ORIGINAL ARTICLE

# Non-invasive prenatal testing of fetal whole chromosome aneuploidy by massively parallel sequencing

Desheng Liang[1], Weigang Lv[1], Hua Wang[2], Liangpu Xu[3], Jing Liu[1], Haoxian Li[1], Liang Hu[1], Ying Peng[1] and Lingqian Wu[1]*

[1]State Key Laboratory of Medical Genetics, Xiangya School of Medicine, Central South University, Changsha, Hunan, China
[2]Women and Children's Hospital of Hunan Province, Changsha, Hunan, China
[3]Fujian Provincial Maternity and Children Health Hospital, Fuzhou, Fujian, China
*Correspondence to: Lingqian Wu. E-mail: wulingqian@sklmg.edu.cn

## ABSTRACT

**Objective** To determine whether non-invasive prenatal testing by maternal plasma DNA sequencing can uncover all fetal chromosome aneuploidies in one simple sequencing event.

**Methods** Plasma samples from 435 pregnant women at high risk for Down syndrome were collected prior to amniocentesis in three hospitals in China between March 2009 and June 2011. We sequenced the plasma DNA extracted from these samples at low coverage. We discovered that the genome representation of each of the 24 chromosomes obeyed a linear relationship to its GC content. Applying this relationship, we analysed the copy number of each of the 24 chromosomes. Full fetal karyotyping was compared with maternal plasma DNA sequencing results.

**Results** Among the 435 samples, 412 samples (94.7%) have full karyotyping and sequencing results. Sixty-seven samples containing a fetal chromosome aneuploidy, including trisomy 21, trisomy 18, trisomy 13, trisomy 9, monosomy X or others, can be accurately identified with a detection sensitivity of 100% and a detection specificity of 99.71%. Normalization of the chromosome representation values against chromosomal guanine/cytosine base content is the key issue to ensure the accuracy.

**Conclusions** Our results indicate that non-invasive detection of fetal chromosome aneuploidies for all 24 chromosomes in one single sequencing event is feasible. © 2013 John Wiley & Sons, Ltd.

*Supporting information may be found in the online version of this article.*

Funding sources: This study was supported by the grants from the National High Technology Research and Development Program of China (863 Program) (No. 2011AA02A112), the National Key Basic Research Program of China (2012CB944600) and the National Key Technology R&D Program of China (2012BAI09B05). The funder was not involved in study design, data collection, data analysis, manuscript preparation or publication decisions.
Conflicts of interest: None declared

## INTRODUCTION

Chromosome aneuploidies represent a major class of genetic defects, with an estimated incidence rate of one out of 300 live births. The common chromosome aneuploidies among live births are trisomy 21, trisomy 18, trisomy 13 and sex chromosome aneuploidies.[1] Most notably, trisomy 21 (Down syndrome) occurs to one in 800 newborns. The incidence rates for trisomy 18 and trisomy 13 among the live births are lower than that for trisomy 21, estimated at one out of 6000 and one out of 10 000, respectively, but the rates among the unborn fetuses are much higher.[2] Sex chromosome aneuploidies, including X chromosome monosomy (45, X or Turner syndrome), XXY syndrome (47, XXY or Klinefelter syndrome), XYY syndrome (47, XYY) and Triple X syndrome (47, XXX), occur at an estimated rate of one out of 1000 each among the live births.

Despite their known genetic causes, serum screening of such an important class of genetic disorders primarily relies on indirect markers.[3] Giving an estimated detection sensitivity of 90% at a false positive rate of 2% to 5%, the current most effective screening protocol for Down syndrome and trisomy 18 is based on the combination of maternal age, ultrasonographic examination of the fetus, and levels of various proteins or hormones in the maternal blood,[4] none of which is a product expressed by or derived from a gene located on chromosome 21 or chromosome 18, and hence, none of which bears direct relationship with trisomy 21 or trisomy 18.[5,6] Today, serum screening for trisomy 13 and other aneuploidies is not commonly performed. Amniocentesis or other invasive methods, including chorionic villus sampling and cordocentesis, coupling with fetal cell karyotyping, do yield definitive answers. However, they come with a risk of miscarriage rate of about 0.5%.[7,8]

The existence of cell-free fetal DNA in maternal circulation[9] and the recent advent of massive parallel sequencing technologies[10–12] have enabled non-invasive prenatal testing

 

(NIPT) of fetal chromosome aneuploidies.[13–16] This was first demonstrated on trisomy 21 samples.[13–15] The results of several recent clinical studies indicated that the detection sensitivity and specificity for trisomy 21 could reach 100% and 99%, respectively.[17–22] Detecting trisomy 18 and trisomy 13 by sequencing was proved to be more difficult than detecting trisomy 21 because the measurement coefficient of variances (CVs) for chromosome 18 and 13 were much larger than that for chromosome 21.[13,14] However, when adjusted with GC content, it was documented that trisomy 18 and trisomy 13 can be detected effectively and accurately.[20,21,23]

## MATERIALS AND METHODS

### Study design and sample collection

Four hundred thirty-five pregnant women were recruited between March 2009 and June 2011 for this study. Approvals were obtained from the institutional review boards of each recruitment site: the Women and Children's Hospital of Hunan Province, the Xiangya Hospital, both located in Changsha, Hunan, China, and the Fujian Provincial Maternity and Children Health Hospital, located in Fuzhou, Fujian, China. All participants gave informed written consent. The fetuses of the participants were screened to be high risks for Down syndrome by conventional serum screening and ultrasound scanning, and the participants were offered the option of amniocentesis or other invasive prenatal diagnostics. Five millilitres maternal peripheral blood was collected into EDTA-K2 Vacutainer tubes (Becton Dickinson, Franklin Lakes, NJ, USA) from each participant prior to amniocentesis or other invasive procedures, centrifuged according to the procedure described herein and the plasma shipped to Berry Genomics in Beijing for DNA extraction and downstream sequencing analysis. Table S1 summarized the patient demographics, including the samples obtained from different indications: 217 samples from serum screening (49.89%), 84 samples from advanced maternal age (19.31%), 67 samples from ultrasound abnormality, 4 samples from prior aneuploidy pregnancy and 64 samples with multiple indications. Except one case (11 weeks), the gestation ages of the samples were from 15 to 39 weeks. One hundred seventy-four of the participants were from 15 + 0 to 20 + 6 weeks gestation, and 260 were from 21 + 0 to 39 + 6 weeks gestation. The karyotyping was conducted in facilities located in the Women and Children's Hospital of Hunan Province, the Xiangya Hospital and the Fujian Provincial Maternity and Children Health Hospital. Table S2 summarized the karyotyping results of the fetal cells included in the study. The sequencing results were obtained independently by Berry Genomics in Beijing, and the karyotype information of all samples was kept blinded until after analysis. In the present study, Berry Genomics, which performed the sequencing analysis for free, is an NIPT and sequencing service provider in China. The authors collected the samples, conducted invasive procedures and karyotyping, unblinded the results and participated in the final data analysis.

### Maternal plasma DNA processing and sequencing

Within 8 h of collection, the maternal blood samples were centrifuged first at 1600*g* for 10 min at 4 °C. The plasma was transferred to microcentrifuge tubes, centrifuged again at 16 000*g* for 10 min at 4 °C to remove the residual cells, transferred to fresh tubes and stored at −80 °C. For each sample, plasma DNA was extracted from 1 mL of the plasma using the QIAamp Circulating Nucleic Acid kit from Qiagen (Hilden, Germany). The resulting plasma DNA was used as the input DNA to make a library for sequencing, using the modified ChIP Seq protocol as reported previously.[13] Plasma DNA libraries of 8 to 12 samples were indexed using 6 bp indexing oligos, quantitated by Kapa SYBR fast qPCR kit from Kapa Biosystems (Woburn, MA, USA), pooled and loaded into one lane in a v2 Illumina HiSeq2000 flow cell. Clustering and sequencing were conducted according to Illumina's instruction, using the single-ended 43 bp sequencing protocol.

### Data analysis

The sequences were binned for each sample according to the index and mapped to the unmasked human genome sequence (hg19). We tested several mapping software, including BWA,[24] Bowtie,[25] and SOAP2.[26] We found that the results were comparable among the three mapping algorithms. To be consistent with the previous reports,[27] the results in this study were derived from the SOAP2 mapping. There were 23 samples that failed to pass the quality control (QC) criteria. Among them, 12 samples were due to sequencing problems and the other 11 samples failed the karyotyping procedure (Supplement Figure S1). For the remaining 412 samples, the sequences mapped to each chromosome were counted, and their GC content was calculated. A reference set of 50 female diploid samples were generated from an independent sample set with normal 46, XX fetal karyotypes.

For each chromosome of each sample, the normalized chromosome representation (NCR) value in the unit of percentage was generated according to following equation: NCR = count of the sequences uniquely mapped to the chromosome of interest / total count of the sequences uniquely mapped to all the autosomal chromosomes.[18] For any given chromosome except chromosome 14 and chromosome Y of the 50 reference samples, we plotted the NCR values against the GC content and calculated the slope by simple linear regression. We used the slope to further normalize the NCR of the 412 QC-passed samples: $NCR_{gc} = NCR - (GC - GC_{average\_ref}) / Slope_{ref}$, where $NCR_{gc}$ is the GC corrected NCR value, NCR is the uncorrected value, GC is the chromosome GC content of the test sample, $GC_{average\_ref}$ and $Slope_{ref}$ are the average chromosome GC content and the slope of the reference samples, respectively. We calculated the mean ($\mu_{ref}$) and the standard deviation ($\delta_{ref}$) of the $NCR_{gc}$ for the reference dataset. The *z* score for chromosome *i* (*i* can be any one of the 24 chromosomes except chromosome 14 and chromosome Y) of sample *j* was calculated according to following equation: $z_{ij} = (x_{ij} - \mu_{ref}) / \delta_{ref}$, where $x_{ij}$ is the $NCR_{gc}$ value for chromosome *i* of sample *j*. For chromosome 14 and chromosome Y, the *z* scores were calculated without GC correction. For any given male fetus sample *j*, the percentage of the fetal DNA was estimated based on the following equation: $2 * (1 - x_j)$, where $x_j$ is the ratio between the $NCR_{gc}$ for chromosome X of sample *j* with a male fetus and the average $NCR_{gc}$ for chromosome X of the reference samples with female fetus.[28]

## RESULTS

### Maternal plasma DNA sequencing

We sequenced the plasma DNA extracted from the maternal peripheral blood samples of 435 pregnant women who underwent invasive prenatal diagnostics and whose fetal cells were karyotyped. We indexed and pooled 8 to 12 samples into one lane of an eight-lane flow cell. To test the lane-to-lane and run-to-run variations, we put in each lane two control samples, whose fetal cell karyotypes were known to us: one diploid sample (D21) and one trisomy 21 sample (T21). We generated 543 data entries, among which 108 data entries were generated from the two control samples, 54 data points each, derived from one single DNA library. On average, we collected 8.3M 36 bp sequence tags per sample. The sequences were mapped to the unmasked human genome (hg19). For each chromosome, we calculated the NCR value by dividing the number of sequences uniquely mapped to the chromosome of interest by the total number of sequences uniquely mapped to all autosomal chromosomes.[18] Analysing the control samples, we noticed a wide range of the NCR values among the data points generated between runs and between lanes of a single run. Searching for the sources of the variance, we plotted the NCR values against the GC content of the sequences that were uniquely mapped to chromosome 21. We observed a linear relationship between the NCR values and the GC content (Figure 1A). We further observed that similar linear relationship existed for all the 24 chromosomes except chromosome 14, each with a different slope and an intercept. To offset the GC bias, we corrected the NCR values against the GC content based on the linear relationship (see Methods). After such GC correction, the CV for each chromosome except chromosomes 12, 21 and Y was significantly lower than the CV without GC correction (Figure 1B). In particular, the CVs for chromosomes 3, 4, 5, 6, 13, 16, 17, 18, 19, 20, 22 and X were reduced by 57% to 82% (Supplement Table S3).

### Non-invasive detection of fetal trisomy 21

After removing the replicated data points for the control samples, we analysed 412 non-duplicate data points representing 412 samples that passed the quality standards (Supplement Figure S1). Using a 50 female sample dataset as the reference and applying the above GC correction, we calculated the $z$ scores for each autosomal chromosome.[13] Using a $z$-score value of 3 as the cutoff value, we identified 40 T21 candidate samples (Figure 2B), which were clearly separated from the D21 samples on the NCR vs. GC content scatter plot (Figure 2A). These 40 T21 samples identified by sequencing correctly matched the 40 T21 samples identified by karyotyping, giving 100% detection sensitivity and 100% specificity (Table 1). These 40 T21 samples included one case of mosaicism (47,XY,+21[26]/46,XY[24]) and two cases of Robertsonian translocation, confirming the previous report that the trisomy 21 samples caused by Robertsonian translocation could be detected by this approach.[29] The lowest $z$ score for the T21 samples was 3.89, and the highest $z$ score for the D21 samples was 2.24.

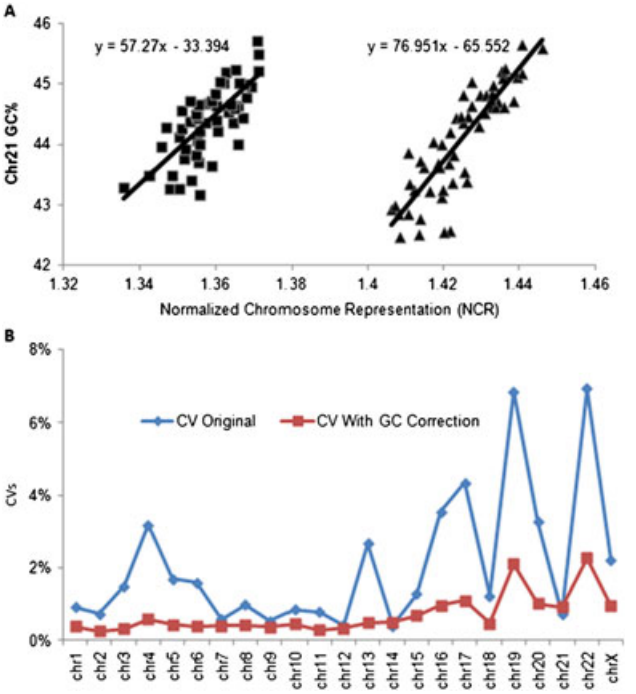


Figure 1 (A) Linear relationship between the normalized chromosome representation (NCR) values and the GC content. Linear relationship exists between the NCR values and the GC content of sequence tags uniquely mapped to chromosome 21 for the replicated control samples. Each square (D21 sample) or triangle (T21 sample) represents one data point. (B) The reduction of measurement coefficient of variances (CVs) using GC correction. GC correction reduces the measurement CVs for all the 24 chromosomes except chromosomes 12, 21 and Y

### Non-invasive detection of fetal trisomy 18

We applied the same approach to chromosome 18. We noticed that the measurement CV for chromosome 18 (0.47%) was less than that for chromosome 21 (0.92%). The $z$-score cutoff value of 3 for chromosome 21 corresponded to fetal DNA percentage of 5.52%. At this level of fetal DNA, we calculated the $z$ score for a trisomy 18 sample to be 5.91. We therefore set the $z$-score cutoff value for trisomy 18 to 5.91 (Table S3).

Analysing the sequencing data, we identified 14 samples that had $z$ scores above 5.91 for chromosome 18 (Figure 2D). We found that these 14 samples correctly identified all the 14 T18 cases determined by karyotyping, including one mosaic sample (46,XX[96]/47,XX,+18[3]), giving a sensitivity of 100% (14/14) and a specificity of 100% (Table 1). The T18 and non-T18 samples were also clearly separated from each other on the NCR vs. GC content plot (Figure 2C). The $z$ scores for the T18 samples ranged from 6.46 to 34.33, whereas the highest $z$ score for the non-T18 samples was 3.72.

### Non-invasive detection of fetal trisomy 13

There were four T13 samples in the study determined by karyotyping, including a twin sample with one of the twin being trisomy 13. As in the case of trisomy 18, we calculated the $z$-score cutoff value for trisomy 13 to be 5.72, based on 5.52% fetal DNA (Table S3). Using 5.72 as the $z$-score cutoff value, sequencing identified five T13 candidates, four of which correctly matched all the four T13 samples by karyotyping,

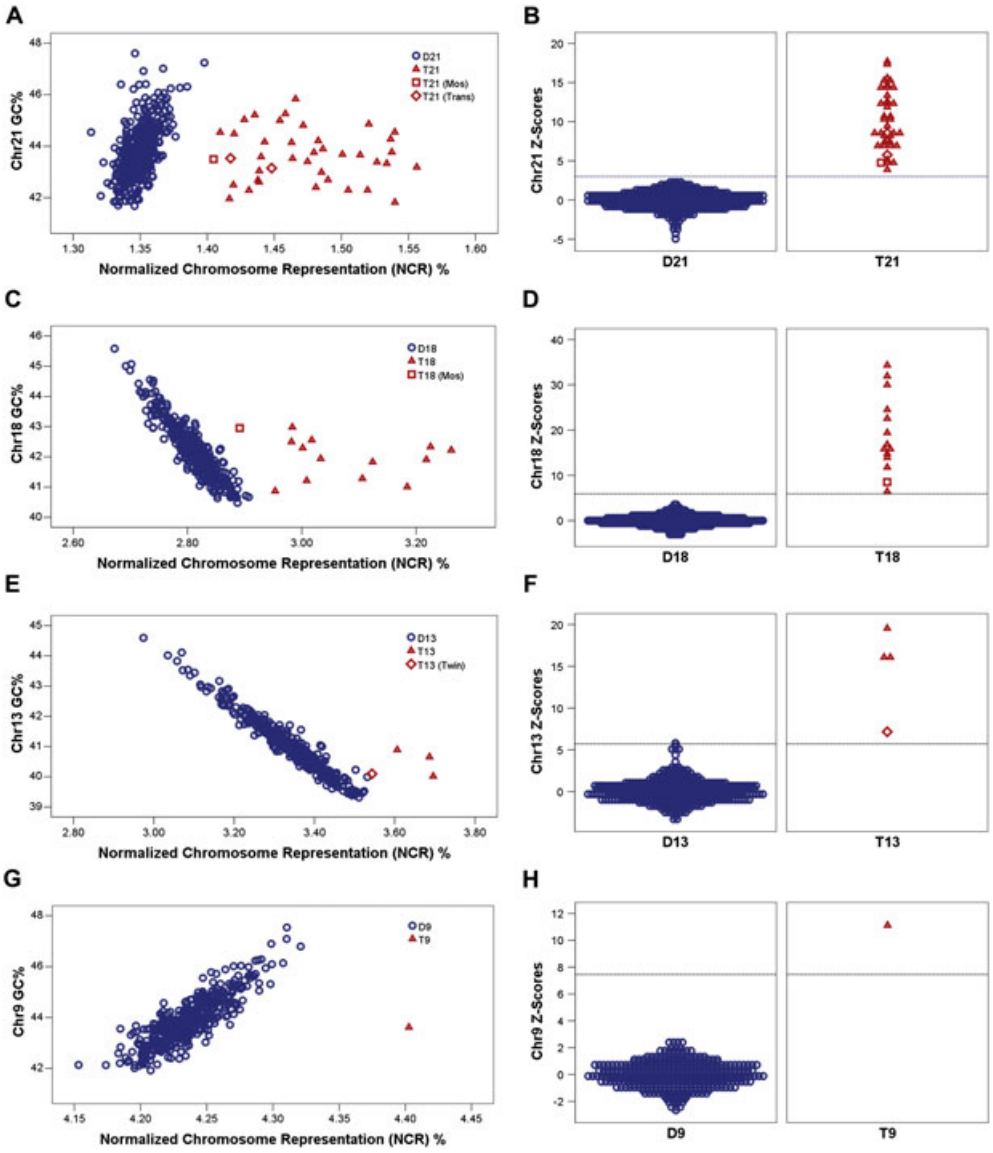


Figure 2 (A) Non-invasive detection of fetal trisomy 21. The normalized chromosome representation (NCR) values are plotted against the GC content of the sequences that were uniquely mapped to chromosome 21. The T21 samples were depicted by the triangles, the diamonds (Robertsonian translocation) and the square (mosaicism), whereas the D21 samples were depicted by the circles. (B) The *z* scores of the D21 (left) and T21 samples (right). The dot line represents the cutoff *z* score of 3. (C) Non-invasive detection of fetal trisomy 18. The T18 samples were depicted by the triangles and the square (mosaicism), whereas the D18 samples were depicted by the circles. (D) The *z* scores of the D18 (left) and T18 samples (right). The dot line represents the cutoff *z* score of 5.91. (E) Non-invasive detection of fetal trisomy 13. The T13 samples were depicted by the triangles and the diamond (the twin sample), whereas the D13 samples were depicted by the circles. (F) The *z* scores of the D13 (left) and T13 samples (right). The dot line represents the cutoff *z* score of 5.72. (G) Non-invasive detection of fetal trisomy 9. The trisomy 9 sample was depicted by the triangle, whereas the D9 samples were depicted by the circles. (H) The *z* scores of the D9 (left) and the T9 sample (right). The dot line represents the cutoff *z* score of 7.45

achieving a detection sensitivity of 100% and a specificity of 99.75% (407/408) (Table 1). The *z* scores for the T13 samples ranged from 7.17 to 19.54, with the twin sample having the lowest *z* score (Figure 2E and F). The false positive T13 sample, a true 45, X sample, had a *z* score of 5.81.

### Non-invasive detection of other autosomal chromosome aneuploidy

The aforementioned principles can be applied to other autosomal chromosomes. Among the 412 samples in the study, there was one sample that had been karyotyped as trisomy 9. Using a *z*-score cutoff value of 7.45 for chromosome 9 (corresponding to 5.52% fetal DNA, Table S3), sequencing correctly identified the trisomy 9 sample without generating any false positive samples (Figure 2G). The *z* score for the T9 sample was 11.11, whereas the *z* scores for the rest of 412 samples ranged from −2.66 to 2.45 (Figure 2H).

### Non-invasive detection of fetal sex chromosome aneuploidies

Similarly, we calculated the NCR values and *z* scores for sex chromosomes. Using a *z*-score cutoff value of 3 for Y

Table 1 **Summary of the aneuploidies detected in this study**

| Syndrome | Detection sensitivity | Detection specificity | PPV | NPV |
|---|---|---|---|---|
| Trisomy 21 | 100% (40/40) | 100% (372/372) | 100% (40/40) | 100% (372/372) |
| Trisomy 18 | 100% (14/14) | 100% (398/398) | 100% (14/14) | 100% (398/398) |
| Trisomy 13[a] | 100% (4/4) | 99.75% (407/408) | 80% (4/5) | 100% (407/407) |
| Trisomy 9 | 100% (1/1) | 100% (411/411) | 100% (1/1) | 100% (411/411) |
| 45, X[b] | 100% (5/5) | 99.75% (406/407) | 83.3% (5/6) | 100% (405/405) |
| XXX | 100% (1/1) | 100% (411/411) | 100% (1/1) | 100% (411/411) |
| XXY | 100% (1/1) | 100% (411/411) | 100% (1/1) | 100% (411/411) |
| XYY | 100% (1/1) | 100% (411/411) | 100%(1/1) | 100% (411/411) |
| Combined detection | 100% (67/67) | 99.71% (344/345) | 98.53% (67/68) | 100% (343/343) |

PPV, positive predictive value; NPV, negative predictive value.
[a]Including a twin sample.
[b]Not including three mosaic 45, X samples.

chromosome, we accurately separated female samples from male samples, with 100% sensitivity and 100% specificity. Using a *z*-score cutoff value of −2.91 for X chromosome (corresponding to 5.52% fetal DNA, Table S3), we correctly called all the 211 female samples, but misclassified two male samples as female samples, yielding a specificity of 99.51%. Considering the copy numbers of sex chromosomes were proportional to the percentage of fetal DNA in the maternal plasma, we figured that the *z* scores of X and Y chromosomes should have a linear relationship with each other for the male samples, accompanying measurement errors and random sampling errors. We plotted the X chromosome *z* scores against Y chromosome *z* scores, and indeed, we found that the *z* scores of sex chromosomes were correlated (Figure 3). For female samples, we expected the *z* scores of sex chromosomes to centre around the origin (0,0) on the plot: for X chromosome, the *z* scores should range from −2.91 to +2.91; for Y chromosome, the *z* scores should be less than 3. The results were as we expected (Figure 3). Those samples that were outliers to these two categories were potential sex chromosome aneuploidy samples. We identified seven Turner syndrome (45, X) candidates that had X chromosome *z* scores lower than −2.91 and Y chromosome *z* scores less than 3, among which five samples correctly matched all the five Turner syndrome samples in the study determined by karyotyping. The seven Turner syndrome candidates by sequencing also correctly identified one mosaic Turner syndrome sample by karyotyping: 45,X [88]/46,XY[13]. The seventh Turner syndrome candidate was a false positive sample of a normal karyotype but with abnormal *z* scores for several chromosomes. Sequencing also identified one XXY sample, one XYY sample and one XXX sample (Figure 3). These results correctly matched the karyotyping results. Sequencing missed two mosaic Turner syndrome samples identified by karyotyping: 45,X[9]/46,XX [41] and 45,X[48]/46,XX[15]. Excluding the mosaic samples, the aforementioned results corresponded to a sensitivity of 100% (8/8) and a specificity of 99.75% (403/404) for sex chromosome aneuploidies (Table 1).

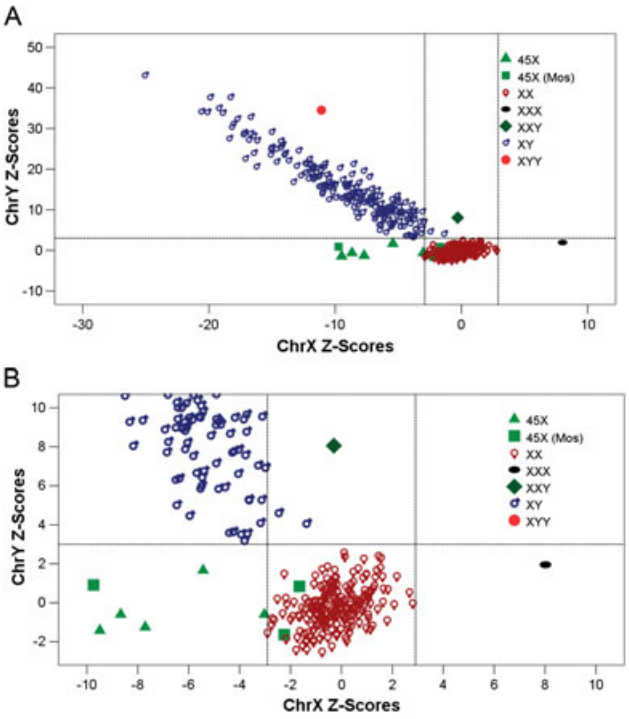


Figure 3 (A) Non-invasive detection of fetal sex chromosome aneuploidy. The *z* scores of X chromosome (*x* axis) and Y chromosome (*y* axis) were plotted to reveal the male (blue male symbols), the female (red female symbols), the XYY (solid red dot), the XXY (green diamond), the 45, X (green triangles and green squares) and the XXX sample (black oval). The green squares are the mosaic 45, X samples. (B) The enlarged version of Figure 3A

## DISCUSSION

Non-invasive prenatal testing of fetal chromosome aneuploidies by sequencing plasma DNA of maternal peripheral blood has been reported previously.[13–23,27–29] These studies were mostly restricted to three autosomes. Using a local GC correction algorithm that was deduced from an empirical dataset by regression, Chen *et al.* recently advanced

the detection to cover trisomy 13 and trisomy 18 cases,[27] despite missing three out of 37 true T18 cases. A follow-up study by Palomaki *et al.* confirmed the effectiveness of the GC correction algorithm in detecting trisomy 13 and trisomy 18 samples.[20]

This study attempted to detect fetal aneuploidies of all the chromosomes although the clinical cases were mainly of the five most clinically relevant chromosomes: chromosomes 21, 18, 13, X and Y. We detected 67 aneuploidy samples covering six chromosomes among the 412 plasma DNA extracted from maternal peripheral blood samples (Table 1). We encountered two false positive cases, one of which was T13 false positive but also a true 45, X positive sample. Considering the six chromosomes as a whole, we detected all the 67 aneuploidies with one false positive, corresponding to a detection sensitivity of 100% (67/67) and a specificity of 99.71% (344/345). These results indicate that with appropriate GC correction algorithm, detecting aneuploidies of all the chromosomes in one single sequencing event is achievable.

The GC correction also improved detection sensitivity substantially. At the early phase of data analysis, we applied the standard *z*-score calculation algorithm developed by Chiu *et al.*,[13] which was without GC correction, and found that we could detected 40 out of the 40 trisomy 21 samples with 2 false positive cases, 13 out of the 14 trisomy 18 samples without false positive sample and 3 out of the 4 trisomy 13 samples without false positive sample. As reported previously,[14,27] NIPT requires a demanding CV for any given chromosome. The GC correction reported here dramatically reduces the CVs for most of the 24 chromosomes. It therefore increases the detection sensitivity, which also depends on the fetal DNA percentage. When we surveyed the fetal DNA percentage of the 201 male samples in this study (Figure S2), we found that the maternal plasma of 199 samples contained more than 5.52% fetal DNA, which was used to set the *z*-score cutoff values in this study. This means that 99% of the samples contain a sufficient amount of plasma DNA for trisomy 21, trisomy 18, trisomy 13, trisomy 9 and sex chromosome aneuploidies to be detected, at the *z*-score cutoff values proposed in this study. In other words, the detection sensitivity for these chromosomes should be more than 99%, a breakthrough comparing with the 90% detection sensitivity for the current screening protocol. Despite lack of data points for the samples collected during the first trimester, we argue that this sensitivity and specificity should be equally applicable to the samples collected at the late stage of the first trimester because the percentage of the fetal DNA was reported to be almost constant between late stage of the first trimester and second trimester (20, 30).

This study also demonstrated the feasibility of clinically implementing NIPT by sequencing. This study employed a 12-plex protocol, where up to 12 samples were indexed and pooled into one sequencing lane in an Illumina v2 flow cell, drastically reducing the sequencing cost to an affordable level, while improving the throughput to handle 30 000 samples annually on one single Illumina HiSeq2000 instrument. In practice, we found that the simple GC correction algorithm presented earlier permits the sequencing analysis be applied to the situations where the sequencing reagents and the sequencing instruments are upgraded. To achieve this, the reference set and the new GC correction parameters must be calculated for each run. As sequencing throughput continues to increase and the cost continues to drop, the cost of the test will certainly become lower.

## CONCLUSION

We discovered that the genome representation of each of the 24 chromosomes obeyed a linear relationship to its GC content. Applying this relationship, in a single sequencing event, we could accurately identify the samples containing a fetal chromosome aneuploidy, including trisomy 21, trisomy 18, trisomy 13, trisomy 9, monosomy X and others, with a detection sensitivity of 100% and a detection specificity of 99.71%. We believe that the sequencing-based NIPT is promising in the future to be implemented in the clinical setting.

## ACKNOWLEDGEMENTS

We thank the team at Berry Genomics for conducting the sequencing and the analysis. We are grateful to all pregnant women for their participation in this study. We thank members of the hospitals for their assistance in collecting the plasma samples and performing karyotyping.

### WHAT'S ALREADY KNOWN ABOUT THIS TOPIC?

- The existence of cell-free fetal DNA in maternal circulation and the advent of massive parallel sequencing (MPS) technologies have enabled non-invasive prenatal testing (NIPT) of fetal chromosome aneuploidies.

### WHAT DOES THIS STUDY ADD?

- We discovered that the genome representation of each of the 24 chromosomes obeyed a linear relationship to its GC content. Applying this relationship, we analysed the copy number of each of the 24 chromosomes from 435 clinical samples, including a case of trisomy 9, with 100% detection sensitivity and 99.71% specificity. The sequencing was conducted using a 12-plex protocol, demonstrating that the MPS-based NIPT is scalable.

## REFERENCES

1. Driscoll DA, Gross S. Clinical practice. Prenatal screening for aneuploidy. N Engl J Med 2009;360:2556–62.
2. Griffiths A, Miller J, Suzuki D, *et al.* Chromosome mutation II: Changes in chromosome number. An Introduction to Genetic Analysis (7th edn). New York: W. H. Freeman, 2000.
3. Wald NJ, Watt HC, Hackshaw AK. Integrated screening for Down's syndrome on the basis of tests performed during the first and second trimesters. N Engl J Med 1999;341:461–7.
4. Malone FD, Canick JA, Ball RH, *et al.* First-trimester or second-trimester screening, or both, for Down's syndrome. N Engl J Med 2005;353:2001–11.

5. Palomaki GE, Haddow JE, Knight GJ, *et al.* Risk-based prenatal screening for trisomy 18 using alpha-fetoprotein, unconjugated oestriol and human chorionic gonadotropin. Prenat Diagn 1995;15:713–23.
6. Crandall BF, Hanson FW, Keener S, *et al.* Maternal serum screening for alpha-fetoprotein, unconjugated estriol, and human chorionic gonadotropin between 11 and 15 weeks of pregnancy to detect fetal chromosome abnormalities. Am J Obstet Gynecol 1993;168:1864–7; discussion 1867-9.
7. Odibo AO, Dicke JM, Gray DL, *et al.* Evaluating the rate and risk factors for fetal loss after chorionic villus sampling. Obstet Gynecol 2008;112:813–9.
8. Odibo AO, Gray DL, Dicke JM, *et al.* Revisiting the fetal loss rate after second-trimester genetic amniocentesis: a single center's 16-year experience. Obstet Gynecol 2008;111:589–95.
9. Lo YM, Corbetta N, Chamberlain PF, *et al.* Presence of fetal DNA in maternal plasma and serum. Lancet 1997;350:485–7.
10. Bentley DR, Balasubramanian S, Swerdlow HP, *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. Nature 2008;456:53–9.
11. Mardis ER. A decade's perspective on DNA sequencing technology. Nature 2011;470:198–203.
12. Metzker ML. Sequencing technologies – the next generation. Nat Rev Genet 2010;11:31–46.
13. Chiu RW, Chan KC, Gao 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 2008;105:20458–63.
14. Fan HC, Blumenfeld YJ, Chitkara U, *et al.* Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. Proc Natl Acad Sci USA 2008;105:16266–71.
15. Chiu RW, Sun H, Akolekar R, *et al.* Maternal plasma DNA analysis with massively parallel sequencing by ligation for noninvasive prenatal diagnosis of trisomy 21. Clin Chem 2010;56:459–63.
16. Lo YM. The quest for accurate measurement of fetal DNA in maternal plasma. Clin Chem 2011;57:522–3.
17. Chiu RW, Akolekar R, Zheng YW, *et al.* Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study. BMJ 2011;342:c7401.
18. Ehrich M, Deciu C, Zwiefelhofer T, *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 e201–11.
19. Sehnert AJ, Rhees B, Comstock D, *et al.* Optimal detection of fetal chromosomal abnormalities by massively parallel DNA sequencing of cell-free fetal DNA from maternal blood. Clin Chem 2011;57:1042–9.
20. Palomaki GE, Deciu C, Kloza EM, *et al.* DNA sequencing of maternal plasma reliably identifies trisomy 18 and trisomy 13 as well as Down syndrome: an international collaborative study. Genet Med 2012;14:296–305.
21. Palomaki GE, Kloza EM, Lambert-Messerlian GM, *et al.* DNA sequencing of maternal plasma to detect Down syndrome: an international clinical validation study. Genet Med 2011;13:913–20.
22. Sparks AB, Wang ET, Struble CA, *et al.* Selective analysis of cell-free DNA in maternal blood for evaluation of fetal trisomy. Prenat Diagn 2012;32:3–9.
23. Bianchi DW, Platt LD, Goldberg JD, *et al.* Genome-wide fetal aneuploidy detection by maternal plasma DNA sequencing. Obstet Gynecol 2012;119:890–901.
24. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009;25:1754–60.
25. Langmead B, Trapnell C, Pop M, *et al.* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009;10:R25.
26. Li R, Yu C, Li Y, *et al.* SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 2009;25:1966–7.
27. Chen EZ, Chiu RW, Sun H, *et al.* Noninvasive prenatal diagnosis of fetal trisomy 18 and trisomy 13 by maternal plasma DNA sequencing. PLoS One 2011;6:e21791.
28. Fan HC, Quake SR. Sensitivity of noninvasive prenatal detection of fetal aneuploidy from maternal plasma using shotgun sequencing is limited only by counting statistics. PLoS One 2010;5:e10439.
29. Lun FM, Jin YY, Sun H, *et al.* Noninvasive prenatal diagnosis of a case of Down syndrome due to Robertsonian translocation by massively parallel sequencing of maternal plasma DNA. Clin Chem 2011;57:917–9.