Simultaneous detection of fetal aneuploidies and single gene diseases by a novel method of noninvasive prenatal testing: Targeted And Genome-wide simultaneous sequencing (TAGs-seq)

Yujing Wu¹, Lin Yang², Zhiyang Hu³, Haiping Zhang⁴, Dandan Pu⁴, Huijuan Yan⁵, Sijia Zhang⁵, Hui Jiang⁴, Qiang Liu⁶, Yuying Yuan⁶, Yan Zhang⁵, Fang Chen⁴, Yanping Lu⁷, Silin Pan⁸, Linhua Lin³, and Ya Gao²

¹BGI Education Center, University of Chinese Academy of Sciences
²BGI-Shenzhen
³Department of Obstetrics, Shenzhen People's Hospital, the Second Clinical Medical School of Jinan University, Shenzhen, 518020, China
⁴BGI-Shenzhen, Shenzhen, 518083, China
⁵MGI, BGI-Shenzhen, Shenzhen 518083, China
⁶Clinical laboratory of BGI Health, BGI-Shenzhen, Shenzhen 518083, China
⁷Department of Obstetrics and Gynecology, Chinese PLA General Hospital, Beijing, 100853, China

⁸ Qingdao Women and Children's Hospital, Qingdao University, Qingdao, China.

May 5, 2020

Abstract

Objective: To simultaneously detect fetal an euploidies and single gene diseases using a novel noninvasive prenatal testing (NIPT) method called the Targeted And Genome-wide simultaneous sequencing (TAGs-seq). Design: Comparison of TAGs-seq NIPT results with conventional NIPT and diagnostic results. Setting: Shenzhen People's Hospital and Chinese PLA General Hospital Population or Sample: 26 normal pregnancies, 7 pregnancies with fetal aneuploidies, 7 pregnancies with fetal achondroplasia (ACH) or thanatophoric dysplasia (TD), 18 pregnancies with ACH/TD-like ultrasound findings, and 10 pregnancies with fetal risk of beta-thalassemia. Methods: Plasma cfDNA was amplified by TAGs-seq to simultaneously obtain the whole-genome sequence of $0.1-3\times$ depth and reads on target genes of $>1000\times$ depth. The whole-genome sequence was analyzed for fetal aneuploidy risk using a binary hypothesis T-score, and the reads on target genes were analyzed for single gene diseases by calculating minor allelic frequency of loci on FGFR3 and HBB. Main Outcome Measures: Concordance between the TAGs-seq NIPT, conventional NIPT and diagnostic results. Results: Consistent to conventional NIPT and diagnostic results, all cases of fetal aneuploidies and fetal ACH/TD were correctly identified from 58 pregnancies by TAGs-seq NIPT with high sensitivities and specificities. Two cases of paternal mutations of beta-thalassemia were correctly identified by TAGs-seq NIPT from 10 pregnancies, although one false-negative result was obtained. Conclusions: The TAGs-seq assay demonstrated a good potential to simultaneously detect fetal aneuploidies and single gene diseases as a novel NIPT method.

Tweetable abstract of 110 characters

Simultaneous detection of fetal chromosome aneuploidies and single gene diseases by TAGs-seq based NIPT

Introduction

Since its discovery in maternal blood, cell-free DNA (cfDNA) originated from placenta trophoblasts was soon used for fetal disease testing, and rapidly adopted in clinical practice known as the noninvasive prenatal testing (NIPT)[1]. NIPT usually analyzes cfDNA in maternal blood by one of following methods: whole-genome massively parallel sequencing (MPS) [2, 3], chromosome-selective region sequencing [4], and single nucleotide polymorphism-based sequencing (SNPs) [5, 6]. With higher sensitivity and specificity than conventional biochemical screening, NIPT was initially employed to detect fetal trisomy 21, 18 and 13 [7, 8]. Not long after that, cfDNA was used to screen for sex chromosome aneuploidy and chromosome copy number variants [9-11]. Recently, several studies detected single gene diseases in fetus by analyzing cfDNA using quantitative polymerase chain reaction (qPCR), digital PCR or haplotype-assisted approaches [12-17]. A major issue of detecting single gene diseases with cfDNA is that it exploits different experimental procedures as detecting chromosome abnormality, and thus it is not possible to detect chromosome abnormality and single gene diseases in one experiment. This may top-up the cost of NIPT and increase the difficulty to the laboratories providing NIPT tests.

We intended to develop a novel NIPT method with improved procedures of library construction and sequencing strategy, so that fetal chromosome aneuploidy and single gene diseases can be detected simultaneously. Herein this study, we reported the proof-of-concept validation of simultaneous detection of fetal trisomy 21 (T21) trisomy 18 (T18), trisomy 13 (T13), achondroplasia (ACH), thanatophoric dysplasia (TD), and beta-thalassemia by using a Targeted And Genome-wide simultaneous sequencing (TAGs-seq) method.

Methods

Patients

Pregnant women participating in this study were recruited through collaborative hospitals. Each hospital was involved in patient recruitment, routine prenatal healthcare, prenatal diagnosis of disease, and follow-up of outcomes. Institutional review board approval was obtained from each hospital, and written informed consent was obtained from each participant before collecting samples. To be included in the study, pregnant women had to be singleton pregnancy with gestational age of 12 weeks or larger at the time of samples collection. Participants were recruited if they have pregnancy with confirmed fetal T21, T18 T13, ACH, or TD. As an internal control, pregnant women who had no *FGFR3* mutations, yet with abnormal ultrasound suspicious to ACH/TD, or previous history of an ACH/TD pregnancy were also recruited. Pregnant women and spouses who were carriers of beta-thalassemia were also recruited to evaluate the fetal risk of inheriting paternal mutations. Women with normal singleton pregnancy were recruited as healthy controls.

Sample collection and processing

Five milliliter of blood sample was taken from each pregnant woman using an ethylene diamine tetraacetic acid (EDTA) containing tube. Plasma was extracted from blood samples by a two-step centrifugation procedure at 4°C within eight hours after blood collection. Briefly, blood sample was initially centrifuged at 1600 g for 10 minutes at 4°C, and then supernatant was transferred to a new Eppendorf tube for a second centrifugation at 16000 g for 10 minutes at 4°C. Plasma and buffy coat of each sample were then separately barcoded, and delivered with dry ice to BGI-Shenzhen for laboratory analysis. CfDNA was extracted from maternal plasma using an QIAamp Circulating Nucleic Acid Kit (Qiagen) following the manufacturer's instructions. As a quality control, only cfDNA with peak size of around 167 base pairs and yield of more than 10ng was used for further experiments.

Sequencing for conventional NIPT

MGIEasy cell-free DNA Library Prep Set (MGI) was used to prepare cfDNA library for conventional NIPT following the manufacturer's instructions. Briefly, cfDNA was end-repaired and ligated with adapters containing 10bp barcode sequences (Table S1) for multiplex sequencing. The ligated products were subjected to 12 cycles of PCR using the Kapa HIFI hotstart ready master mix (Kapa Biosystems). PCR products were then quantitated using the dsDNA HS Assay kit 2.0 (Invitrogen) by Qubit2.0 and Agilent High Sensitivity DNA Kit (Agilent Technologies) with a 2100 Bioanalyzer. PCR products were then normalized and processed for

circularization [18] (Figure S1). Briefly, PCR products were heat-denatured at 95 degC for 3 minutes to make a single strand DNA circle (ssDNA circle), which were then mixed reagents of MGIEasyTM DNA Library Prep Kit (MGI) and incubated at 37 degC for 30 minutes to complete the circularization. The resulting ssDNA circle was then used to generate DNA nanoballs (DNBs) by rolling circle amplification (RCA) [19]. After RCA and the formation of DNBs, the final product was measured by Qubit using the ssDNA HS Assay kit (Invitrogen), and loaded on a DNBSEQ-500 platform (MGI) for sequencing [20] following the manufacturer's instructions.

TAGs-seq

TAGs-seq used a proprietary library construction and sequencing strategy to simultaneously amplify the whole-genome region of 0.1-3x depth and target genes of >1000x depth. CfDNA was firstly processed with end-repair and dA tailing using T4 polynucleotide kinase (Enzymatic), Klenow polymerase (Enzymatic), and Bst polymerase (NEB), following by ligation with DNBSEQ-500 adapters (Table S1). The ligated products were subjected to two rounds of nested PCR, each round of 10 to 20 cycles, using a Kapa HIFI hotstart ready master mix kit (Kapa Biosystems) for the amplification of whole genome and gene specific regions. The first round of PCR (PCR1) uses gene-specific primers (GSP1 pool) and a primer complementary to universal ligated adapter (UP1, blue) to amplify target genes, meanwhile cfDNA fragments of whole-genome region are liner-amplified using UP1. PCR conditions were 2 minutes of denaturing at 98, followed by 10 cycles at 98 for 10 seconds, 60 for 2 minutes, 72 for 30 seconds; and 72 for 5 minutes. At the second round of PCR, the universal primer 1 (UP1), universal primer 2 (UP2), and a set of 3' nested gene-specific primers (GSP2 pool) were used for amplification. PCR conditions were 2 minutes of denaturing at 98, followed by 10 cycles at 98 for 10 seconds, 60 for 2 minutes, 72 for 30 seconds; and 72 for 5 minutes. The PCR product contained the mixture of genome-wide and targeted amplicons.

The amplification libraries were then normalized and processed for circularization and sequencing on a DNBSEQ-500 platform, following the same procedures as described above. After sequencing, the wholegenome region with low sequencing depth was used to analyze chromosome aneuploidy, while the chromosome selective regions with an ultra-high sequencing depth were used to analyze singe gene disorders (Figure 1). All TAGs-seq sequencing data and 16 NIPT sequencing data were uploaded to China National Gene Bank Nucleotide Sequence Archive (CNSA) upon the request of accession code. The remaining 50 NIPT sequencing data were not uploaded due to lack of patient approval.

Primer design

Nested gene-specific primers used in TAGs-seq were designed to amplify 18 reported mutations on FGFR3 responsible for ACH and TD, and 26 reported mutations on HBB responsible for beta-thalassemia, respectively (Table S2 and Table S3). Primer3 (v. 0.4.0) and BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) were used to design the primers with the length of 20 to 30 nucleotides and annealing temperature of 58 to 62. 3'-terminals of primers were designed to avoid single nucleotide variants of minor allelic frequency higher than 0.5%. The primers were obtained from the Invitrogen.

Data analysis

Data analysis was conducted in blind of the diagnostic results. All data were sequenced with PE50 or PE100. Paired-end (PE) reads were generated with a read length of 50 bp or 100 bp using DNBSEQ-500. The primary raw reads were trimmed to remove adapter and primer sequences. The trimmed reads were aligned to the human hg19 reference genome using the BWA tool (version 0.7.15). PE reads of low quality were removed from analysis, including shorter than 50bp, longer than 500bp, poor sequencing quality, and non-unique alignments.

The aligned reads were then divided into two parts: the reads mapping to the target genes using for genotyping analysis, and the reads of whole-genome region using for chromosome analysis. For the reads mapping to the target genes, the minor allelic frequency (MAF) of loci known to cause ACH, TD, and beta-thalassemia was calculated. A paternal point mutation was determined to exist when the MAF was between 2% to 20%. MAF higher than 20% was determined as the presence of maternal point mutation. iTools (version: 0.19) was used to call the sequencing depth and allelic frequency of detected mutations. Chromosome analysis using the genome-wide reads followed the method described previously [21]. Briefly, the GC bias introduced by the sequencing process was corrected, and the k -mer (k refers to the length of the sequencing reads) coverage was calculated with the corrected GC content for each chromosome. A binary hypothesis T-test and logarithmic likelihood ratio L -score between the two t-tests were used to determine the risk of chromosome aneuploidy. Fetal fraction was calculated in male pregnancies using method described before [21]. Briefly, formula Fetal fraction = $2*\frac{\text{UR}_Y}{\text{UR}}$ was used to calculate the fetal fraction estimate by chromosome Y, where UR_Y represents the unique reads on chromosome Y, and UR represents the unique reads on autosome

Result:

chromosomes.

A total of 68 pregnant women were recruited (Figure S2), including 26 normal pregnant women as healthy controls, 7 pregnant women with fetal trisomies, 7 pregnant women with fetal ACH or TD, 18 pregnant women with ACH/TD-like ultrasound findings, and 10 pregnant women with fetal risk of beta-thalassemia. The mean maternal age was 31.1 years old, and the mean gestational week was 24.62 weeks (Table S4).

Fetal chromosome analysis

Plasma cfDNA of 26 normal pregnancies and 7 pregnancies of fetal trisomies were tested by TAGs-seq and conventional NIPT in parallel, generating about 54 million and 40 million of clean reads per sample, respectively (Table S5). In each sample, about 98.95% of the TAGs-seq reads was distributed on 23 pairs of chromosomes as well as mitochondrial genome, with a mean sequencing depth of 1.66 +- 0.18x (95%CI). The coefficient variance (CV) and chromosome distribution percentage between TAGs-seq and conventional NIPT showed no statistical difference (Figure 2). Using the genome-wide reads of TAGs-seq, we calculated the fetal fraction of each sample, which increased by 20% comparing with the fetal fraction of conventional NIPT (Figure S3). Consistently, reads generated by TAGs-seq contained higher percentage of short cfDNA fragments than conventional NIPT (Figure S3).

Among 33 plasma samples, 3 cases of T21, 2 cases of T18, 2 cases of T13 were correctly identified by TAGsseq using the T-score of 3 as the cut-off value. This was consistent to the results of conventional NIPT (Figure 3A-B) and diagnostic results by amniocentesis (Table S5). Meanwhile, fetal genders deduced using TAGs-seq data were also consistent to that of conventional NIPT (Table S5).

ACH and TD analysis

Plasma cfDNA of 7 pregnant women with fetal ACH or TD, and 18 pregnant women with ACH/TD-like ultrasound findings or pregnant history were tested with TAGs-seq. On average, TAGs-seq generated 44.7 million clean reads, and the mean mapping rate on human genome was 93.0% (Table S6). In each sample, around 98.8% to 99.3% of the TAGs-seq reads were distributed on the whole-genome region with mean sequencing depth of 1.38 + 0.21x (95%CI). Around 0.5% to 1.2% of reads were aligned to the target region of *FGFR3* with 100% of coverage and mean sequencing depth of 55625.30 + 8717.07x (95%CI) (Figure S4A-B). Chromosome analysis using the reads on the whole-genome region showed no samples with T-scores above 3 on chromosome 13, 18, and 21 (Table S6). Hence, all 25 samples were determined as low-risk of fetal trisomy (Figure 3C). Thirteen samples were determined as male pregnancy, and the fetal fraction calculated with TAGs-seq data ranged from 14.3% to 35.9% (Table S6).

The target region reads of TAGs-seq data were used to analyze the allelic frequency of 18 FGFR3 mutations, which identified seven samples with allelic frequency higher than 2%, including five samples of FGFR3: c.1138G>A, one sample of FGFR3:c.1118A>G, and one sample of FGFR3c.742C>T (Figure 4A). No remaining samples were found to contain FGFR3 mutations with the allelic frequency higher than 2% (Figure 4A and Table S6). All TAGs-seq results were consistent with diagnostic results. In 26 normal pregnancies, 0.4% to 1.9% of the TAGs-seq reads were mapped to the FGFR3 gene, and no allelic frequency higher than 2% was detected (Table S5 and Figure 4A).

Analysis of beta-thalassemia using TAGs-seq

Ten pregnancies with high risk of fetal beta-thalassemia were tested by TAGs-seq, each generating 53.9 million of mean clean reads (Table S7). Around 98.22% of the TAGs-seq reads in each sample were distributed on the whole-genome region with mean sequencing depth of 1.99 + 0.68x (95%CI). Chromosome analysis showed no samples with T-score of 3 or above, and thus no fetal aneuploidy of T21, T18, and T13 was detected (Figure 3D). In each case, about 1.12% to 2.49% (mean=1.78%) of clean reads were mapped to the coding region of *HBB* with mean sequencing depth of 42603.88 + 18901.82x (95%CI), and were used to calculate the MAF of mutations (Table S7).

In two samples (H170 and X391), paternal mutations were identified by TAGs-seq in plasma cfDNA, showing 9.81% of c.52A>T in H170 and 2.96% of c.-78A>G, respectively (Figure 4B and Table S7). Both H170 and X391 also contained high MAF of maternal mutations in plasma cfDNA, yet it was impossible to determine if fetus inherited the maternal mutation due to the presence of strong maternal background in plasma cfDNA. Thus, the fetuses of H170 and X391 were only classified as risky of beta-thalassemia. Sanger sequencing of amniotic fluid cells confirmed that the fetus of H170 carried both paternal and maternal mutations, while the fetus of X391 only carried the paternal mutation (Table S7).

Eight samples were detected by TAGs-seq to only contain maternal mutation in plasma cfDNA with the MAF ranging from 17.6% to 46.59% and the fetal fraction from 5.06% to 10.40% (Figure 4B). In total of eight samples, seven (L830, H431, F462, H643, C251, L219, P503) were couples carrying different HBB mutations (Table S7). Hence, the fetuses of these seven pregnancies were determined as low-risk of beta-thalassemia due to the lack of paternal mutations. In Sample K079, both parents carry the same HBB mutation (c.126_129delCTTT), and the possibility of inheriting the paternal mutation in fetus could not be ruled out. Thus, fetal risk of beta-thalassemia could not be determined by TAGs-seq. Sanger sequencing of amniotic fluid in these eight pregnancies confirmed the low risk of beta-thalassemia (Table S7). However, the Sanger sequencing result in F462 revealed the existence of paternal mutation c.-78A>G in fetus, of which the plasma cfDNA sequencing by TAGs-seq identified an MAF of 1.7%, close to but not beyond the 2% cut-off (Figure 4B and Table S7). Thus, although the fetus of F462 was not affected by beta-thalassemia, it contained a paternal mutation missed by TAGs-seq.

Detection performance

All 68 pregnancies were analyzed for fetal aneuploidy using the TAGs-seq method, in which 7 aneuploidies were correctly identified with no false-positive and false-negative results. Thus, the sensitivity and specificity of detecting aneuploidies by TAGs-seq were 100% (56% to 100%, 95%CI) and 100% (93% to 100%, 95%CI), respectively. Fifty-eight pregnancies were analyzed for *FGFR3 de novo* mutations using the TAGs-seq method, and seven positive cases were correctly identified with no false-positive and falsenegative results. Thus, the sensitivity and specificity of detecting *FGFR3 de novo* mutations were 100% (56% to 100%, 95%CI) and 100% (91% to 100%, 95%CI), respectively. In ten pregnancies with fetal risk of beta-thalassemia, three carried the paternal mutations of *HBB*, in which two were correctly identified by TAGs-seq. Thus, the accuracy was 66%.

Discussion:

Main findings

In this study, we demonstrated that TAGs-seq could amplify plasma cfDNA to simultaneously obtain the reads on whole genome and selective regions. In each sample of 68 pregnancies, approximate 98.87% of TAGs-seq reads covered the whole-genome region with sequencing depth of 1.60 + 0.16x (95%CI), while the remaining TAGs-seq reads were mapped to the selected genes with sequencing depth of 9624-149949x (62685.6 + 7293.9x 95%CI). Using confirmed diagnostic results as gold standards, we showed that three cases of T21, two cases of T18, two cases of T13, five cases of ACH, and two cases of TD were correctly identified in the same experiment with no false positive or false negative results. Furthermore, among 10 pregnancies with

fetal risk of beta-thalassemia, TAGs-seq was able to identify the two fetuses carrying paternal mutations on *HBB* at the same time of fetal chromosome analysis. The TAGs-seq NIPT demonstrated the high sensitivity and specificity in detecting fetal aneuploidies and *de novo* mutations. Nonetheless, the TAGs-seq NIPT had one false-negative result of detecting the paternal mutation on HBB, resulting in the 66% of accuracy.

Strengths and limitations

Several techniques have been described to detect single gene diseases in NIPT by using droplet digital PCR [22-27], targeted sequencing [28], COLD-PCR and microarray [29], allele specific real-time PCR [30,31], and cSMART [32,33]. However, these methods require additional experiments not simultaneously coped with chromosome analysis of NIPT. In contrast, TAGs-seq had the strength to accomplish the analysis of both chromosome aneuploidy and single gene diseases without dramatically changing the current NIPT protocol. With low-pass sequencing depth, TAGs-seq amplified plasma cfDNA comparable to conventional NIPT at whole genome level. This reassured consistent performance of chromosome analysis by TAGs-seq NIPT with conventional NIPT. Meanwhile, TAGs-seq NIPT enabled targeted sequencing of multiple genes for single gene disease testing using specific primers. Thus, flexible medical choices could be provided on different clinical scenario. Since no need for additional experiments and deep sequencing, the cost for TAGs-seq NIPT was expected to be close to conventional NIPT.

Another strength of conducting NIPT with TAGs-seq may be the early detection of single gene diseases in fetus. Recent studies have emphasized the importance of prenatal screening for genetically dominant disorders caused by *de novo* or paternally inherited mutations, as they contribute to nearly 60% of severe postnatal monogenic disease [34]. Many single gene diseases, especially those caused by *de novo* mutations, are typically diagnosed at mid-gestation because in lack of family history and asymptomatic until ultrasound examination at 20-22 gestational weeks. Early diagnosis of ACH and TD from other skeletal dysplasia, chromosomal abnormalities, and fetal growth restriction would greatly improve the perinatal management, since different treatments could be performed accordingly [35, 36]. Using the TAGs-seq NIPT, it is possible to identify single gene diseases when screening for chromosome aneuploidy at early pregnancy.

Plasma cfDNA is the mixture of the majority of maternal cfDNA and a small proportion of trophoblastderived cfDNA [37-39]. Because of maternal cfDNA background, our method could not determine the presence of maternal mutation in fetus. Hence, when used for NIPT, TAGs-seq should be restricted to detect *de novo* or paternal mutations of fetus. This makes it impossible to screen for autosome recessive disease by TAGs-seq NIPT in general pregnancies without prior disease history. However, among couples known to carry different alleles of beta-thalassemia, TAGs-seq could be used to noninvasively evaluate the fetal risk by identifying the paternal mutation.

Our study has a small sample size and very few positive cases of disease. This makes it difficult to evaluate the clinical performance. Further assessment is required for proper clinical validation. Nonetheless, as a proof-of-concept study, the data collected from 68 pregnancies demonstrated the stability and repeatability of TAGs-seq in simultaneously generating low-pass whole-genome data and high-depth data of target genes.

Interpretations

The TAGs-seq data obtained from 68 pregnancies were consisted of the low-coverage whole-genome region and high-depth selected genes, which were used for chromosome analysis and allelic mutation analysis, respectively. The TAGs-seq's capability of evenly amplifying whole genome at low coverage was not interfered by the concurrent targeted amplification of the selected genes. This is confirmed by no significant difference of CV and chromosome distribution percentage when comparing the data of TAGs-seq and convention NIPT. As a result, the previously described binary hypothesis t -test and logarithmic likelihood ratio L -score [21] could be used to analyze the TAGs-seq data for chromosome analysis. Interestingly, we observed that the TAGs-seq NIPT led to increased fetal fraction for about 1.2 folds comparing to conventional NIPT, possibly because the linear amplification of multiplex PCR preferably amplified shorter fragments of cfDNA. Since fetal fraction is a key factor to NIPT performance [40, 41], this might bring extra advantages in the situations of low fetal fraction such as early pregnancies or certain aneuploidies [42-44]. Currently, we used the MAF>2% as the cut-off value to call for *de novo* or paternal mutations in plasma cfDNA. However, the cut-off value might be interfered by low fetal fraction, biased amplification preference or sequencing errors. We observed some loci with the frequency close to 2%, which might be caused by biased amplification preference or sequencing errors. Thus, given the circumstance of extremely low fetal fraction or higher sequencing errors, there is the possibility of false-positive or false-negative results using the current MAF cut-off, as being suggested previously [45]. This may be the explanation of the false-negative result of one paternal *HBB* mutation. One way to overcome this problem is defining a MAF 'gray zone' in which samples undergo a retest procedure. Nevertheless, the TAGs-seq NIPT demonstrated very high sensitivities and specificities in detecting fetal aneuploidies and *de novo*mutations causing ACH and TD. Thus, the TAGs-seq demonstrated potentials for clinical use, and further evaluation with larger samples is still required to determine the clinical utility of TAGs-seq NIPT.

Conclusion

We developed a novel NIPT method to simultaneously generate low-coverage genome-wide reads for analyzing chromosome abnormality, and high-depth targeted reads for detecting single gene diseases. With no false positive results, this method correctly identified all the fetal trisomies and mutations on FGFR3 and HBB in plasma cfDNA, showing good potential for clinical use. Further validation is required to evaluate the clinical utility of this method.

Disclosure of interests:

None declared. Completed disclosure of interests form available to view online as supporting information.

Contribution to authorship:

YG, LY and LL designed the study. LY developed the experiment assay. ZH, YL, SP, and LL collected clinical samples, conducted prenatal diagnosis, and analyzed clinical data. YW, LY, HY, SZ, YZ and FC conducted the sample processing for TAGs-seq and sequencing experiments. HZ, QL, DP and YY analyzed the sequencing data and performed statistics analysis and produced the figures for this manuscript. YW and LY wrote the first version of manuscript. HJ, FC and YG oversaw the study and finalized the manuscript.

Details of ethics approval: This study was approved by the Ethics Committee of the Shenzhen People's Hospital (no. 20170710) in Jul. 2017, the Ethics Committee of the PLA General Hospital (no.20151101) in Nov. 2015, the Ethics Committee of the Qingdao Women and Children's Hospital (QFELL-KY-2018-004) in Sept. 2018, and the Institute Review Board of BGI-Shenzhen (BGI-IRB 17151) in Apr. 2018.

Funding: This work was supported by the National Natural Science Foundation of China (No.81601294), Shenzhen Municipal Government of China (JCYJ20170412153136375, JCYJ20180703093402288), and Key-Area Research and Development Program of Guangdong Province (2019B020227001).

Acknowledgements:

We thank all the participants and clinicians involved in this study, without whom the study could not be finished. We appreciate the laboratory assistance by Jia Zhao and Yicong Wang. Thank also goes to Dongsheng Chen for manuscript editing.

Supporting information

Reference

 Lo, Y.M.D., et al., Presence of fetal DNA in maternal plasma and serum. The Lancet, 1997. 350 (9076): p. 485-487.

2. Fan, H.C., et al., Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood. Proc Natl Acad Sci U S A, 2008. **105** (42): p. 16266-71.

3. Chiu, R.W., et al., Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. Proc Natl Acad Sci U S A, 2008. **105** (51): p. 20458-63.

4. Sparks, A.B., et al., Noninvasive prenatal detection and selective analysis of cell-free DNA obtained from maternal blood: evaluation for trisomy 21 and trisomy 18. Am J Obstet Gynecol, 2012.206 (4): p. 319 e1-9.

5. Zimmermann, B., et al., Noninvasive prenatal aneuploidy testing of chromosomes 13, 18, 21, X, and Y, using targeted sequencing of polymorphic loci. Prenat Diagn, 2012. **32** (13): p. 1233-41.

6. Dhallan, R., et al., A non-invasive test for prenatal diagnosis based on fetal DNA present in maternal blood: a preliminary study. The Lancet, 2007. **369** (9560): p. 474-481.

7. Norton, M.E., et al., *Cell-free DNA analysis for noninvasive examination of trisomy*. N Engl J Med, 2015. **372** (17): p. 1589-97.

8. Bianchi, D.W., et al., *DNA sequencing versus standard prenatal aneuploidy screening*. N Engl J Med, 2014. **370** (9): p. 799-808.

9. Lefkowitz, R.B., et al., *Clinical validation of a noninvasive prenatal test for genomewide detection of fetal copy number variants*. Am J Obstet Gynecol, 2016. **215** (2): p. 227 e1-227 e16.

10. Pescia, G., et al., Cell-free DNA testing of an extended range of chromosomal anomalies: clinical experience with 6,388 consecutive cases. Genet Med, 2016.

11. Wapner, R.J., et al., *Expanding the scope of noninvasive prenatal testing: detection of fetal microdeletion syndromes.* Am J Obstet Gynecol, 2015. **212** (3): p. 332 e1-9.

12. Lun, F.M., et al., Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. Proc Natl Acad Sci U S A, 2008. **105** (50): p. 19920-5.

13. Papasavva, T., et al., Prevalence of RhD status and clinical application of non-invasive prenatal determination of fetal RHD in maternal plasma: a 5 year experience in Cyprus. BMC Res Notes, 2016.9 : p. 198.

14. Gu, W., et al., Noninvasive prenatal diagnosis in a fetus at risk for methylmalonic acidemia. Genet Med, 2014. 16 (7): p. 564-7.

15. Hudecova, I., et al., Noninvasive detection of F8 int22h-related inversions and sequence variants in maternal plasma of hemophilia carriers. Blood, 2017. **130** (3): p. 340-347.

16. You, Y., et al., Integration of targeted sequencing and NIPT into clinical practice in a Chinese family with maple syrup urine disease. Genet Med, 2014. **16** (8): p. 594-600.

17. Xu, Y., et al., Haplotype-based approach for noninvasive prenatal tests of Duchenne muscular dystrophy using cell-free fetal DNA in maternal plasma. Genet Med, 2015. **17** (11): p. 889-96.

18. Xu, Y., et al., A new massively parallel nanoball sequencing platform for whole exome research. BMC Bioinformatics, 2019.20 (1): p. 153.

19. Drmanac, R., et al., Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science, 2010.327 (5961): p. 78-81.

20. Huang, J., et al., A reference human genome dataset of the BGISEQ-500 sequencer. Gigascience, 2017. 6 (5): p. 1-9.

21. Jiang, F., et al., Noninvasive Fetal Trisomy (NIFTY) test: an advanced noninvasive prenatal diagnosis methodology for fetal autosomal and sex chromosomal aneuploidies. BMC Med Genomics, 2012. 5 : p. 57.

22. Perlado S, Bustamante-Aragones A, Donas M, Lorda-Sanchez I, PlazaJ, Rodriguez de Alba M. Fetal genotyping in maternal blood by digital PCR: towards NIPD of monogenic disorders independently of parental origin. PLoS One.2016;11(4):e0153258. 23. Lun FM, Tsui NB, Chan KC, Leung TY, Lau TK, Charoenkwan P, et al. Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. Proc Natl Acad Sci U S A.2008;105(50):19920-5.

24. Lam KW, Jiang P, Liao GJ, Chan KC, Leung TY, Chiu RW, et al. Noninvasive prenatal diagnosis of monogenic diseases by targeted massively parallel sequencing of maternal plasma: application to beta-thalassemia. Clin Chem. 2012;58(10):1467-75.

25. Chitty LS, Mason S, Barrett AN, McKay F, Lench N, Daley R, et al. Non-invasive prenatal diagnosis of achondroplasia and thanatophoric dysplasia: next-generation sequencing allows for a safer, more accurate, and comprehensive approach. Prenat Diagn.2015;35(7):656-62.

26. Camunas-Soler J, Lee H, Hudgins L, Hintz SR, Blumenfeld YJ, El-Sayed YY, et al. Noninvasive prenatal diagnosis of single-gene disorders by use of droplet digital PCR. Clin Chem.2018:64(2):336-45.

27. Barrett AN, McDonnell TC, Chan KC, Chitty LS. Digital PCR analysis of maternal plasma for noninvasive detection of sickle cell anemia. Clin Chem. 2012;58(6):1026-32.

28. Zhang J, Li J, Saucier JB, Feng Y, Jiang Y, Sinson J, et al. Non-invasive prenatal sequencing for multiple Mendelian monogenic disorders using circulating cell-free fetal DNA. Nat Med.

29. Galbiati S, Monguzzi A, Damin F, et al. COLD-PCR and microarray: two independent highly sensitive approaches allowing the identification of fetal paternally inherited mutations in maternal plasma. J Med Genet 2016;53:481–487.

30. Zafari M, Gill P, Kowsaryan M, et al. High-resolution melting analysis for noninvasive prenatal diagnosis of IVS-II-I (G-A) fetal DNA in minor betathalassemia mothers. J Matern-Fetal Neonatal Med 2016; 29:3323–3328.

31. Ramezanzadeh M, Salehi M, Farajzadegan Z, et al. Detection of paternally inherited fetal point mutations for b-thalassemia in maternal plasma using simple fetal DNA enrichment protocol with or without whole genome amplification: an accuracy assessment. J Matern-Fetal Neonatal Med 2016;29:2645–2649.

32. Lv W, Wei X, Guo R, et al. Noninvasive prenatal testing for Wilson disease by use of circulating single-molecule amplification and resequencing technology (cSMART). Clin Chem 2015; 61:172–181.

33. Lv W, Li Z, Wei X, et al. Noninvasive fetal genotyping in pregnancies at risk for PKU using a comprehensive quantitative cSMART assay for PAH gene mutations: a clinical feasibility study[J]. BJOG: An International Journal of Obstetrics & Gynaecology, 2019, 126(12): 1466-1474.

34. Zhang J, Li J, Saucier JB, Feng Y, Jiang Y, Sinson J, et al. Non-invasive prenatal sequencing for multiple Mendelian monogenic disorders using circulating cell-free fetal DNA. Nat Med. 2019;25:701–2.

35. Horton, W.A., J.G. Hall, and J.T. Hecht, Achondroplasia.Lancet, 2007. 370 (9582): p. 162-172.

36. Chitty, L.S., et al., Non-invasive prenatal diagnosis of achondroplasia and thanatophoric dysplasia: nextgeneration sequencing allows for a safer, more accurate, and comprehensive approach. Prenat Diagn, 2015. **35** (7): p. 656-62.

37. Lo, Y.M.D. et al. 1998. Quantitative analysis of fetal DNA in maternal plasma and serum: implications for noninvasive prenataldiagnosis. Am.J.Hum.Genet. 62: 768–775.

38. 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: 913-920.

39. Sparks AB, Struble CA, Wang ET, et al. Noninvasive prenatal detection and selective analysis of cell-free DNA obtained from maternal blood: evaluation for trisomy 21 and trisomy 18. Am J Obstet Gynecol 2012; 206: 319 e311-319.

40. Palomaki, G.E. and E.M. Kloza, Prenatal cell-free DNA screening test failures: a systematic review of failure rates, risks of Down syndrome, and impact of repeat testing. Genet Med, 2018.20 (11): p. 1312-1323.

41. Benn, P., et al., Factors Associated With Informative Redraw After an Initial No Result in Noninvasive Prenatal Testing. Obstet Gynecol, 2018. **132** (2): p. 428-435.

42. Norton ME, Jacobsson B, Swamy GK, et al. Cell-free DNA analysis for noninvasive examination of trisomy. N Engl J Med. 2015 Apr 23;372(17):1589–97.

43. Zhou, Y., Zhu, Z., Gao, Y., Yuan, Y., Guo, Y., Zhou, L., ... & Chen, Z. (2015). Effects of maternal and fetal characteristics on cell-free fetal DNA fraction in maternal plasma. Reproductive Sciences, 22(11), 1429-1435.

44. Canick JA, Palomaki GE, Kloza EM, Lambert-Messerlian GM, Haddow JE. The impact of maternal plasma DNA fetal fraction on next generation sequencing tests for common fetal aneuploidies. Prenat Diagn. 2013;33(7):667–674. doi:10.1002/pd.4126.

45. Xiong, L., et al., Non-invasive prenatal diagnostic testing for beta-thalassaemia using cell-free fetal DNA and next generation sequencing. Prenat Diagn, 2015. **35** (3): p. 258-65.



Figure 1. A schematic illustration of simultaneous amplification of whole genome and target regions of cfDNAby TAGs-seq. BC, barcode; GSP1, gene specific primer1; GSP2, gene specific primer2; UP1, universal primer1; UP2, universal primer2.



Figure 2. Comparison of the whole-genome reads obtained by TAGs-seq to conventional NIPT. (A) comparison of genome distribution on each chromosome (1 to 22, X, Y) and mitochondria genome (M) between TAGs-seq (red) and conventional NIPT (blue); (B) coefficient of variation (CV) of sequencing reads on different chromosomes between TAGs-seq (purple bars) to conventional NIPT























