Journal of Maternal-Fetal and Neonatal Medicine, 2012; Early Online: 1–5

ISSN 1476-7058 print/ISSN 1476-4954 online
DOI: 10.3109/14767058.2011.635730

informa
healthcare

Noninvasive prenatal diagnosis of common fetal chromosomal aneuploidies by maternal plasma DNA sequencing

Tze Kin Lau[1*], Fang Chen[2*], Xiaoyu Pan[2,3], Ritsuko K . Pooh[4], Fuman Jiang[2], Yihan Li[2], Hui Jiang[2], Xuchao Li[2], Shengpei Chen[2] & Xiuqing Zhang[2]

[1]Department of Obstetrics and Gynaecology, The Chinese University of Hong Kong, Hong Kong, [2]Guangdong Province Key Laboratory of Genome, BGI-Shenzhen, Shenzhen, China, [3]School of Bioscience and Bioengineering, South China University of Technology, Guangzhou, China, and [4]CRIFM Clinical Research Institute of Fetal Medicine PMC, Osaka, Japan

***Objective*: To develop a new bioinformatic method in the noninvasive prenatal identification of common fetal aneuploidies using massively parallel sequencing on maternal plasma.** ***Methods*: Massively parallel sequencing was performed on plasma DNA samples from 108 pregnant women (median gestation: 12^{+5} week) immediately before chorionic villus sampling (CVS) or amniocentesis. Data were analysed using a novel z-score method with internal reference chromosome. The diagnostic accuracies of the fetal karyotyping status were compared against two previously reported z-score methods – one without adjustment and the other with GC correction.** ***Results*: A total of 32 cases with fetal aneuploidy were confirmed by conventional karyotyping, including 11 cases of Trisomy 21, 10 cases of Trisomy 18, 2 cases of Trisomy 13, 8 cases of Turner syndrome (45, XO) and one case of Klinefelter syndrome (47, XXY). Using the z-score method without reference adjustment, the detection rate for Trisomy 21, Trisomy 18, Trisomy 13, Turner syndrome, and Klinefelter's syndrome is 100%, 40%, 0%, 88% and 0% respectively. Using the z-score method with GC correction, the detection rate increased to 100% for Trisomy 21, 90% for Trisomy 18, 100% for Trisomy 13. By using the z-score method with internal reference, the detection rate increased to 100% for all aneuploidies. The false positive rate was 0% for all three methods.** ***Conclusion*: This massively parallel sequencing-based approach, combined with the improved z-score test methodology, enables the prenatal diagnosis of most common aneuploidies with a high degree of accuracy, even in the first trimester of pregnancy.**

Keywords: Fetal aneuploidy, first trimester, massively parallel sequencing

Introduction

The discovery of cell-free fetal DNA in maternal peripheral plasma and serum [1] marked the start of noninvasive prenatal diagnosis. Clinical application for single gene disorders was quickly established and widespread even at national level, such as the fetal rhesus genotyping service [2]. On the other hand, its application for fetal aneuploidy, the most common indication for prenatal diagnosis and screening, has been far less successful. The real breakthrough was in 2008, when two independent research groups reported the successful prenatal diagnosis of fetal Trisomies from maternal plasma using massively parallel sequencing in cohorts of 28 and 18 cases, including 14 and 12 cases of aneuploidies respectively, with 100% accuracy [3,4]. Larger studies have been reported recently, confirming that noninvasive prenatal diagnosis of fetal Trisomy 21 using this approach was highly accurate, with a detection rate of 100% and a specificity of 97.9–99.7% [5,6].

However, most of the early studies are focused mainly on the detection of Trisomy 21 only, and many included advanced pregnancies as late as 36 weeks. Also, the detection of fetal Trisomy 21 was based on the calculation of the fractional genomic representation (FGR) of chromosome 21 (also known as percentage of chromosome 21) by dividing the number of unique reads from chromosome 21 by the total number of unique reads from the sample [5,6]. The deviation of this FGR from the expected value was measured using the z-score, calculated by subtracting the mean of the control group and dividing by the SD of the control group. However, applying this approach to other aneuploidies was less successful because of the larger coefficient of variation (CV), and hence lower precision, in the estimation of FGR of other chromosomes such as 13 and 18. This variation is known to be related to the different sequencing efficacy as a result of differences in size and GC contents between chromosomes.

In this study, we investigated (1) whether the inclusion of an internal reference could negate the negative effect of GC content bias, hence (2) whether this technology could be extended for the prenatal detection of all common aneuploidies, including aneuploidies of chromosome 13, 18, X and Y, and (3) the efficacy when the test is limited to a cohort of pregnancies predominantly in the first trimester.

Methods

Clinical samples

This study was approved by the Institutional Review Board of BGI-Shenzhen. Patients presenting at the CRIFM Clinical Research Institute of Fetal Medicine PMC, Osaka, Japan for a CVS or amniocentesis were invited to participate. Informed consent

*The authors contributed equally to this work.

Correspondence: Xiuqing Zhang, Guangdong Province Key Laboratory of Genome, BGI-Shenzhen, Shenzhen 518083, China.
E-mail: zhangxq@genomics.org.cn

(Received 28 June 2011; revised 14 September 2011; accepted 24 October 2011)

was obtained from each participant. Five milliliter maternal peripheral blood samples were obtained from each participant and collected into collection tubes containing EDTA, before the invasive procedure (either CVS or amniocentesis). Conventional karyotyping was performed in all cases on either the chorionic villus or amniotic fluid samples, and the results were considered the gold-standard of the fetal chromosomal status for further analysis.

Blood samples were treated on site by a two-step centrifugation protocol immediately after blood sampling. The samples were first centrifuged at 1600*g* for 10 min, and the resultant plasma was centrifuged again at 16000*g* for 10 min. All centrifugation was performed at 4°C and all plasma samples were stored at −80°C until further processing.

Investigators who carried out the subsequent DNA extraction, sequencing, and data analysis were blind to the karyotyping information, and results were generated automatically according to the algorithm mentioned below.

DNA library construction and sequencing

100 µl maternal plasma was used to extract DNA for library construction. Briefly, plasma DNA was used for blunt-end, "A" tailing, and adapter-ligation; 17 cycles of PCR were performed to enrich the adapter-ligated DNA fragments. After PCR purification, the libraries were analysed on an Agilent 2100 Bioanalyzer for size distribution and quantified with Real-time PCR. 12 six-base barcoded libraries were pooled and sequenced by single-end 50+8 cycles on one lane by the Illumina Hiseq™ 2000 according to the manufacturer's instructions.

Sequencing data analysis

All 50-bp sequence reads were aligned to the repeat-masked human genomic reference sequences (NCBI build 36.1) using Efficient Large-Scale Alignment of Nucleotide Databases. The reads that uniquely mapped to the reference without a mismatch were termed unique reads. To detect aneuploidies for each sample, we first determined the percentage of unique reads, also known as the FGR, that mapped to the chromosome of interest. This was determined by dividing the count of unique reads of a specific chromosome by the total number of unique reads generated by the sample in question, denoted as $\%UR_{ChrN}$, where the chromosome of interest is ChrN.

Sequencing data analysis with internal reference

To minimize this sequencing bias, we introduce a z-score method with internal reference. First, we performed a simulation analysis to identify the best internal reference for each chromosome of interest. Using samples with known karyotypes, 22 autosomes were used to serve as internal reference, one by one, to analyse the detection rate, False Positive Rate (FPR), sensitivity and specificity. For example, when evaluating different internal reference for Chr18, we have chosen 13 Trisomy 18 cases and 87 normal cases (a total of 100 cases) in the simulation analysis. According to result (see supplementary Table S1.), using Chr8 as internal reference achieves the best accuracy, and therefore Chr8 was chosen as internal reference for Chr18. The best internal reference for Chr13, 21, X and Y were identified using similar method, and the number of abnormal and normal samples used in the simulation were 3/97,17/83, 4/96 and 1/99 for Chr13, 21 X and Y, respectively.

In this study, Chr4, Chr8 and Chr14 were chosen as the internal reference chromosomes for detection of Trisomy 13, Trisomy 18 and Trisomy 21, respectively. These internal reference chromosomes share similar GC content with the corresponding autosome of interest. We then calculated the relative FGR using the internal reference (relative FGR) as $\%UR_{ChrN}/\%UR_{ChrIR}$, where ChrIR represents the internal reference chromosome for ChrN. We then calculated the z-score of the Relative FRG in the test sample by subtracting the mean $\%UR_{ChrN}/\%UR_{ChrIR}$ of a reference set of euploid pregnancies (see below for description of the reference set) from $\%UR_{ChrN}/\%UR_{ChrIR}$ of the test sample and divided by the standard deviation of $\%UR_{ChrN}/\%UR_{ChrIR}$ of the reference set. We used a z-score_{IR} of ≥3 as the threshold for the diagnosis of Trisomy in ChrN of the test sample, indicating a genomic representation greater than that of the 99.9thpercentile of the reference set for a one-tailed distribution.

To detect aneuploidy in sex chromosomes, for each sample, we first estimated the fetal gender according to $\%UR_{ChrY}$, so that appropriate reference samples of the same gender could be selected to calculate the z-score_{IR}. Cases with $\%UR_{ChrY}$ less than the threshold (0.006) were classified as female, while cases with $\%UR_{ChrY}$ greater than the threshold were classified as male. Chr7 and Chr20 were chosen as the internal control for the detection of sex chromosome aneuploidies. We then calculated the relative FGR of the sex chromosome by the following equation:

$$(\%UR_{ChrX} + \%UR_{ChrY})/\%UR_{Chr7}$$

Then, we calculated the z-score_{IR} for each test sample similar to that for the autosomes. For cases of female fetuses, those with a $\text{z-score}_{IR} \geq 3$ were classified as 47, XXX, while those with $\text{z-score}_{IR} \leq -3$ as Turner syndrome (45, XO). Cases of male fetuses with $\text{z-score}_{IR} \geq 3$ would be classified as Klinefelter's syndrome (47, XXY).

For each sample, a set of euploid samples (totally 400) with fluctuating GC content in a range of 3%, compared to the corresponding sample, were chosen as references.

Sequencing data analysis with GC correction

The data analysis was repeated following previously reported method with GC correction [7], 50 kb bin with relatively stable CVs for measuring aneuploidies in maternal plasma were chosen. Then locally weighted scatter plot smoothing (LOESS) regression was applied to correct the GC bias for each bin. After GC correction, the z-score statistic was calculated with the same equation.

Sequencing data analysis without internal reference

The data analysis was repeated using previously reported method without internal reference [3]. In brief, the FGR of ChrN, denoted as $\%UR_{ChrN}$, was calculated followed by the z-score. The z-score was determined by simply subtracting the $\%UR_{ChrN}$ of the test sample by the mean $\%UR_{ChrN}$ from the same reference sample set used to calculate the z-score_{IR} and then divided by the SD of the same reference samples. A sample with z-score of > 3 for ChrN was classified as having Trisomy of ChrN.

Final comparison

The fetal karyotype prediction from the z-score, z-score_{IR} and $\text{z-score}_{GC\ correction}$ values was compared to the conventional karyotyping results, using the latter as the gold-standard. The detection rate and false positive rate of three methods were calculated and compared.

Results

A total of 108 cases were recruited. The pregnancy and maternal characteristics were summarized in Table I. Over 89.8% (97 cases) were in the first trimester; less than 14 week of gestation. The average maternal age was 37 ± 4.3. The most common indications for CVS or amniocentesis were due to positive first trimester screening (47.2%), the presence of first trimester sonographic markers (22.2%), and the presence of other structural anomalies (28.5%). Six participants had amniocentesis, while and the remaining underwent CVS. There was a total of 32 cases with fetal aneuploidies confirmed by conventional karyotyping, including 11 cases of Trisomy 21, 10 cases of Trisomy 18, 2 case of Trisomy 13, 8 cases of Turner syndrome (45, XO) and one case of Klinefelter's syndrome (47, XXY).

An average of 7.3 million reads was obtained from each sample, and the mean number of uniquely mapped reads was 2.7 million. No case required repeated blood sampling.

The z-score method with internal reference can detect fetal aneuploidies as early as the 11th week of gestation. For the eleven cases of Trisomy 21, the average z-score$_{IR}$ was 6.60 ± 1.52 (median: 6.59; interquartile range (IQR): 6.04–7.52). For the ten cases of Trisomy 18, the average z-score$_{IR}$ was 5.50 ± 1.73 (median: 5.51; IQR: 4.21–6.74). For the two cases of Trisomy 13, the z-score$_{IR}$ was 6.06 and 4.99, respectively.

Of the 108 cases in this study, 47 cases were identified as female fetuses and the remaining 61 were male fetuses. There was complete concordance with the fetal karyotype. Among the 47 female fetuses, 8 cases were identified as Turner syndrome (45, XO) with an average z-score$_{IR}$ of −4.5 ± 1.21 (median: −3.92; IQR: −5.12 to −3.73). For the 61 cases of male fetuses, we identified one case of Klinefelter's syndrome (47, XXY), whose z-score$_{IR}$ was 3.19 (Figure 1).

Table II showed the individual z-score values for each case with aneuploidy. Using the z-score method without reference

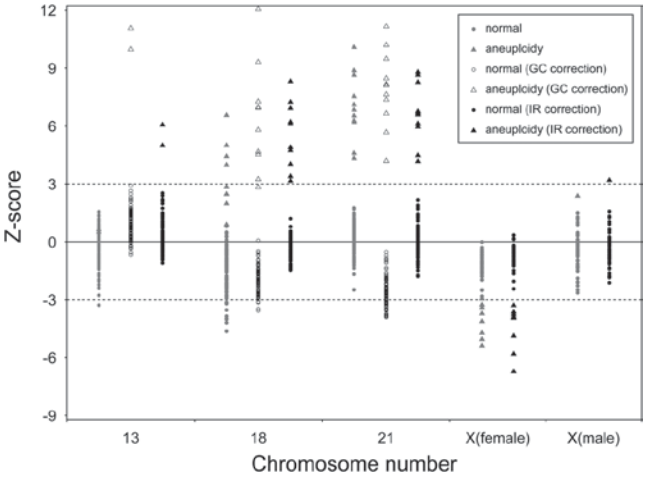

Figure 1. Comparison of the z-score values of three approaches among the 108 cases.

Table I. Clinical characteristics of the study population (n = 108)

Clinical characteristics	
Maternal age (mean, SD)	37 ± 4.3
Maternal age ≥ = 35 (%)	81 (75.0%)
Gestation at blood sampling	
Median (week)	12^{+5}
Range (week)	11^{+4} – 28^{+0}
11^{+0} to 11^{+6} week (*n*, %)	7 (6.5%)
12^{+0} to 12^{+6} week (*n*, %)	56 (51.9%)
13^{+0} to 13^{+6} week (n, %)	34 (31.5%)
14^{+0} to 14^{+6} week (*n*, %)	5 (4.6%)
15^{+0} to 19^{+6} week (*n*, %)	5 (4.6%)
20^{+0} to 28^{+0} week (*n*, %)	1 (0.9%)
Diagnostic procedures	
Chorionic villus sampling	102 (94.4%)
Amniocentesis	6 (5.6%)
Indication for diagnostic procedures	102 (94.4%)
High risk for first trimester screening	51 (47.2%)
Positive first trimester sonographic markers	24 (22.2%)
Other structural anomalies	20 (18.5%)
Maternal anxiety	12 (11.1%)
Previous T21	1 (0.9%)
Fetal karyotype	
Normal	76
Trisomy 21	11
Trisomy 18	10
Trisomy 13	2
45 XO	8
47 XXY	1

Table II. Comparison of the z-score values of three approaches among the cases with fetal aneuploidies.

Case No.	Karyotype	Gestational age (week)	z-score$_{IR}$	z-score	z-score$_{GC\ correction}$
49	Trisomy 13	12	6.06	**0.51**	11.05
103		13	4.99	**0.64**	9.97
14	Trisomy 21	12	8.79	10.08	9.49
15		13	6.76	7.11	7.64
43		13	5.96	6.53	8.17
53		16	4.17	4.33	5.67
55		11	8.27	8.64	10.20
57		12	6.59	7.54	8.47
61		13	4.48	4.59	4.20
82		13	8.65	8.87	11.15
84		13	6.67	6.84	8.11
95		12	6.12	6.28	6.65
108		12	6.13	6.21	7.37
6	Trisomy 18	12	6.12	4.43	6.94
10		12	6.92	5.00	7.27
12		12	4.89	**2.47**	4.67
16		12	3.15	**2.46**	**2.83**
18		28	8.31	6.56	12.06
74		12	3.39	**1.99**	3.24
100		17	6.21	3.99	6.97
102		13	4.03	**0.89**	4.53
105		11	4.74	**2.84**	5.79
107		12	7.25	**2.84**	9.32
11	45, XO	12	−3.64	**−2.94**	/
21		11	−3.29	−3.42	/
54		13	−5.83	−5.05	/
60		12	−3.95	−3.72	/
68		13	−3.77	−3.26	/
86		13	−6.72	−5.38	/
87		14	−4.88	−4.73	/
65		12	−3.90	−4.13	/
80	47, XXY	11	3.19	**2.36**	/

Values below the threshold are bolded.

Table III. Detection rate and false positive rate of massively parallel sequencing in the prenatal detection of fetal aneuploidies from maternal plasma.

	z-scoreIR		z-score		$z\text{-score}_{GC\ correction}$	
	Detection rate	False positive rate	Detection rate	False positive rate	Detection rate	False positive rate
Trisomy 21	100% (11/11)	0% (0/97)	100% (11/11)	0% (0/97)	100% (11/11)	0% (0/97)
Trisomy 18	100% (10/10)	0% (0/98)	40% (4/10)	0% (0/98)	90% (9/10)	0% (0/98)
Trisomy 13	100% (2/2)	0% (0/106)	0% (0/2)	0% (0/106)	100%(2/2)	0% (0/106)
45, XO	100% (8/8)	0% (0/100)	88% (7/8)	0% (0/100)	/	/
47, XXY	100% (1/1)	0% (0/107)	0% (0/1)	0% (0/107)	/	/

adjustment, the detection rate was 100% for Trisomy 21, but only 40% for Trisomy 18, 0% for Trisomy 13, 88% for Turner syndrome (45, XO), and 0% for Klinefelter's syndrome (47, XXY). The false positive rate was 0%. Using the z-score method with adjustment by GC correction, the detection rate increased to 100% for Trisomy 21, 90% for Trisomy 18, 100% for Trisomy 13, and the false positive rate was 0%, as well. By using the z-score method with adjustment by internal reference in this study, the detection rate increased to 100% for all aneuploidies, while the false positive rate remained 0%. In the 76 cases with normal karyotype, all z-score, $z\text{-score}_{IR}$ and $z\text{-score}_{GC\ correction}$ values were within the threshold limit. Table III summarized and compared the detection rate and false positive rate of these three approaches.

Discussion

Over the last 5 years, there have been major discoveries and advances in prenatal diagnosis of fetal aneuploidies from maternal plasma. Initially, success came from the study of differences between maternal and fetal nucleic acids, such as the use of differential methylation markers, or the RNA allelic ratio [8,9]. However, the widespread application of these approaches was limited by their reliance on polymorphic markers, and the complexity associated with methylation and RNA studies. The reported success of massively parallel sequencing for noninvasive prenatal diagnosis was promising, because it could be applied to all pregnant women in a population regardless of ethnicity; also working with DNA is generally easier than with RNA.

Reports using sequencing technology were focused mainly on fetal Trisomy 21, which was highly accurate with a detection rate of 100% and a specificity of 97.9–99.7% [5,6]. In this study, we attempted to develop a method that enables us to detect all common aneuploidies, including Trisomy 13, 18, 21, Turner syndrome, triple X syndrome, Klinefelter syndrome, and 47, XYY.

To minimize the sequencing bias, we developed a method introducing an internal reference.

The interesting part is that Chr4, 8 and 14 share the similar GC content with Chr13, 18 and 21, respectively, while Chr7 and 20 share the similar chromosome size, rather than GC content, with ChrX and Y, respectively. Recently, the influence of GC bias on the sensitivity of aneuploidy detection has also been reported [4]. We believe that GC content is an important contributor to the sequencing bias, but as to sex chromosomes, chromosome size variation may be another contributor to the sequencing bias. However, the relationship between chromosome size and sequencing bias is unclear.

In our study, we showed that the use of simple z-score method allowed the accurate detection of Trisomy 21, even as early as 11 weeks of gestation. However, the diagnostic accuracy was far from acceptable for other aneuploidies, being 0% for Trisomy 13 and 40% for Trisomy 18. Even with GC correction with LOESS, the detection rate for Trisomy 18 was only 90% and could not detect any sex chromosome abnormalities. However, using z-scores with an internal reference, which corrects for GC bias and sequencing efficiency, substantially improved the performance of the test. In our cohort, such corrections enabled the accurate identification of all cases with autosomal aneuploidies at a false positive rate of 0%. The identification of sex chromosome abnormalities using massively parallel sequencing has not been previously reported. Using the modified z-score, we were able to achieve a 100% detection rate for all sex chromosome abnormalities in this cohort as well. The ability to detect all common aneuploidies in a single noninvasive step – instead of Trisomy 21 only – makes this test much more attractive to pregnant women over the conventional invasive tests (CVS and amniocentesis) that carries 0.5–1% fetal loss rate [10].

The demand from pregnant women for early prenatal diagnosis is enormous, which enables early reassurance, as well as earlier and safer intervention in case of fetal abnormalities. In this study, over 85% of the cases were pregnancies in the first trimester, as early as 11 weeks of gestation. We have confirmed that this test is robust even at this early gestation. It would be interesting and clinically relevant to further investigate whether this test could be extended to even earlier gestations, such as 9–10 weeks.

Although this test achieved high accuracy in this cohort, this study was limited by its small sample size. Large studies are required to precisely estimate the detection rate and false positive rate for the detection of various chromosomal abnormalities.

In conclusion, this study has shown that massively parallel sequencing with a z-score methodology combined with internal reference is a highly accurate noninvasive test for all common aneuploidies.

Declaration of interest: The authors report no conflicts of interest.

Reference

1. Lo YM, Corbetta N, Chamberlain PF, Rai V, Sargent IL, Redman CW, Wainscoat JS. Presence of fetal DNA in maternal plasma and serum. Lancet 1997;350:485–487.
2. Finning K, Martin P, Daniels G. A clinical service in the UK to predict fetal Rh (rhesus) D blood group using free fetal DNA in maternal plasma. Ann N Y Acad Sci 2004;1022:119–123.
3. Chiu RW, Chan KC, Gao Y, Lau VY, Zheng W, Leung TY, Foo CH, et al. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. Proc Natl Acad Sci USA 2008;105:20458–20463.
4. Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR. Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. Proc Natl Acad Sci USA 2008;105:16266–16271.
5. Chiu RW, Akolekar R, Zheng YW, Leung TY, Sun H, Chan KC, Lun FM, et al. Non-invasive prenatal assessment of Trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study. BMJ 2011;342:c7401.
6. Ehrich M, Deciu C, Zwiefelhofer T, Tynan JA, Cagasan L, Tim R, Lu V, et al. Noninvasive detection of fetal Trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting. Am J Obstet Gynecol 2011;204:205.e1–205.11.
7. Eric Z. Chen, Rossa W.K. Chiu, Hao Sun, Ranjit Akolekar, Allen Chan KC, Tak Y. Leung, Peiyong Jiang, Yama W.L. Zheng, Fiona M.F. Lun, Lisa Y.S. Chan, Yongjie Jin, Attie TJIG, Elizabeth T. Lau, William

W.K. To, Wing C. Leung, Rebecca Y.K. Tang, Sidney K.C. Au-Yeung, Helena Lam, Yu Y. Kung, Xiuqing Zhang, John M.G. van Vugt, Ryoko Minekawa, Mary H.Y. Tang, Jun Wang, Cees B.M. Oudejans, Tze K. Lau, Kypros H. Nicolaides, Dennis Lo YM. Noninvasive prenatal diagnosis of fetal Trisomy 18 and Trisomy 13 by maternal plasma DNA sequencing. Plos one 2011;6:e21791.

8. Tong YK, Ding C, Chiu RW, Gerovassili A, Chim SS, Leung TY, Leung TN, et al. Noninvasive prenatal detection of fetal Trisomy 18 by epigenetic allelic ratio analysis in maternal plasma: theoretical and empirical considerations. Clin Chem 2006;52: 2194–2202.
9. Lo YM, Tsui NB, Chiu RW, Lau TK, Leung TN, Heung MM, Gerovassili A, et al. Plasma placental RNA allelic ratio permits noninvasive prenatal chromosomal aneuploidy detection. Nat Med 2007;13: 218–223.
10. Kong CW, Leung TN, Leung TY, Chan LW, Sahota DS, Fung TY, Lau TK. Risk factors for procedure-related fetal losses after mid-trimester genetic amniocentesis. Prenat Diagn 2006;26:925–930