniquely mapped, before it can be used. The data quality of the PI chip is far behind the other three in many ways. However, rigid quality control makes it possible to obtain the right results.

In conclusion, all Ion chips are applicable in noninvasive prenatal fetal aneuploidy diagnosis, and we recommend researchers and clinicians to use appropriate chips with barcoding technology, based on the sample number.

ACKNOWLEDGEMENTS

This work was supported by the Shanghai Committee of Science and Technology, China (grant no. 134119a4600).

- 1 Driscoll, D. A. & Gross, S. Clinical practice. Prenatal screening for aneuploidy. *N. Engl. J. Med.* **360**, 2556–2562 (2009).
- 2 Lo, Y. M. D., Corbetta, N., Chamberlain, P. F., Rai, V., Sargent, I. L., Redman, C. W. G. *et al.* Presence of fetal DNA in maternal plasma and serum. *Lancet* **350**, 485–487 (1997).
- 3 Chiu, R. W., Akolekar, R., Zheng, Y. W., Leung, T. Y., Sun, H., Chan, K. C. *et al.* Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study. *BMJ* **342**, c7401 (2011).
- 4 Chachkin, C. J. What potent blood: non-invasive prenatal genetic diagnosis and the transformation of modern prenatal care. *Am. J. Law. Med.* **33**, 9–53 (2007).
- 5 Chiu, R. W., Chan, K. C., Gao, Y., Lau, V. Y., Zheng, W., Leung, T. Y. *et al.* Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc. Natl Acad. Sci. USA* **105**, 20458–20463 (2008).

- 6 Fan, H. C., Blumenfeld, Y. J., Chitkara, U., Hudgins, L. & Quake, S. R. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. *Proc. Natl Acad. Sci. USA* **105**, 16266–16271 (2008).
- 7 Lun, F. M., Chiu, R. W., Allen Chan, K. C., Yeung Leung, T., Kin Lau, T. & Dennis, Lo YM. Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clin. Chem.* **54**, 1664–1672 (2008).
- 8 Yuan, Y., Jiang, F., Hua, S., Du, B., Hao, Y., Ye, L. *et al.* Feasibility study of semiconductor sequencing for noninvasive prenatal detection of fetal aneuploidy. *Clin. Chem.* **59**, 846–849 (2013).
- 9 Lo, Y. M., Tein, M. S., Lau, T. K., Haines, C. J., Leung, T. N., Poon, P. M. *et al.* Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenatal diagnosis. *Am. J. Hum. Genet.* **62**, 768–775 (1998).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)